PURE and APPLIED ANALYSIS

PAA

msp

vol. 2 no. 2 2020

PURE and APPLIED
ANALYSIS

msp.org/paa

Cover image: The figure shows the outgoing scattered field produced by scattering a plane wave, coming from the northwest, off of the (stylized) letters P A A. The total field satisfies the homogeneous Dirichlet condition on the boundary of the letters. It is based on a numerical computation by Mike O'Neil of the Courant Institute.

See inside back cover or msp.org/paa for submission instructions.

The subscription price for 2020 is US $505/year for the electronic version, and $565/year (+$25, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Pure and Applied Analysis (ISSN 2578-5885 electronic, 2578-5893 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

PAA peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY
mathematical sciences publishers
nonprofit scientific publishing

PURE and APPLIED
ANALYSIS

Vol. 2, No. 2, 2020
https://doi.org/10.2140/paa.2020.2.203

HYPOCOERCIVITY WITHOUT CONFINEMENT

EMERIC BOUIN, JEAN DOLBEAULT, STÉPHANE MISCHLER,
CLÉMENT MOUHOT AND CHRISTIAN SCHMEISER

Hypocoercivity methods are applied to linear kinetic equations with mass conservation and without confinement in order to prove that the solutions have an algebraic decay rate in the long-time range, which the same as the rate of the heat equation. Two alternative approaches are developed: an analysis based on decoupled Fourier modes and a direct approach where, instead of the Poincaré inequality for the Dirichlet form, Nash's inequality is employed. The first approach is also used to provide a simple proof of exponential decay to equilibrium on the flat torus. The results are obtained on a space with exponential weights and then extended to larger function spaces by a factorization method. The optimality of the rates is discussed. Algebraic rates of decay on the whole space are improved when the initial datum has moment cancellations.

1. Introduction

We consider the Cauchy problem

$$\partial_t f + v \cdot \nabla_x f = \mathsf{L} f, \quad f(0,x,v) = f_0(x,v) \tag{1}$$

for a distribution function $f(t,x,v)$, with *position* variable $x \in \mathbb{R}^d$, *velocity* variable $v \in \mathbb{R}^d$, and with *time* $t \geq 0$. Concerning the *collision operator* L, we shall consider two cases:

(a) *Fokker–Planck* collision operator:

$$\mathsf{L} f = \nabla_v \cdot [M \, \nabla_v (M^{-1} f)].$$

(b) *Scattering* collision operator:

$$\mathsf{L} f = \int_{\mathbb{R}^d} \sigma(\,\cdot\,, v')(f(v')M(\,\cdot\,) - f(\,\cdot\,)M(v'))\, dv'.$$

We shall make the following assumptions on the *local equilibrium* $M(v)$ and on the *scattering rate* $\sigma(v,v')$:

$$\int_{\mathbb{R}^d} M(v)\,dv = 1, \quad \nabla_v \sqrt{M} \in \mathrm{L}^2(\mathbb{R}^d), \quad M \in C(\mathbb{R}^d),$$
$$M = M(|v|), \quad 0 < M(v) \leq c_1 e^{-c_2|v|} \quad \text{for all } v \in \mathbb{R}^d, \quad \text{for some } c_1, c_2 > 0. \tag{H1}$$

MSC2010: primary 82C40; secondary 76P05, 35H10, 35K65, 35P15, 35Q84.
Keywords: hypocoercivity, linear kinetic equations, Fokker–Planck operator, scattering operator, transport operator, Fourier mode decomposition, Nash's inequality, factorization method, Green's function, micro-/macrodecomposition, diffusion limit.

$$1 \le \sigma(v, v') \le \bar{\sigma} \quad \text{for all } v, v' \in \mathbb{R}^d, \quad \text{for some } \bar{\sigma} \ge 1. \tag{H2}$$

$$\int_{\mathbb{R}^d} (\sigma(v, v') - \sigma(v', v)) M(v')\, dv' = 0 \quad \text{for all } v \in \mathbb{R}^d. \tag{H3}$$

Before stating our main results, let us list some preliminary observations.

(i) A typical example of a *local equilibrium* satisfying (H1) is the Gaussian

$$M(v) = \frac{e^{-\frac{|v|^2}{2}}}{(2\pi)^{\frac{d}{2}}}. \tag{2}$$

(ii) With $\sigma \equiv 1$, Case (b) includes the relaxation operator $\mathsf{L} f = M\rho_f - f$, also known as the *linear BGK operator*, with position density defined by

$$\rho_f(t, x) := \int_{\mathbb{R}^d} f(t, x, v)\, dv.$$

(iii) Positivity and exponential decay of the local equilibrium are essential for our approach. The assumption on the gradient and continuity are technical and only needed for some of our results. Rotational symmetry is not important, but assumed for computational convenience. However the property

$$\int_{\mathbb{R}^d} v M(v)\, dv = 0,$$

i.e., *zero flux in local equilibrium*, is essential.

(iv) Since microreversibility (or detailed balance), i.e., symmetry of σ, is not required, assumption (H3) is needed for *mass conservation*, i.e.,

$$\int_{\mathbb{R}^d} \mathsf{L} f\, dv = 0,$$

in Case (b). The boundedness away from zero of σ in (H2) guarantees coercivity of L relative to its null space (such bound can always be written $\sigma \ge 1$ by scaling).

Since $e^{t\mathsf{L}}$ propagates probability densities, i.e., conserves mass and nonnegativity, L dissipates convex relative entropies, implying in particular

$$\int_{\mathbb{R}^d} \mathsf{L} f \frac{f}{M}\, dv \le 0.$$

This suggests to use the L^2-space with the measure $d\gamma_\infty := \gamma_\infty\, dv$, where $\gamma_\infty(v) = M(v)^{-1}$, as a functional-analytic framework (the subscript ∞ will make sense later). We shall need the *microscopic coercivity* property

$$-\int_{\mathbb{R}^d} f\, \mathsf{L} f\, d\gamma_\infty \ge \lambda_m \int_{\mathbb{R}^d} (f - M\rho_f)^2\, d\gamma_\infty, \tag{H4}$$

with some $\lambda_m > 0$. In Case (a) it is equivalent to the Poincaré inequality with weight M,

$$\int_{\mathbb{R}^d} |\nabla_v h|^2 M\, dv \ge \lambda_m \int_{\mathbb{R}^d} \left(h - \int_{\mathbb{R}^d} h M\, dv \right)^2 M\, dv$$

for all $h = f/M \in \mathrm{H}^1(M\,dv)$. It holds as a consequence of the exponential decay assumption in (H1); see, e.g., [Nash 1958; Bakry et al. 2008]. For the normalized Gaussian (2) the optimal constant is known to be $\lambda_m = 1$; see for instance [Beckner 1989]. In Case (b), (H4) means

$$\frac{1}{2}\iint_{\mathbb{R}^d\times\mathbb{R}^d} \sigma(v,v')M(v)M(v')(u(v)-u(v'))^2 dv'\,dv \ge \lambda_m \int_{\mathbb{R}^d} (u-\rho_{uM})^2 M\,dv$$

for all $u = f/M \in \mathrm{L}^2(M\,dv)$, and it holds with $\lambda_m = 1$ as a consequence of the lower bound for σ in assumption (H2).

Although the transport operator does not contribute to entropy dissipation, its dispersion in the x-direction in combination with the dissipative properties of the collision operator yields the desired decay results. In order to perform a *mode-by-mode hypocoercivity* analysis, we introduce the Fourier representation with respect to x,

$$f(t,x,v) = \int_{\mathbb{R}^d} \hat{f}(t,\xi,v)e^{+ix\cdot\xi}\,d\mu(\xi),$$

where $d\mu(\xi) = (2\pi)^{-d}\,d\xi$ and $d\xi$ is the Lebesgue measure on $\mathbb{R}^d$. The normalization of $d\mu(\xi)$ is chosen such that Plancherel's formula reads

$$\|f(t,\cdot,v)\|_{\mathrm{L}^2(dx)} = \|\hat{f}(t,\cdot,v)\|_{\mathrm{L}^2(d\mu(\xi))},$$

with a straightforward abuse of notation. The Cauchy problem (1) in Fourier variables is now decoupled in the ξ-direction:

$$\partial_t \hat{f} + i(v\cdot\xi)\hat{f} = \mathsf{L}\hat{f}, \quad \hat{f}(0,\xi,v) = \hat{f}_0(\xi,v). \tag{3}$$

Our main results are devoted to *hypocoercivity without confinement*: when the variable x is taken in $\mathbb{R}^d$, we assume that there is no potential preventing the runaway corresponding to $|x| \to +\infty$. So far, hypocoercivity results have been obtained either in the compact case corresponding to a bounded domain in x, for instance $\mathbb{T}^d$, or in the whole Euclidean space with an external potential V such that the measure $e^{-V}\,dx$ admits a Poincaré inequality. Usually other technical assumptions are required on V and there are many variants (for instance one can assume a stronger logarithmic Sobolev inequality instead of a Poincaré inequality), but the common property is that some growth condition on V is assumed and in particular the measure $e^{-V}\,dx$ is bounded. Here we consider the case $V \equiv 0$, which is obviously a different regime. By replacing the Poincaré inequality by Nash's inequality or using direct estimates in Fourier variables, we adapt the L^2 hypocoercivity methods and prove that an appropriate norm of the solution decays at a rate which is the rate of the heat equation. This observation is compatible with diffusion limits, which have been a source of inspiration for building Lyapunov functionals and establishing the L^2 hypocoercivity method of [Dolbeault et al. 2015]. Before stating any results, we need some notation to implement the *factorization* method of [Gualdani et al. 2017] and obtain estimates in large functional spaces.

Let us consider the measures

$$d\gamma_k := \gamma_k(v)\,dv, \quad \text{where} \quad \gamma_k(v) = (1+|v|^2)^{\frac{k}{2}} \quad \text{and} \quad k > d, \tag{4}$$

such that $1/\gamma_k \in \mathrm{L}^1(\mathbb{R}^d)$. The condition $k \in (d,\infty]$ then covers the case of weights with a growth of the order of $|v|^k$, when k is finite, and we denote by $k=\infty$ the case when the weight $\gamma_\infty = M^{-1}$ grows at least exponentially fast.

Theorem 1. *Assume (H1)–(H4), $x \in \mathbb{R}^d$, and $k \in (d,\infty]$. Then there exists a constant $C>0$ such that solutions f of (1) with initial datum $f_0 \in \mathrm{L}^2(dx\,d\gamma_k) \cap \mathrm{L}^2(d\gamma_k;\mathrm{L}^1(dx))$ satisfy, for all $t \ge 0$,*

$$\|f(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_k)} \le C\,\frac{\|f_0\|^2_{\mathrm{L}^2(dx\,d\gamma_k)} + \|f_0\|^2_{\mathrm{L}^2(d\gamma_k;\mathrm{L}^1(dx))}}{(1+t)^{\frac{d}{2}}}.$$

For the heat equation improved decay rates can be shown by Fourier techniques, if the modes with slowest decay are eliminated from the initial data. The following two results are in this spirit.

Theorem 2. *Let the assumptions of Theorem 1 hold, and let*

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} f_0\,dx\,dv = 0.$$

Then there exists $C>0$ such that solutions f of (1) with initial datum f_0 satisfy, for all $t \ge 0$,

$$\|f(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_k)} \le C\,\frac{\|f_0\|^2_{\mathrm{L}^2(d\gamma_{k+2};\mathrm{L}^1(dx))} + \|f_0\|^2_{\mathrm{L}^2(d\gamma_k;\mathrm{L}^1(|x|\,dx))} + \|f_0\|^2_{\mathrm{L}^2(dx\,d\gamma_k)}}{(1+t)^{\frac{d}{2}+1}},$$

with $k \in (d,\infty)$.

The case of Theorem 2, but with $k=\infty$, is covered in Theorem 3 under the stronger assumption that M is a Gaussian. For the formulation of a result corresponding to the cancellation of higher-order moments, we introduce the set $\mathbb{R}_\ell[X,V]$ of polynomials of order at most ℓ in the variables $X, V \in \mathbb{R}^d$ (the sum of the degrees in X and in V is at most ℓ). We also need that the kernel of the collision operator is spanned by a Gaussian function in order to keep polynomial spaces invariant. This means that for any $P \in \mathbb{R}_\ell[X,V]$, one has $(\mathsf{L}-\mathsf{T})(PM) \in \mathbb{R}_\ell[X,V]M$. Since the transport operator mixes both variables x and v, one needs moments with respect to both x and v variables.

Theorem 3. *In Case* (a), *let M be the normalized Gaussian (2). In Case* (b), *we assume that $\sigma \equiv 1$. Let $k \in (d,\infty]$, $\ell \in \mathbb{N}$, and assume that the initial datum $f_0 \in \mathrm{L}^1(\mathbb{R}^d\times\mathbb{R}^d)$ is such that*

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} f_0(x,v)\,P(x,v)\,dx\,dv = 0 \tag{5}$$

for all $P \in \mathbb{R}_\ell[X,V]$. Then there exists a constant $c_k>0$ such that any solution f of (1) with initial datum f_0 satisfies, for all $t \ge 0$,

$$\|f(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_k)} \le c_k\,\frac{\|f_0\|^2_{\mathrm{L}^2(d\gamma_{k+2};\mathrm{L}^1(dx))} + \|f_0\|^2_{\mathrm{L}^2(d\gamma_k;\mathrm{L}^1(|x|\,dx))} + \|f_0\|^2_{\mathrm{L}^2(dx\,d\gamma_k)}}{(1+t)^{\frac{d}{2}+1+\ell}}.$$

The *outline of this paper* goes as follows. In Section 2, we slightly strengthen the *abstract hypocoercivity* result of [Dolbeault et al. 2015] by allowing complex Hilbert spaces and by providing explicit formulas for the coefficients in the decay rate (Proposition 4). In Corollary 5, this result is applied for fixed ξ to the Fourier transformed problem (3), where integrals are computed with respect to the measure $d\gamma_\infty$ in the velocity variable v. Since the frequency ξ can be considered as a parameter, we shall speak of a *mode-by-mode hypocoercivity* result. It provides exponential decay, however, with a rate deteriorating as $\xi \to 0$.

In Section 3, we state a special case (Proposition 6) of the *factorization* result of [Gualdani et al. 2017] with explicit constants which corresponds to an *enlargement* of the space, and also a *shrinking* result (Proposition 7) which will be useful in Section 6.2. By the enlargement result, the estimate corresponding to the exponential weight γ_∞ is extended in Corollary 8 to larger spaces corresponding to the algebraic weights γ_k with $k \in (d, \infty)$. As a straightforward consequence, in Section 4, we recover an *exponential convergence rate* in the case of the flat torus $\mathbb{T}^d$ (Corollary 9), and then give a first proof of the *algebraic decay rate* of Theorem 1 in the whole space without confinement.

In Section 5, a hypocoercivity method, where the Poincaré inequality, or the so-called *macroscopic coercivity* condition, is replaced by the *Nash inequality*, provides an alternative proof of Theorem 1. Such a direct approach is also applicable to problems with nonconstant coefficients like scattering operators with x-dependent scattering rates σ, or Fokker–Planck operators with x-dependent diffusion constants like $\nabla_v \cdot (\mathscr{D}(x) M \nabla_v (M^{-1} f))$.

The *improved algebraic decay rates* of Theorems 2 and 3 are obtained by direct Fourier estimates in Section 6. As we shall see in Appendix A, the rates of Theorem 1 are optimal: the decay rate is the rate of the heat equation on $\mathbb{R}^d$. Our method is consistent with the *diffusion limit* and provides estimates which are asymptotically uniform in this regime: see Appendix B. We also check that the results of Theorems 2 and 3 are uniform in the diffusive limit in Appendix B.

We conclude this introduction by a brief *review of the literature*: On the whole Euclidean space, we refer to [Vázquez 2017] for recent lecture notes on available techniques for capturing the large-time asymptotics of the heat equation. Some of our results make a clear link with the heat flow seen as the diffusion limit of the kinetic equation. We also refer to [Iacobucci et al. 2019] for recent results on the diffusion limit, or overdamped limit (see Appendix B).

The mode-by-mode analysis is an extension of the hypocoercivity theory of [Dolbeault et al. 2015], which has been inspired by [Hérau 2006], but is also close to the Kawashima compensating function method [1990]; see also [Glassey 1996, Chapter 3, Section 3.9]. We also refer to [Duan 2011] where the Kawashima approach is applied to the Fokker–Planck operator (a) and to a particular case of the scattering model (b).

The word *hypocoercivity* was coined by T. Gallay and widely disseminated in the context of kinetic theory by C. Villani. In [Mouhot and Neumann 2006; Villani 2006; 2009], the method deals with large-time properties of the solutions by considering an H^1-norm (in x and v variables) and taking into account cross-terms. This is very well explained in [Villani 2006, Section 3], but was already present in earlier works like [Hérau and Nier 2004]. Hypocoercivity theory is inspired by and related to the earlier *hypoellipticity* theory. The latter has a long history in the context of the kinetic Fokker–Planck equation. One can refer for instance to [Eckmann and Hairer 2003; Hérau and Nier 2004] and much earlier to Hörmander's theory [1967]. The seed for such an approach can even be traced back to Kolmogorov's computation [1934] of Green's kernel for the kinetic Fokker–Planck equation, which has been reconsidered in [Ilin and Hasminskii 1964] and successfully applied, for instance, to the study of the Vlasov–Poisson–Fokker–Planck system in [Victory and O'Dwyer 1990; Bouchut 1993].

Linear Boltzmann equations and BGK (Bhatnagar–Gross–Krook, see [Bhatnagar et al. 1954]) models also have a long history: we refer to [Degond et al. 2000; Cáceres et al. 2003] for key mathematical properties, and to [Mouhot and Neumann 2006; Hérau 2006] for first hypocoercivity results. In this paper we will mostly rely on [Dolbeault et al. 2009; 2015]. However, among more recent contributions, one has to quote [Han-Kwan and Léautaud 2015; Achleitner et al. 2016; Bouin et al. 2017] and also an approach based on the Fisher information which has recently been implemented in [Evans 2017; Monmarché 2017].

With the *exponential weight* $\gamma_\infty = M^{-1}$, Corollary 9 can be obtained directly by the method of [Dolbeault et al. 2015]. In this paper we also obtain a result for weights with polynomial growth in the velocity variable based on [Gualdani et al. 2017]. For completeness, let us mention that recently the exponential growth issue was overcome for the Fokker–Planck case in [Kavian and Mischler 2015; Mischler and Mouhot 2016] by a different method. The improved decay rates established in Theorems 2 and 3 generalize to kinetic models similar results known for the heat equation; see for instance [Mischler and Mouhot 2016, Remark 3.2(7)] or [Bartier et al. 2011].

2. Mode-by-mode hypocoercivity

Let us consider the evolution equation

$$\frac{dF}{dt} + \mathsf{T}F = \mathsf{L}F, \tag{6}$$

where T and L are respectively a general *transport operator* and a general *linear collision operator*. We shall use the abstract approach of [Dolbeault et al. 2015]. Although the extension of the method to Hilbert spaces over complex numbers is rather straightforward, we carry it out here for completeness. For details on the Cauchy problem or, e.g., on the domains of the operators, we refer to [Dolbeault et al. 2015]. Notice that we do not ask that L is a Hermitian operator but simply assume that $\mathsf{L}^*\mathsf{A} = 0$.

Proposition 4. *Let L and T be closed unbounded linear operators on the complex Hilbert space $(\mathscr{H}, \langle\cdot,\cdot\rangle)$ with dense domains $\mathscr{D}(L)$ and $\mathscr{D}(T)$. Assume that T is anti-Hermitian. Let Π be the orthogonal projection onto the null space of L and define*

$$\mathsf{A} := (1 + (\mathsf{T}\Pi)^*\mathsf{T}\Pi)^{-1}(\mathsf{T}\Pi)^*,$$

where * denotes the adjoint with respect to $\langle\cdot,\cdot\rangle$. We assume that $\mathsf{L}^*\mathsf{A} = 0$ and that there are positive constants λ_m, λ_M, and C_M such that, for any $F \in \mathscr{H}$, the following properties hold:

- Microscopic coercivity:

$$-\langle \mathsf{L}F, F\rangle \ge \lambda_m \|(1-\Pi)F\|^2 \quad \textit{for all } F \in \mathscr{D}(L). \tag{A1}$$

- Macroscopic coercivity:

$$\|\mathsf{T}\Pi F\|^2 \ge \lambda_M \|\Pi F\|^2 \quad \textit{for all } F \in \mathscr{D}(T). \tag{A2}$$

- Parabolic macroscopic dynamics:

$$\Pi \mathsf{T} \Pi F = 0 \quad \textit{for all } F \in \mathscr{D}(T). \tag{A3}$$

- Bounded auxiliary operators:

$$\|\mathsf{AT}(1-\Pi)F\| + \|\mathsf{AL}F\| \le C_M \|(1-\Pi)F\| \quad \textit{for all } F \in \mathscr{D}(L) \cap \mathscr{D}(T). \tag{A4}$$

Then $L - T$ generates a C_0-semigroup and for any $t \ge 0$ we have

$$\|e^{(\mathsf{L}-\mathsf{T})t}\|^2 \le 3e^{-\lambda t}, \quad \textit{where } \lambda = \frac{\lambda_M}{3(1+\lambda_M)} \min\left\{1, \lambda_m, \frac{\lambda_m \lambda_M}{(1+\lambda_M)C_M^2}\right\}. \tag{7}$$

Proof. For some $\delta > 0$ to be determined later, the Lyapunov functional

$$\mathsf{H}[F] := \tfrac{1}{2}\|F\|^2 + \delta \operatorname{Re}\langle \mathsf{A}F, F\rangle$$

is such that $\frac{d}{dt}\mathsf{H}[F] = -\mathsf{D}[F]$ if F solves (6), with

$$\mathsf{D}[F] := -\langle \mathsf{L}F, F\rangle + \delta\langle \mathsf{AT}\Pi F, F\rangle + \delta \operatorname{Re}\langle \mathsf{AT}(1-\Pi)F, F\rangle - \delta \operatorname{Re}\langle \mathsf{TA}F, F\rangle - \delta \operatorname{Re}\langle \mathsf{AL}F, F\rangle.$$

Note that we have used the fact that $\operatorname{Re}\langle \mathsf{A}F, \mathsf{L}F\rangle = 0$ because of the assumption $\mathsf{L}^*\mathsf{A} = 0$, and also that $\langle \mathsf{AT}\Pi F, F\rangle$ is real because $\mathsf{AT}\Pi$ is self-adjoint by construction. Since the Hermitian operator $\mathsf{AT}\Pi$ can be interpreted as the application of the map $z \mapsto (1+z)^{-1}z$ to $(\mathsf{T}\Pi)^*\mathsf{T}\Pi$ and as a consequence of the spectral theorem [Reed and Simon 1980, Theorem VII.2, p. 225], the conditions (A1) and (A2) imply

$$-\langle \mathsf{L}F, F\rangle + \delta\langle \mathsf{AT}\Pi F, F\rangle \ge \lambda_m \|(1-\Pi)F\|^2 + \frac{\delta\lambda_M}{1+\lambda_M}\|\Pi F\|^2.$$

As in [Dolbeault et al. 2015, Lemma 1], if $G = \mathsf{A}F$, i.e., $G + (\mathsf{T}\Pi)^*\mathsf{T}\Pi G = (\mathsf{T}\Pi)^* F$, one has

$$\|\mathsf{A}F\|^2 + \|\mathsf{TA}F\|^2 = \langle G, G + (\mathsf{T}\Pi)^*\mathsf{T}\Pi G\rangle = \langle G, (\mathsf{T}\Pi)^* F\rangle = \langle \mathsf{TA}F, (1-\Pi)F\rangle,$$

where we have used $A = \Pi A$ and $\Pi T \Pi = 0$. Using $|\langle \mathsf{TA}F, (1-\Pi)F\rangle| \le \|\mathsf{TA}F\|^2 + \frac{1}{4}\|(1-\Pi)F\|^2$, one gets

$$\|\mathsf{A}F\|^2 \le \tfrac{1}{4}\|(1-\Pi)F\|^2, \tag{8}$$

which implies that $|\operatorname{Re}\langle \mathsf{A}F, F\rangle| \le \|AF\|\|F\| \le \frac{1}{2}\|F\|^2$ and provides us with the norm equivalence of $\mathsf{H}[F]$ and $\|F\|^2$,

$$\tfrac{1}{2}(1-\delta)\|F\|^2 \le \mathsf{H}[F] \le \tfrac{1}{2}(1+\delta)\|F\|^2. \tag{9}$$

With $X := \|(1-\Pi)F\|$ and $Y := \|\Pi F\|$, it follows from (A4) that

$$\mathsf{D}[F] \ge (\lambda_m - \delta)X^2 + \frac{\delta\lambda_M}{1+\lambda_M}Y^2 - \delta C_M XY.$$

The choice

$$\delta = \frac{1}{2}\min\left\{1, \lambda_m, \frac{\lambda_m\lambda_M}{(1+\lambda_M)C_M^2}\right\}$$

implies

$$\mathsf{D}[F] \geq \frac{\lambda_m}{4} X^2 + \frac{\delta \lambda_M}{2(1+\lambda_M)} Y^2 \geq \frac{1}{4} \min\left\{\lambda_m, \frac{2\delta\lambda_M}{1+\lambda_M}\right\} \|F\|^2 \geq \frac{2\delta\lambda_M}{3(1+\lambda_M)} \mathsf{H}[F].$$

With λ defined in (7), using $\delta \leq \frac{1}{2}$ and $(1+\delta)/(1-\delta) \leq 3$, we get

$$\|F(t)\|^2 \leq \frac{2}{1-\delta} \mathsf{H}[F](t) \leq \frac{1+\delta}{1-\delta} e^{-\lambda t} \|F(0)\|^2 \leq 3 e^{-\lambda t} \|F(0)\|^2. \qquad \square$$

For any fixed $\xi \in \mathbb{R}^d$, let us apply Proposition 4 to (3) with $F = \hat{f}$ and

$$\mathscr{H} = \mathrm{L}^2(d\gamma_\infty), \quad \|F\|^2 = \int_{\mathbb{R}^d} |F|^2 \, d\gamma_\infty, \quad \Pi F = M \int_{\mathbb{R}^d} F \, dv = M \rho_F, \quad \mathsf{T} F = i(v \cdot \xi) F.$$

Here we are in a mode-by-mode framework in which the transport operator T is a simple multiplication operator.

Corollary 5. *Assume (H1)–(H4), and take $\xi \in \mathbb{R}^d$. If $\hat{f}$ is a solution of (3) such that $\hat{f}_0(\xi,\cdot) \in \mathrm{L}^2(d\gamma_\infty)$, then for any $t \geq 0$ we have*

$$\|\hat{f}(t,\xi,\cdot)\|^2_{\mathrm{L}^2(d\gamma_\infty)} \leq 3 e^{-\mu_\xi t} \|\hat{f}_0(\xi,\cdot)\|^2_{\mathrm{L}^2(d\gamma_\infty)},$$

where

$$\mu_\xi := \frac{\Lambda |\xi|^2}{1+|\xi|^2} \quad \textit{and} \quad \Lambda = \frac{1}{3} \min\{1,\Theta\} \min\left\{1, \frac{\lambda_m \Theta^2}{K + \Theta \kappa^2}\right\}, \tag{10}$$

with

$$\Theta := \int_{\mathbb{R}^d} (v\cdot \mathrm{e})^2 M(v)\, dv, \quad K := \int_{\mathbb{R}^d} (v\cdot \mathrm{e})^4 M(v)\, dv, \quad \theta := \frac{4}{d} \int_{\mathbb{R}^d} |\nabla_v \sqrt{M}|^2 \, dv \tag{11}$$

for an arbitrary $\mathrm{e} \in \mathbb{S}^{d-1}$, and with $\kappa = \sqrt{\theta}$ in Case (a) *and $\kappa = 2\bar{\sigma}\sqrt{\Theta}$ in Case* (b).

Proof. We check that the assumptions of Proposition 4 are satisfied with $F = \hat{f}$. The property $\mathsf{L}^* \mathsf{A} = 0$ is a consequence of the mass conservation $\int_{\mathbb{R}^d} \mathsf{L} f \, dv = 0$ because $\Pi \mathsf{A} = \mathsf{A}$. Assumption (H4) implies (A1). Concerning the macroscopic coercivity (A2), since

$$\mathsf{T}\Pi F = i(v\cdot\xi) \rho_F M,$$

one has

$$\|\mathsf{T}\Pi F\|^2 = |\rho_F|^2 \int_{\mathbb{R}^d} |v\cdot\xi|^2 M(v)\, dv = \Theta |\xi|^2 |\rho_F|^2 = \Theta |\xi|^2 \|\Pi F\|^2,$$

and thus (A2) holds with $\lambda_M = \Theta|\xi|^2$. By assumption $M(v)$ depends only on $|v|$, so it is *unbiased*, i.e., $\int_{\mathbb{R}^d} v M(v)\, dv = 0$, which means that (A3) holds.

Let us now prove (A4). Since $(\mathsf{T}\Pi)^* F = -\Pi \mathsf{T} F = -i\big(\xi \int_{\mathbb{R}^d} v' F(v')\, dv'\big) M$, we obtain

$$(1 + (\mathsf{T}\Pi)^* \mathsf{T}\Pi) \rho M = \left(1 + \int_{\mathbb{R}^d} (\xi \cdot v')^2 M(v')\, dv'\right) \rho M = (1 + \Theta|\xi|^2) \rho M$$

and the operator A, defined in Proposition 4, is given mode-by-mode by

$$\mathsf{A}F = \frac{-i\xi \int_{\mathbb{R}^d} v' F(v')\,dv'}{1+\Theta|\xi|^2} M.$$

As a consequence, A satisfies the estimate

$$\begin{aligned}\|\mathsf{A}F\| = \|\mathsf{A}(1-\Pi)F\| &\le \frac{1}{1+\Theta|\xi|^2}\int_{\mathbb{R}^d} \frac{|(1-\Pi)F|}{\sqrt{M}} |v\cdot\xi|\sqrt{M}\,dv \\ &\le \frac{\|(1-\Pi)F\|}{1+\Theta|\xi|^2}\left(\int_{\mathbb{R}^d}(v\cdot\xi)^2 M\,dv\right)^{\frac12} = \frac{\sqrt{\Theta}|\xi|}{1+\Theta|\xi|^2}\|(1-\Pi)F\|.\end{aligned}$$

In Case (b) the collision operator L is obviously bounded,

$$\|\mathsf{L}F\| \le 2\bar{\sigma}\|(1-\Pi)F\|$$

and, as a consequence,

$$\|\mathsf{AL}F\| \le \frac{2\bar{\sigma}\sqrt{\Theta}|\xi|}{1+\Theta|\xi|^2}\|(1-\Pi)F\|.$$

We also notice that $\mathsf{L}^*\mathsf{A} = 0$ according to (H3). For estimating AL in Case (a), we note that

$$\int_{\mathbb{R}^d} v\mathsf{L}F\,dv = 2\int_{\mathbb{R}^d} \nabla_v\sqrt{M}\,\frac{F}{\sqrt{M}}\,dv$$

and obtain as above that

$$\|\mathsf{AL}F\| \le \frac{2}{1+\Theta|\xi|^2}\int_{\mathbb{R}^d}\frac{|(1-\Pi)F|}{\sqrt{M}}\,|\xi\cdot\nabla_v\sqrt{M}|\,dv \le \frac{\sqrt{\theta}|\xi|}{1+\Theta|\xi|^2}\|(1-\Pi)F\|.$$

For both cases we finally obtain

$$\|\mathsf{AL}F\| \le \frac{\kappa|\xi|}{1+\Theta|\xi|^2}\|(1-\Pi)F\|.$$

Similarly we can estimate

$$\mathsf{AT}(1-\Pi)F = \frac{\int_{\mathbb{R}^d}(v'\cdot\xi)^2(1-\Pi)F(v')\,dv'}{1+\Theta|\xi|^2}M$$

by

$$\begin{aligned}\|\mathsf{AT}(1-\Pi)F\| &= \frac{\left|\int_{\mathbb{R}^d}(v'\cdot\xi)^2(1-\Pi)F(v')\,dv'\right|}{1+\Theta|\xi|^2} \\ &\le \frac{\left(\int_{\mathbb{R}^d}(v'\cdot\xi)^4 M(v')\,dv'\right)^{\frac12}}{1+\Theta|\xi|^2}\|(1-\Pi)F\| = \frac{\sqrt{K}|\xi|^2}{1+\Theta|\xi|^2}\|(1-\Pi)F\|,\end{aligned}$$

meaning that we have proven (A4) with

$$C_M = \frac{\kappa|\xi| + \sqrt{K}|\xi|^2}{1+\Theta|\xi|^2}.$$

With the elementary estimates

$$\frac{\Theta|\xi|^2}{1+\Theta|\xi|^2} \geq \min\{1,\Theta\}\frac{|\xi|^2}{1+|\xi|^2} \quad \text{and} \quad \frac{\lambda_M}{(1+\lambda_M)C_M^2} = \frac{\Theta(1+\Theta|\xi|^2)}{(\kappa+\sqrt{K}|\xi|)^2} \geq \frac{\Theta^2}{K+\Theta\kappa^2},$$

the proof is completed using (7). □

3. Enlarging and shrinking spaces by factorization

Square integrability against the inverse of the *local equilibrium* M is a rather restrictive assumption on the initial datum. In this section it will be relaxed with the help of the abstract *factorization method* of [Gualdani et al. 2017] in a simple case (factorization of order 1). Here we state the result and sketch a proof in a special case, for the convenience of the reader. We shall then give a result based on similar computations in the opposite direction: how to establish a rate in a stronger norm, which corresponds to a *shrinking* of the functional space. We will conclude with an application to the problem studied in Corollary 5. Let us start by *enlarging* the space.

Proposition 6. *Let $\mathscr{B}_1$, $\mathscr{B}_2$ be Banach spaces and let $\mathscr{B}_2$ be continuously imbedded in $\mathscr{B}_1$, i.e., $\|\cdot\|_1 \leq c_1\|\cdot\|_2$. Let $\mathfrak{B}$ and $\mathfrak{A}+\mathfrak{B}$ be the generators of the strongly continuous semigroups $e^{\mathfrak{B}t}$ and $e^{(\mathfrak{A}+\mathfrak{B})t}$ on $\mathscr{B}_1$. Assume that there are positive constants c_2, c_3, c_4, λ_1 and λ_2 such that, for all $t \geq 0$,*

$$\|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{2\to 2} \leq c_2 e^{-\lambda_2 t}, \quad \|e^{\mathfrak{B}t}\|_{1\to 1} \leq c_3 e^{-\lambda_1 t}, \quad \|\mathfrak{A}\|_{1\to 2} \leq c_4,$$

where $\|\cdot\|_{i\to j}$ denotes the operator norm for linear mappings from $\mathscr{B}_i$ to $\mathscr{B}_j$. Then there exists a positive constant $C = C(c_1, c_2, c_3, c_4)$ such that, for all $t \geq 0$,

$$\|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{1\to 1} \leq \begin{cases} C(1+|\lambda_1-\lambda_2|^{-1})e^{-\min\{\lambda_1,\lambda_2\}t} & \text{for } \lambda_1 \neq \lambda_2, \\ C(1+t)e^{-\lambda_1 t} & \text{for } \lambda_1 = \lambda_2. \end{cases}$$

Proof. Integrating the identity $\frac{d}{ds}(e^{(\mathfrak{A}+\mathfrak{B})s}e^{\mathfrak{B}(t-s)}) = e^{(\mathfrak{A}+\mathfrak{B})s}\mathfrak{A}e^{\mathfrak{B}(t-s)}$ with respect to $s \in [0,t]$ gives

$$e^{(\mathfrak{A}+\mathfrak{B})t} = e^{\mathfrak{B}t} + \int_0^t e^{(\mathfrak{A}+\mathfrak{B})s}\mathfrak{A}e^{\mathfrak{B}(t-s)}\,ds.$$

The proof is completed by the straightforward computation

$$\begin{aligned}\|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{1\to 1} &\leq c_3 e^{-\lambda_1 t} + c_1\int_0^t \|e^{(\mathfrak{A}+\mathfrak{B})s}\mathfrak{A}e^{\mathfrak{B}(t-s)}\|_{1\to 2}\,ds \\ &\leq c_3 e^{-\lambda_1 t} + c_1c_2c_3c_4 e^{-\lambda_1 t}\int_0^t e^{(\lambda_1-\lambda_2)s}\,ds.\end{aligned}$$
□

The second statement of this section is devoted to a result on the *shrinking* of the functional space. It is based on a computation which is similar to the one of the proof of Proposition 6.

Proposition 7. *Let $\mathscr{B}_1$, $\mathscr{B}_2$ be Banach spaces and let $\mathscr{B}_2$ be continuously imbedded in $\mathscr{B}_1$, i.e., $\|\cdot\|_1 \leq c_1\|\cdot\|_2$. Let $\mathfrak{B}$ and $\mathfrak{A}+\mathfrak{B}$ be the generators of the strongly continuous semigroups $e^{\mathfrak{B}t}$ and $e^{(\mathfrak{A}+\mathfrak{B})t}$*

on $\mathscr{B}_1$. Assume that there are positive constants c_2, c_3, c_4, λ_1 and λ_2 such that, for all $t \geq 0$,

$$\|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{1\to 1} \leq c_2 e^{-\lambda_1 t}, \quad \|e^{\mathfrak{B}t}\|_{2\to 2} \leq c_3 e^{-\lambda_2 t}, \quad \|\mathfrak{A}\|_{1\to 2} \leq c_4,$$

where $\|\cdot\|_{i\to j}$ denotes the operator norm for linear mappings from $\mathscr{B}_i$ to $\mathscr{B}_j$. Then there exists a positive constant $C = C(c_1, c_2, c_3, c_4)$ such that, for all $t \geq 0$,

$$\|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{2\to 2} \leq \begin{cases} C(1+|\lambda_2-\lambda_1|^{-1})e^{-\min\{\lambda_2,\lambda_1\}t} & \text{for } \lambda_2 \neq \lambda_1, \\ C(1+t)e^{-\lambda_1 t} & \text{for } \lambda_1 = \lambda_2. \end{cases}$$

Proof. Integrating the identity $\frac{d}{ds}(e^{\mathfrak{B}(t-s)}e^{(\mathfrak{A}+\mathfrak{B})s}) = e^{\mathfrak{B}(t-s)}\mathfrak{A}e^{(\mathfrak{A}+\mathfrak{B})s}$ with respect to $s \in [0,t]$ gives

$$e^{(\mathfrak{A}+\mathfrak{B})t} = e^{\mathfrak{B}t} + \int_0^t e^{\mathfrak{B}(t-s)}\mathfrak{A}e^{(\mathfrak{A}+\mathfrak{B})s}\,ds.$$

The proof is completed by the straightforward computation

$$\begin{aligned}\|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{2\to 2} &\leq c_3 e^{-\lambda_2 t} + \int_0^t \|e^{\mathfrak{B}(t-s)}\mathfrak{A}e^{(\mathfrak{A}+\mathfrak{B})s}\|_{2\to 2}\,ds \\ &\leq c_3 e^{-\lambda_2 t} + c_1\int_0^t \|e^{\mathfrak{B}(t-s)}\mathfrak{A}e^{(\mathfrak{A}+\mathfrak{B})s}\|_{1\to 2}\,ds \\ &\leq c_3 e^{-\lambda_2 t} + c_1\int_0^t \|e^{\mathfrak{B}(t-s)}\|_{2\to 2}\|\mathfrak{A}\|_{1\to 2}\|e^{(\mathfrak{A}+\mathfrak{B})s}\|_{1\to 1}\,ds \\ &\leq c_3 e^{-\lambda_2 t} + c_1 c_2 c_3 c_4 e^{-\lambda_2 t}\int_0^t e^{(\lambda_2-\lambda_1)s}\,ds. \end{aligned}$$ □

We will use Proposition 7 in Section 6.2. Coming back to the problem studied in Corollary 5, Proposition 6 applies to (3) with the spaces $\mathscr{B}_1 = \mathrm{L}^2(d\gamma_k)$, $k \in (d,\infty)$, and $\mathscr{B}_2 = \mathrm{L}^2(d\gamma_\infty)$ corresponding to the weights defined by (4). The exponential growth of γ_∞ guarantees that $\mathscr{B}_2$ is continuously imbedded in $\mathscr{B}_1$.

Corollary 8. *Assume (H1)–(H4), $k \in (d,\infty]$, and $\xi \in \mathbb{R}^d$. Then there exists a constant $C > 0$ such that solutions $\hat{f}$ of (3) with initial datum $\hat{f}_0(\xi,\cdot) \in \mathrm{L}^2(d\gamma_k)$ satisfy, with μ_ξ given by (10),*

$$\|\hat{f}(t,\xi,\cdot)\|^2_{\mathrm{L}^2(d\gamma_k)} \leq C e^{-\mu_\xi t}\|\hat{f}_0(\xi,\cdot)\|^2_{\mathrm{L}^2(d\gamma_k)} \quad \text{for all } t \geq 0.$$

Proof. In Case (a), let us define $\mathfrak{A}$ and $\mathfrak{B}$ by $\mathfrak{A}F = N\chi_R F$ and $\mathfrak{B}F = -i(v\cdot\xi)F + \mathsf{L}F - \mathfrak{A}F$, where N and R are two positive constants, χ is a smooth function such that $\mathbb{1}_{B_1} \leq \chi \leq \mathbb{1}_{B_2}$, and $\chi_R := \chi(\cdot/R)$. Here B_r is the centered ball of radius r. It has been established in [Mischler and Mouhot 2016, Lemma 3.8] that if $k > d$, then the inequality

$$\int_{\mathbb{R}^d}(\mathsf{L}-\mathfrak{A})(F)F\,d\gamma_k \leq -\lambda_1\int_{\mathbb{R}^d} F^2\,d\gamma_k$$

holds for some $\lambda_1 > 0$. Moreover, λ_1 can be chosen arbitrarily large for R and N large enough. The boundedness of $\mathfrak{A}:\mathscr{B}_1 \to \mathscr{B}_2$ follows from the compactness of the support of χ and Proposition 6 applies with $\lambda_2 = \frac{1}{2}\mu_\xi \leq \frac{1}{4}$, where μ_ξ is given by (10).

In Case (b), we consider $\mathfrak{A}$ and $\mathfrak{B}$ such that

$$\mathfrak{A}F(v) = M(v)\int_{\mathbb{R}^d}\sigma(v,v')F(v')\,dv',$$
$$\mathfrak{B}F(v) = -\Big[i(v\cdot\xi)+\int_{\mathbb{R}^d}\sigma(v,v')M(v')\,dv'\Big]F(v).$$

The boundedness of $\mathfrak{A}:\mathscr{B}_1\to\mathscr{B}_2$ follows from (H2) and

$$\|\mathfrak{A}F\|_{\mathrm{L}^2(d\gamma_\infty)} \le \bar\sigma\|F\|_{\mathrm{L}^1(dv)} \le \bar\sigma\Big(\int_{\mathbb{R}^d}\gamma_k^{-1}\,dv\Big)^{\frac12}\|F\|_{\mathrm{L}^2(d\gamma_k)}.$$

Proposition 6 applies with $\lambda_2=\frac12\mu_\xi\le\frac14$ and $\lambda_1=1$ because $\int_{\mathbb{R}^d}\sigma(v,v')M(v')\,dv'\ge 1$. □

4. Asymptotic behavior based on mode-by-mode estimates

In this section we consider (1) and use the estimates of Corollary 5 with weight $\gamma_\infty = 1/M$ and Corollary 8 for weights with $O(|v|^k)$ growth to get decay rates with respect to t. We shall consider two cases for the spatial variable x. In Section 4.1, we assume that $x\in\mathbb{T}^d$, where $\mathbb{T}^d$ is the flat d-dimensional torus (represented by $[0,2\pi)^d$ with periodic boundary conditions) and prove an exponential convergence rate. In Section 4.2, we assume that $x\in\mathbb{R}^d$ and establish algebraic decay rates.

4.1. ***Exponential convergence to equilibrium in*** $\mathbb{T}^d$**.** In the periodic case $x\in\mathbb{T}^d$ there is a unique nonzero normalized equilibrium given by

$$f_\infty(x,v)=\rho_\infty M(v)\quad\text{with}\quad \rho_\infty=\frac{1}{|\mathbb{T}^d|}\iint_{\mathbb{T}^d\times\mathbb{R}^d} f_0\,dx\,dv.$$

Corollary 9. *Assume (H1)–(H4) and $k\in(d,\infty]$. Then there exists a constant $C>0$ such that the solution f of (1) on $\mathbb{T}^d\times\mathbb{R}^d$ with initial datum $f_0\in\mathrm{L}^2(dx\,d\gamma_k)$ satisfies, with Λ given by (10),*

$$\|f(t,\cdot,\cdot)-f_\infty\|_{\mathrm{L}^2(dx\,d\gamma_k)}\le C\|f_0-f_\infty\|_{\mathrm{L}^2(dx\,d\gamma_k)}e^{-\Lambda\frac{t}{4}}\quad \textit{for all } t\ge 0.$$

Proof. We represent the flat torus $\mathbb{T}^d$ by $[0,2\pi)^d$ with periodic boundary conditions, and the Fourier variable is denoted by $\xi\in\mathbb{Z}^d$. For $\xi=0$, the microscopic coercivity (see Section 2) implies

$$\|\hat f(t,0,\cdot)-\hat f_\infty(0,\cdot)\|_{\mathrm{L}^2(d\gamma_\infty)}\le\|\hat f_0(0,\cdot)-\hat f_\infty(0,\cdot)\|_{\mathrm{L}^2(d\gamma_\infty)}e^{-t}.$$

For all other modes, $\hat f_\infty(\xi,\cdot)=0$ for any $\xi\ne0$ (that is, for any ξ such that $|\xi|\ge1$). We can use Corollary 5 with $\mu_\xi\ge\frac12\Lambda$, with the notation of (10). An application of Parseval's identity then proves the result for $k=\infty$ and $C=\sqrt{3}$. If k is finite, the result with the weight γ_k follows from Corollary 8. □

Note that the latter result can also alternatively be proved by directly applying Proposition 4 to (1), as in [Dolbeault et al. 2015].

4.2. *Algebraic decay rates in* $\mathbb{R}^d$. With the result of Corollaries 5 and 8 we obtain a first proof of Theorem 1 as follows. Let $C > 0$ be a generic constant which is going to change from line to line. Plancherel's formula implies

$$\|f(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_k)} \le C \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} e^{-\mu_\xi t} |\hat{f}_0|^2 \, d\xi \right) d\gamma_k.$$

We know that

$$\int_{|\xi|\le 1} e^{-\mu_\xi t} \, d\xi \le \int_{\mathbb{R}^d} e^{-\frac{\Lambda}{2}|\xi|^2 t} \, d\xi = \left(\frac{2\pi}{\Lambda t}\right)^{\frac{d}{2}}$$

and thus, for all $v \in \mathbb{R}^d$,

$$\int_{|\xi|\le 1} e^{-\mu_\xi t} |\hat{f}_0|^2 \, d\xi \le C \|f_0(\cdot,v)\|^2_{\mathrm{L}^1(dx)} \int_{\mathbb{R}^d} e^{-\frac{\Lambda}{2}|\xi|^2 t} \, d\xi \le C \|f_0(\cdot,v)\|^2_{\mathrm{L}^1(dx)} t^{-\frac{d}{2}}.$$

Using the fact that $\mu_\xi \ge \frac{1}{2}\Lambda$ when $|\xi| \ge 1$ and Plancherel's formula, we know that, for all $v \in \mathbb{R}^d$,

$$\int_{|\xi|>1} e^{-\mu_\xi t} |\hat{f}_0|^2 \, d\xi \le C e^{-\frac{\Lambda}{2} t} \|f_0(\cdot,v)\|^2_{\mathrm{L}^2(dx)},$$

which completes a first proof of Theorem 1.

5. Hypocoercivity and the Nash inequality

In view of the proof of Theorem 1 in Section 4.2 and of the rate, it is natural to wonder if the hypocoercivity can be controlled by the use of Nash's inequality. Here we temporarily abandon the Fourier variable ξ and consider the direct variable $x \in \mathbb{R}^d$: throughout this section, the *transport operator* on the position space is defined as

$$\mathsf{T} f = v \cdot \nabla_x f.$$

We rely on the abstract setting of Section 2, applied to (1) with the scalar product $\langle \cdot,\cdot \rangle$ on $\mathrm{L}^2(dx\,d\gamma_\infty)$ and the induced norm $\|\cdot\|$. Notice that this norm includes the x variable, which was not the case in the mode-by-mode analysis of Section 2. It is then easy to check that $(\mathsf{T}\Pi) f = M \mathsf{T} \rho_f = v \cdot \nabla_x \rho_f M$, $(\mathsf{T}\Pi)^* f = -\nabla_x \cdot \big(\int_{\mathbb{R}^d} v f \, dv\big) M$, and $(\mathsf{T}\Pi)^*(\mathsf{T}\Pi) f = -\Theta \Delta_x \rho_f M$ so that

$$g = \mathsf{A} f = (1 + (\mathsf{T}\Pi)^* \mathsf{T}\Pi)^{-1} (\mathsf{T}\Pi)^* f \quad \Longleftrightarrow \quad g = u M,$$

where $u - \Theta \Delta u = -\nabla_x \cdot \big(\int_{\mathbb{R}^d} v f \, dv\big)$. Since M is unbiased, $\mathsf{A} f = \mathsf{A}(1-\Pi) f$. For some $\delta > 0$ to be chosen later, we redefine the entropy by $\mathsf{H}[f] := \frac{1}{2}\|f\|^2 + \delta \langle \mathsf{A} f, f \rangle$.

Proof of Theorem 1. If f solves (1), the time derivative of $\mathsf{H}[f(t,\cdot,\cdot)]$ is given by

$$\frac{d}{dt}\mathsf{H}[f] = -\mathsf{D}[f], \tag{12}$$

where, as in the proof of Proposition 4,

$$\mathsf{D}[f] := -\langle \mathsf{L} f, f \rangle + \delta \langle \mathsf{A}\mathsf{T}\Pi f, f \rangle + \delta \,\mathrm{Re}\langle \mathsf{A}\mathsf{T}(1-\Pi) f, f \rangle - \delta \,\mathrm{Re}\langle \mathsf{T}\mathsf{A} f, f \rangle - \delta \,\mathrm{Re}\langle \mathsf{A}\mathsf{L} f, f \rangle.$$

Here we use the fact that $\langle \mathsf{A} f, \mathsf{L} f \rangle = 0$. The first term in $\mathsf{D}[f]$ satisfies the microscopic coercivity condition

$$-\langle \mathsf{L} f, f \rangle \geq \lambda_m \|(1-\Pi) f\|^2.$$

The second term in (12) is computed as follows. Solving $g = \mathsf{AT}\Pi f$ is equivalent to solving $(1 + (\mathsf{T}\Pi)^* \mathsf{T}\Pi) g = (\mathsf{T}\Pi)^* \mathsf{T}\Pi f$, i.e.,

$$v_f - \Theta \Delta_x v_f = -\Theta \Delta_x \rho_f, \tag{13}$$

where $g = v_f M$. Hence

$$\langle \mathsf{AT}\Pi f, f \rangle = \int_{\mathbb{R}^d} v_f \, \rho_f \, dx.$$

A direct application of the hypocoercivity approach of [Dolbeault et al. 2015] to the whole-space problem fails by lack of a *macroscopic coercivity* condition. Although the second term in (12) is not coercive, we observe that the last three terms in (12) can still be dominated by the first two for $\delta > 0$, small enough, as follows.

(1) As in [Dolbeault et al. 2015], we use the adjoint operators to compute

$$\langle \mathsf{AT}(1-\Pi) f, f \rangle = -\langle (1-\Pi) f, \mathsf{TA}^* f \rangle.$$

We observe that

$$\mathsf{A}^* f = \mathsf{T}\Pi (1 + (\mathsf{T}\Pi)^* \mathsf{T}\Pi)^{-1} f = \mathsf{T}(1 + (\mathsf{T}\Pi)^* \mathsf{T}\Pi)^{-1} \Pi f = M \mathsf{T} u_f = v M \cdot \nabla_x u_f,$$

where u_f is the solution in $\mathrm{H}^1(dx)$ of

$$u_f - \Theta \Delta_x u_f = \rho_f. \tag{14}$$

With K defined by (11), we obtain

$$\|\mathsf{TA}^* f\|^2 \leq K \|\nabla_x^2 u_f\|^2_{\mathrm{L}^2(dx)} = K \|\Delta_x u_f\|^2_{\mathrm{L}^2(dx)}.$$

On the other hand, we observe that $v_f = -\Theta \Delta u_f$ solves (13). Hence by multiplying (14) by $v_f = -\Theta \Delta u_f$ and integrating by parts, we know that

$$\Theta \|\nabla_x u_f\|^2_{\mathrm{L}^2(dx)} + \Theta^2 \|\Delta_x u_f\|^2_{\mathrm{L}^2(dx)} = \int_{\mathbb{R}^d} v_f \, \rho_f \, dx = \langle \mathsf{AT}\Pi f, f \rangle. \tag{15}$$

Notice that a central feature of our method is the fact that quantities of interest involving the operator A can be computed by solving an elliptic equation (for instance (13) in case of $\mathsf{AT}\Pi f$ or (14) in case of $\mathsf{A}^* f$). Altogether we obtain

$$|\langle \mathsf{AT}(1-\Pi) f, f \rangle| \leq \|(1-\Pi) f\| \, \|\mathsf{TA}^* f\| \leq \frac{\sqrt{K}}{\Theta} \|(1-\Pi) f\| \langle \mathsf{AT}\Pi f, f \rangle^{\frac{1}{2}}.$$

(2) By (8), we have

$$|\langle \mathsf{TA} f, f | = |\langle \mathsf{TA}(1-\Pi) f, (1-\Pi) f \rangle| \leq \|(1-\Pi) f\|^2.$$

(3) It remains to estimate the last term on the right-hand side of (12). Let us consider the solution u_f of (14). If we multiply (13) by u_f and integrate, we observe that

$$\Theta\|\nabla_x u_f\|^2_{\mathrm{L}^2(dx)} = \int_{\mathbb{R}^d} u_f v_f\,dx \le \int_{\mathbb{R}^d} u_f v_f\,dx + \int_{\mathbb{R}^d} |v_f|^2\,dx = \int_{\mathbb{R}^d} v_f \rho_f\,dx$$

because $v_f = -\Theta\Delta u_f$, so that

$$\|\mathsf{A}^* f\|^2 = \Theta\|\nabla_x u_f\|^2_{\mathrm{L}^2(dx)} \le \langle \mathsf{AT}\Pi f, f\rangle.$$

In Case (a), we compute

$$\langle \mathsf{AL} f, f\rangle = \langle \mathsf{L}(1-\Pi)f, \mathsf{A}^* f\rangle = \iint_{\mathbb{R}^d\times\mathbb{R}^d} \nabla_x u_f \cdot \frac{\nabla_v M}{M}(1-\Pi)f\,dx\,dv.$$

It follows from the Cauchy–Schwarz inequality that

$$\int_{\mathbb{R}^d} |\nabla_v M|\,|(1-\Pi)f|\,d\gamma_\infty \le \|\nabla_v M\|_{\mathrm{L}^2(d\gamma_\infty)}\|(1-\Pi)f\|_{\mathrm{L}^2(d\gamma_\infty)} = \sqrt{d\theta}\,\|(1-\Pi)f\|_{\mathrm{L}^2(d\gamma_\infty)}$$

and

$$|\langle \mathsf{AL} f, f\rangle| \le \|\nabla_x u_f\|_{\mathrm{L}^2(dx)}\left(\int_{\mathbb{R}^d}\left(\frac{1}{d}\int_{\mathbb{R}^d} |\nabla_v M|\,|(1-\Pi)f|\,d\gamma\right)^2 dx\right)^{\frac12}.$$

Altogether, we obtain

$$|\langle \mathsf{AL} f, f\rangle| \le \sqrt{\frac{\theta}{\Theta}}\,\|(1-\Pi)f\|\,\langle \mathsf{AT}\Pi f, f\rangle^{\frac12}.$$

In Case (b), we use (H2) to get

$$|\langle \mathsf{AL} f, f\rangle| \le \|\mathsf{L} f\|\,\|\mathsf{A}^* f\| \le 2\bar\sigma\|(1-\Pi)f\|\,\|\mathsf{A}^* f\| \le 2\bar\sigma\|(1-\Pi)f\|\,\langle \mathsf{AT}\Pi f, f\rangle^{\frac12}.$$

In both cases, (a) and (b), the estimate can be written as

$$|\langle \mathsf{AL} f, f\rangle| \le 2\bar\sigma\|(1-\Pi)f\|\,\langle \mathsf{AT}\Pi f, f\rangle^{\frac12},$$

with the convention that $\bar\sigma = \frac12\sqrt{\theta/\Theta}$ in Case (a).

Summarizing, we know that

$$-\frac{d}{dt}\mathsf{H}[f] \ge (\lambda_m - \delta)X^2 + \delta Y^2 + 2\delta \mathsf{b} XY,$$

with $X := \|(1-\Pi)f\|$, $Y := \langle \mathsf{AT}\Pi f, f\rangle^{1/2}$, and $\mathsf{b} := K/(2\Theta) + 2\bar\sigma$. The largest $\mathsf{a} > 0$ such that

$$(\lambda_m - \delta)X^2 + \delta Y^2 + 2\delta \mathsf{b} XY \ge \mathsf{a}(X^2 + 2Y^2)$$

holds for any $X, Y \in \mathbb{R}$ is given by the conditions

$$\mathsf{a} < \lambda_m - \delta, \quad 2\mathsf{a} < \delta, \quad \delta^2\mathsf{b}^2 - (\lambda_m - \delta - \mathsf{a})(\delta - 2\mathsf{a}) \le 0 \tag{16}$$

and it is easy to check that there exists a positive solution if $\delta > 0$ is small enough. To fulfill the additional constraint $\delta < 1$, we can for instance choose

$$\delta = \frac{4\min\{1,\lambda_m\}}{8\mathsf{b}^2+5} \quad \text{and} \quad \mathsf{a} = \frac{\delta}{4}.$$

Altogether we obtain

$$-\frac{d}{dt}\mathsf{H}[f] \geq \mathsf{a}\big(\|(1-\Pi)f\|^2 + 2\langle \mathsf{AT}\Pi f, f\rangle\big).$$

Using (14) and (15), we control $\|\Pi f\|^2 = \|\rho_f\|^2_{\mathrm{L}^2(dx)}$ by $\langle \mathsf{AT}\Pi f, f\rangle$ according to

$$\|\Pi f\|^2 = \|u_f\|^2_{\mathrm{L}^2(dx)} + 2\Theta\|\nabla_x u_f\|^2_{\mathrm{L}^2(dx)} + \Theta^2\|\Delta_x u_f\|^2_{\mathrm{L}^2(dx)} \leq \|u_f\|^2_{\mathrm{L}^2(dx)} + 2\langle \mathsf{AT}\Pi f, f\rangle.$$

We observe that, for any $t \geq 0$,

$$\|u_f(t,\cdot)\|_{\mathrm{L}^1(dx)} = \|\rho_f(t,\cdot)\|_{\mathrm{L}^1(dx)} = \|f_0\|_{\mathrm{L}^1(dx\,dv)}, \quad \|\nabla_x u_f\|^2_{\mathrm{L}^2(dx)} \leq \frac{1}{\Theta}\langle \mathsf{AT}\Pi f, f\rangle.$$

We recall the *Nash inequality* [1958]

$$\|u\|^2_{\mathrm{L}^2(dx)} \leq \mathscr{C}_{\mathrm{Nash}}\|u\|^{\frac{4}{d+2}}_{\mathrm{L}^1(dx)}\|\nabla u\|^{\frac{2d}{d+2}}_{\mathrm{L}^2(dx)} \tag{17}$$

for any function $u \in \mathrm{L}^1 \cap \mathrm{H}^1(\mathbb{R}^d)$. We use (17) with $u = u_f$ to get

$$\|\Pi f\|^2 \leq \Phi^{-1}(2\langle \mathsf{AT}\Pi f, f\rangle), \quad \text{with } \Phi^{-1}(y) := y + \left(\frac{y}{\mathsf{c}}\right)^{\frac{d}{d+2}} \quad \text{for all } y \geq 0,$$

where

$$\mathsf{c} = 2\Theta\mathscr{C}_{\mathrm{Nash}}^{-1-\frac{2}{d}}\|f_0\|^{-\frac{4}{d}}_{\mathrm{L}^1(dx\,dv)}.$$

The function $\Phi : [0,\infty) \to [0,\infty)$ satisfies $\Phi(0) = 0$ and $0 < \Phi' < 1$, so that

$$\|(1-\Pi)f\|^2 + 2\langle \mathsf{AT}\Pi f, f\rangle \geq \Phi(\|f\|^2) \geq \Phi\left(\frac{2}{1+\delta}\mathsf{H}[f]\right),$$

where the last inequality holds as a consequence of (9). From

$$z = \Phi^{-1}(y) = y + \left(\frac{y}{\mathsf{c}}\right)^{\frac{d}{d+2}} \leq y_0^{\frac{2}{d+2}}y^{\frac{d}{d+2}} + \left(\frac{y}{\mathsf{c}}\right)^{\frac{d}{d+2}} = \left(y_0^{\frac{2}{d+2}} + \mathsf{c}^{-\frac{d}{d+2}}\right)y^{\frac{d}{d+2}},$$

as long as $y \leq y_0$, for y_0 to be chosen later, we have

$$y = \Phi(z) \geq \left(\Phi(z_0)^{\frac{2}{d+2}} + \mathsf{c}^{-\frac{d}{d+2}}\right)^{-\frac{d+2}{d}} z^{1+\frac{2}{d}},$$

as long as $z \leq z_0 := \Phi^{-1}(y_0)$. Since $\frac{d}{dt}\mathsf{H}[f] \leq 0$, we have

$$\frac{2}{1+\delta}\mathsf{H}[f] \leq \frac{2}{1+\delta}\mathsf{H}[f_0].$$

We thus apply the previous inequalities with

$$z_0 = \frac{2}{1+\delta}\mathsf{H}[f_0]$$

together with the fact that

$$\Phi(z_0) \geq z_0 \geq \frac{1-\delta}{1+\delta}\|f_0\|^2$$

and that c is proportional to $\|f_0\|_{\mathrm{L}^1(dx\,dv)}^{-4/d}$, to get

$$\Phi\left(\frac{2}{1+\delta}\mathsf{H}[f]\right) \gtrsim \left(\|f_0\|_{\mathrm{L}^2(dx\,d\gamma_\infty)}^{\frac{4}{d+2}} + \|f_0\|_{\mathrm{L}^1(dx\,dv)}^{\frac{4}{d+2}}\right)^{-\frac{d+2}{d}} \mathsf{H}[f]^{1+\frac{2}{d}}.$$

We deduce the entropy decay inequality

$$-\frac{d}{dt}\mathsf{H}[f] \gtrsim \left(\|f_0\|_{\mathrm{L}^2(dx\,d\gamma_\infty)}^{\frac{4}{d+2}} + \|f_0\|_{\mathrm{L}^1(dx\,dv)}^{\frac{4}{d+2}}\right)^{-\frac{d+2}{d}} \mathsf{H}[f]^{1+\frac{2}{d}}. \tag{18}$$

A simple integration from 0 to t shows that

$$\mathsf{H}[f] \lesssim \left[\mathsf{H}[f_0]^{-\frac{2}{d}} + \left(\|f_0\|_{\mathrm{L}^2(dx\,d\gamma_\infty)}^{\frac{4}{d+2}} + \|f_0\|_{\mathrm{L}^1(dx\,dv)}^{\frac{4}{d+2}}\right)^{-\frac{d+2}{d}} t\right]^{-\frac{d}{2}}.$$

The result of Theorem 1 then follows from elementary considerations. □

Using moments instead of the mass, it is possible to state an *improved Nash inequality*: there exists a positive constant $\mathscr{C}_\star$ such that

$$\|u\|_{\mathrm{L}^2(dx)}^2 \leq \mathscr{C}_\star \|x\,u\|_{\mathrm{L}^1(dx)}^{\frac{4}{d+4}} \|\nabla u\|_{\mathrm{L}^2(dx)}^{\frac{d+2}{d+4}}$$

for any $u \in \mathrm{H}^1(dx) \cap \mathrm{L}^1((1+|x|)\,dx)$ such that $\int_{\mathbb{R}^d} u\,dx = 0$. The proof follows from a minor modification of the original proof (attributed by Nash himself to Stein) in [Nash 1958] and uses Fourier variables. As a consequence, any solution of the heat equation with zero average decays in $\mathrm{L}^2(dx)$ like $O(t^{-1-d/2})$ as $t \to +\infty$. It is the topic of the following section to use Fourier variables in the spirit of Nash's proof to get improved rates of decay at the level of the kinetic equation.

6. Algebraic decay rates in $\mathbb{R}^d$ by Fourier estimates and improvements

We prove Theorem 2 in Section 6.1 and Theorem 3 in Section 6.2.

6.1. ***Improved decay rates.*** Let us prove Theorem 2 by Fourier methods inspired by the proof of Nash's inequality.

Step 1: decay of the average in space by a factorization argument. We define

$$f_\bullet(t,v) := \int_{\mathbb{R}^d} f(t,x,v)\,dx \tag{19}$$

and observe that $f_\bullet$ solves

$$\partial_t f_\bullet = \mathsf{L} f_\bullet.$$

As a consequence, we have that $0 = \int_{\mathbb{R}^d} f_\bullet(t,v)\,dv$. From the *microscopic coercivity property* (H4), we deduce that

$$\|f_\bullet(t,\cdot)\|_{\mathrm{L}^2(d\gamma_\infty)}^2 = \int_{\mathbb{R}^d}\left|\frac{f_\bullet(t,v)}{M}\right|^2 M\,dv \leq \|f_\bullet(0,\cdot)\|_{\mathrm{L}^2(d\gamma_\infty)}^2 e^{-\lambda_m t} \quad \text{for all } t \geq 0.$$

With $k \in (d, \infty)$, Proposition 6 applies like in the proof of Corollary 8 or in [Mischler and Mouhot 2016]. We observe that $\|f_\bullet(0,\cdot)\|_{\mathrm{L}^2(|v|^2\,d\gamma_k)} \le \|f_0\|_{\mathrm{L}^2(|v|^2\,d\gamma_k;\mathrm{L}^1(dx))}$. For some positive constants C and λ, we get

$$\|f_\bullet(t,\cdot)\|^2_{\mathrm{L}^2(|v|^2\,d\gamma_k)} \le C\,\|f_0\|^2_{\mathrm{L}^2(|v|^2\,d\gamma_k;\mathrm{L}^1(dx))} e^{-\lambda t} \quad \text{for all } t \ge 0. \tag{20}$$

Step 2: improved decay of f. Let us define $g(t,x,v) := f(t,x,v) - f_\bullet(t,v)\varphi(x)$, where φ is a given positive function satisfying

$$\int_{\mathbb{R}^d} \varphi(x)\,dx = 1, \quad \text{e.g.,} \quad \varphi(x) := (2\pi)^{-\frac{d}{2}} e^{-\frac{|x|^2}{2}}, \quad \text{for all } x \in \mathbb{R}^d.$$

Since $\partial_t f_\bullet = \mathsf{L} f_\bullet$, the Fourier transform $\hat g(t,\xi,v)$ of $g(t,x,v)$ solves

$$\partial_t \hat g + \mathsf{T}\hat g = \mathsf{L}\hat g - f_\bullet \mathsf{T}\hat\varphi,$$

where $\mathsf{T}\hat\varphi = i(v\cdot\xi)\hat\varphi$. Using Duhamel's formula

$$\hat g = e^{(\mathsf{L}-\mathsf{T})t}\hat g_0 - \int_0^t e^{(\mathsf{L}-\mathsf{T})(t-s)} f_\bullet(s,v)\mathsf{T}\hat\varphi(\xi)\,ds,$$

Corollary 5, and Proposition 6, for some generic constant $C > 0$ which will change from line to line, we get

$$\|\hat g(t,\xi,\cdot)\|_{\mathrm{L}^2(d\gamma_k)} \le C e^{-\frac12 \mu_\xi t}\|\hat g_0(\xi,\cdot)\|_{\mathrm{L}^2(d\gamma_k)} + C\int_0^t e^{-\frac{\mu_\xi}{2}(t-s)}\|f_\bullet(s,\cdot)\|_{\mathrm{L}^2(|v|^2\,d\gamma_k)}|\xi|\,|\hat\varphi(\xi)|\,ds. \tag{21}$$

The key observation is $\hat g_0(0,v) = 0$, so that

$$\hat g_0(\xi,v) = \int_0^{|\xi|} \frac{\xi}{|\xi|}\cdot\nabla_\xi \hat g_0\Big(\eta\frac{\xi}{|\xi|}, v\Big)\,d\eta$$

yields

$$|\hat g_0(\xi,v)| \le |\xi|\,\|\nabla_\xi \hat g_0(\cdot,v)\|_{\mathrm{L}^\infty(d\xi)} \le |\xi|\,\|g_0(\cdot,v)\|_{\mathrm{L}^1(|x|\,dx)} \quad \text{for all } (\xi,v) \in \mathbb{R}^d\times\mathbb{R}^d.$$

We know from (10) that $\mu_\xi = \Lambda|\xi|^2/(1+|\xi|^2)$. The first term of the right-hand side of (21) can therefore be estimated for any $t \ge 1$ by

$$\begin{aligned}\bigg(\int_{|\xi|\le 1}\int_{\mathbb{R}^d} |e^{(\mathsf{L}-\mathsf{T})t}\hat g_0|^2\,d\gamma_k\,d\xi\bigg)^{\frac12} &\le \bigg(\int_{\mathbb{R}^d} |\xi|^2 e^{-\frac{\Lambda}{2}|\xi|^2 t}\,d\xi\bigg)^{\frac12}\|g_0\|_{\mathrm{L}^2(d\gamma_k;\mathrm{L}^1(|x|\,dx))}\\ &\le \frac{C}{(1+t)^{1+\frac{d}{2}}}\|g_0\|_{\mathrm{L}^2(d\gamma_k;\mathrm{L}^1(|x|\,dx))},\end{aligned}$$

which is the leading-order term as $t \to \infty$, and we have

$$\int_{|\xi|>1} e^{-\mu_\xi t}\|\hat g_0(\xi,\cdot)\|^2_{\mathrm{L}^2(d\gamma_k)}\,d\xi \le C e^{-\frac{\Lambda}{2}t}\|g_0\|^2_{\mathrm{L}^2(dx\,d\gamma_k)}$$

for any $t \ge 0$, using the fact that $\mu_\xi \ge \frac12\Lambda$ when $|\xi| \ge 1$ and Plancherel's formula.

Using (20), the second term of the right-hand side of (21) is estimated by

$$\int_{\mathbb{R}^d}\left(\int_0^t e^{-\frac{\mu_\xi}{2}(t-s)}\|f_\bullet(s,\cdot)\|_{\mathrm{L}^2(|v|^2\,d\gamma_k)}|\xi|\,|\hat{\varphi}(\xi)|\,ds\right)^2 d\xi$$
$$\le C\|f_0\|^2_{\mathrm{L}^2(|v|^2\,d\gamma_k;\mathrm{L}^1(dx))}\int_{\mathbb{R}^d}|\xi|^2|\hat{\varphi}(\xi)|^2\left(\int_0^t e^{-\frac{\mu_\xi}{2}(t-s)}e^{-\frac{\lambda}{2}s}\,ds\right)^2 d\xi.$$

On the one hand, we use the Cauchy–Schwarz inequality to get

$$\int_{|\xi|\le 1}|\xi|^2|\hat{\varphi}(\xi)|^2\left(\int_0^t e^{-\frac{\mu_\xi}{2}(t-s)}e^{-\frac{\lambda}{2}s}\,ds\right)^2 d\xi$$
$$\le \|\varphi\|^2_{\mathrm{L}^1(dx)}\int_{|\xi|\le 1}|\xi|^2\left(\int_0^t e^{-\mu_\xi(t-s)}e^{-\frac{\lambda}{2}s}\,ds\right)\left(\int_0^t e^{-\frac{\lambda}{2}s}\,ds\right)d\xi$$
$$\le \frac{2}{\lambda}\|\varphi\|^2_{\mathrm{L}^1(dx)}\int_0^t\left(\int_{|\xi|\le 1}|\xi|^2 e^{-\frac{\Lambda}{2}|\xi|^2(t-s)}\,d\xi\right)e^{-\frac{\lambda}{2}s}\,ds \le C_1 t^{-\frac{d}{2}-1}+C_2 e^{-\frac{\lambda}{4}t},$$

where the last inequality is obtained by splitting the integral in s on $\left(0,\frac{1}{2}t\right)$ and $\left(\frac{1}{2}t,t\right)$. On the other hand, using $\mu_\xi\ge\frac{1}{2}\Lambda$ when $|\xi|\ge 1$, we obtain

$$\int_{|\xi|\ge 1}|\xi|^2|\hat{\varphi}(\xi)|^2\left(\int_0^t e^{-\frac{\mu_\xi}{2}(t-s)}e^{-\frac{\lambda}{2}s}\,ds\right)^2 d\xi \le t^2 e^{-\min\{\frac{1}{2}\Lambda,\lambda\}t}\|\nabla\varphi\|^2_{\mathrm{L}^2(dx)}.$$

By collecting all terms, we deduce that $\|g(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_k)}$ is bounded by

$$C\big(\|g_0\|^2_{\mathrm{L}^2(d\gamma_k;\mathrm{L}^1(|x|\,dx))}+\|f_0\|^2_{\mathrm{L}^2((|v|^2\,d\gamma_k;\mathrm{L}^1(dx))}\big)(1+t)^{-(1+\frac{d}{2})}$$

for some constant $C>0$. Recalling that $f=g+f_\bullet\varphi$, the proof of Theorem 2 is completed using (20).

6.2. ***Improved decay rates with higher-order cancellations.*** We prove Theorem 3, which means that from now on we assume in Case (a) that M is a normalized Gaussian (2), and in Case (b) that $\sigma\equiv 1$. Moreover, the initial data satisfies (5); that is,

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} f_0 P\,dx\,dv = 0 \quad \text{for all } P\in\mathbb{R}_\ell[X,V].$$

For any $P\in\mathbb{R}_\ell[X]$, let

$$P[f](t,v) := \int_{\mathbb{R}^d} P(x)f(t,x,v)\,dx,$$

so that $\int_{\mathbb{R}^d} P[f](0,v)\,dv = 0$.

In this section we use the notation $\lesssim_k$ to express inequalities up to a constant which depends on k.

Step 1: conservation of zero moments. For a solution f of (1) we compute

$$\frac{d}{dt}\iint_{\mathbb{R}^d\times\mathbb{R}^d} f(t,x,v)P(x,v)\,dx\,dv = -\iint_{\mathbb{R}^d\times\mathbb{R}^d}(v\cdot\nabla_x f)P\,dx\,dv+\iint_{\mathbb{R}^d\times\mathbb{R}^d}(\mathsf{L}f)P\,dx\,dv$$
$$=\iint_{\mathbb{R}^d\times\mathbb{R}^d}(v\cdot\nabla_x P)f\,dx\,dv+\iint_{\mathbb{R}^d\times\mathbb{R}^d}(\mathsf{L}f)P\,dx\,dv.$$

In Case (a) of a Fokker–Planck operator, we may write

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} (\mathsf{L} f)P\,dx\,dv = \iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{M}\nabla_v\cdot(M\nabla_v P)f\,dx\,dv = \iint_{\mathbb{R}^d\times\mathbb{R}^d} (\Delta_v P - v\cdot\nabla_v P)f\,dx\,dv.$$

By the definition of $\mathbb{R}_\ell[X,V]$, it turns out that $\Delta_v P - v\cdot\nabla_v P \in \mathbb{R}_\ell[X,V]$. For the scattering operator of Case (b), one has

$$\begin{aligned}\iint_{\mathbb{R}^d\times\mathbb{R}^d} &(\mathsf{L} f)P\,dx\,dv\\ &= \iint_{\mathbb{R}^d\times\mathbb{R}^d}\left(\int_{\mathbb{R}^d}(M(v)f(t,x,v')-M(v')f(t,x,v))\,dv'\right)P(x,v)\,dx\,dv\\ &= \iiint_{\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d}(M(v)f(t,x,v')-M(v')f(t,x,v))P(x,v)\,dx\,dv\,dv'\\ &= \iint_{\mathbb{R}^d\times\mathbb{R}^d}\left(\int_{\mathbb{R}^d} M(v)P(x,v)\,dv\right)f(t,x,v')\,dx\,dv' - \iint_{\mathbb{R}^d\times\mathbb{R}^d} f(x,v)P(x,v)\,dx\,dv.\end{aligned}$$

One can check that $\int_{\mathbb{R}^d} M(v)P(x,v)\,dv \in \mathbb{R}_\ell[X]$. Since also $v\cdot\nabla_x P \in \mathbb{R}_\ell[X,V]$, the evolution of moments of order lower than or equal to ℓ is equivalent to a linear ODE of the form $\dot Y(t) = QY(t)$, where Q is a matrix resulting from the previous computations. Consequently, if $Y(0)=0$ initially, it remains null for all times.

Step 2: decay of polynomial averages in space. We claim that for any $j\le\ell$, there exists $\lambda>0$ such that, for any $P\in\mathbb{R}_j[X]$ and $q\in\mathbb{N}$,

$$\|P[f](t,\cdot)\|_{\mathrm{L}^2(d\gamma_{k+q})} \lesssim_{j,q} \|f_0\|_{\mathrm{L}^2(d\gamma_{k+q+2j};\mathrm{L}^1((1+|x|^j)\,dx))}(1+t)^j e^{-\lambda t} \quad \text{for all } t\ge 0. \tag{22}$$

Let us prove it by induction.

(1) *The case* $j=0$. Notice that $j=0$ means that P is a real number and $P[f]=f_\bullet$ as defined in (19), up to a multiplication by a constant. Since $\int_{\mathbb{R}^d} f_\bullet(t,v)\,dv=0$ for any $t\ge0$, one has $\partial_t f_\bullet = \mathsf{L} f_\bullet$; thus we deduce from the *microscopic coercivity property* as above that

$$\|f_\bullet(t,\cdot)\|_{\mathrm{L}^2(d\gamma_\infty)} \le \|f_\bullet(0,\cdot)\|_{\mathrm{L}^2(d\gamma_\infty)}e^{-\lambda_m t} \quad \text{for all } t\ge0.$$

We also obtain

$$\|f_\bullet(t,\cdot)\|_{\mathrm{L}^2(d\gamma_{k+q})} \lesssim_q \|f_0\|_{\mathrm{L}^2(d\gamma_{k+q};\mathrm{L}^1(dx))}e^{-\lambda t} \quad \text{for all } t\ge0, \tag{23}$$

but this requires some comments. The case $k\in(d,\infty)$ is covered by Corollary 8.

The case $k=\infty$ in (23) is given by the following lemma.

Lemma 10. *Under the assumptions of Theorem 3, one has*

$$\|f_\bullet(t,\cdot)\|_{\mathrm{L}^2((1+|v|^q)\,d\gamma_\infty)} \lesssim_q \|f_0\|_{\mathrm{L}^2((1+|v|^q)\,d\gamma_\infty;\mathrm{L}^1(dx))}e^{-\lambda t} \quad \textit{for all } t\ge0.$$

Proof. We rely on Proposition 7 with the Banach spaces $\mathscr{B}_1=\mathsf{L}^2(d\gamma_\infty)$ and $\mathscr{B}_2=\mathsf{L}^2((1+|v|^q)\,d\gamma_\infty)$. In Case (a), let us define $\mathfrak{A}$ and $\mathfrak{B}$ by $\mathfrak{A}F=N\chi_R F$ and $\mathfrak{B}F=\mathsf{L}F-\mathfrak{A}F$. In Case (b), we consider $\mathfrak{A}$

and $\mathfrak{B}$ such that

$$\mathfrak{A}F(v) = M(v) \int_{\mathbb{R}^d} F(v')\,dv',$$
$$\mathfrak{B}F(v) = -\int_{\mathbb{R}^d} M(v')\,dv'\,F(v).$$

The semigroup generated by $\mathfrak{A}+\mathfrak{B}$ is exponentially decreasing in $\mathscr{B}_1$ by the microscopic coercivity property, as above. The semigroup generated by $\mathfrak{B}$ is exponentially decreasing in $\mathscr{B}_2$. In Case (b), it is straightforward. In Case (a), $F(t) = e^{\mathfrak{B}t}F_0$ is such that

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}^d} |F|^2(1+|v|^q)\,d\gamma_\infty \\
&= \int_{\mathbb{R}^d} (\mathfrak{B}F)F(1+|v|^q)\,d\gamma_\infty \\
&= \int_{\mathbb{R}^d} \nabla_v\Big(M\nabla_v\Big(\frac{F}{M}\Big)\Big)F(1+|v|^q)\,d\gamma_\infty - \int_{\mathbb{R}^d} N\chi_R(v)|F|^2(1+|v|^q)\,d\gamma_\infty \\
&= -\int_{\mathbb{R}^d}\Big|\nabla_v\Big(\frac{F}{M}\Big)\Big|^2(1+|v|^q)M\,dv - \int_{\mathbb{R}^d} q|v|^{q-2}v\cdot\nabla_v\Big(\frac{F}{M}\Big)\frac{F}{M}M\,dv - \int_{\mathbb{R}^d} N\chi_R(v)|F|^2(1+|v|^q)\frac{dv}{M} \\
&\le \int_{\mathbb{R}^d}\Big\{\frac{q}{2}\frac{\nabla_v\cdot(|v|^{q-2}vM)}{(1+|v|^q)M} - N\chi_R(v)\Big\}|F|^2(1+|v|^q)\frac{dv}{M} \le -\frac{\lambda}{2}\int_{\mathbb{R}^d}|F|^2(1+|v|^q)\,d\gamma_\infty
\end{aligned}$$

for some $\lambda > 0$, by choosing N and R large enough.

The operator $\mathfrak{A} : \mathscr{B}_1 \to \mathscr{B}_2$ is bounded. This is straightforward in Case (a) and follows from the boundedness of $\int_{\mathbb{R}^d} M(v)(1+|v|^q)\,d\gamma_\infty$ in Case (b). Proposition 7 applies, which concludes the proof. □

(2) *Induction*. Let us assume that (22) is true for some $j \ge 0$, consider $P \in \mathbb{R}_{j+1}[X]$, and observe that $P[f]$ solves

$$\partial_t P[f] = \mathsf{L}P[f] - \int_{\mathbb{R}^d} (v\cdot\nabla_x P)f\,dx.$$

Since $\nabla_x P \in \mathbb{R}_j[X]$, the induction hypothesis at step j (applied with q replaced by $q+2$) gives

$$\begin{aligned}
\Big\| v\int_{\mathbb{R}^d} (\nabla_x P)[f]\,dx\Big\|_{\mathrm{L}^2(d\gamma_{k+q})} &\lesssim \Big\|\int_{\mathbb{R}^d} (\nabla_x P)[f]\,dx\Big\|_{\mathrm{L}^2(d\gamma_{k+q+2})} \\
&\lesssim_{j,q} \|f_0\|_{\mathrm{L}^2(d\gamma_{k+q+2(j+1)};\mathrm{L}^1((1+|x|^j)\,dx))}(1+t)^j e^{-\lambda t}.
\end{aligned}$$

By Duhamel's formula, we have

$$P[f](t,v) = e^{\mathsf{L}t}P[f](0,v) - \int_0^t e^{\mathsf{L}(t-s)}\Big(v\int_{\mathbb{R}^d}(\nabla_x P)[f_s]\,dx\Big)\,ds.$$

Note that

$$\int_{\mathbb{R}^d} v\int_{\mathbb{R}^d}(\nabla_x P)[f]\,dx\,dv = \iint_{\mathbb{R}^d\times\mathbb{R}^d}(v\cdot\nabla_x P)[f]\,dx\,dv = 0$$

for all $t \geq 0$ since $v \cdot \nabla_x P \in \mathbb{R}_\ell[X, V]$. As a consequence, the decay of the semigroup associated with L can be estimated by

$$\left\| e^{\mathsf{L}(t-s)} \left(v \int_{\mathbb{R}^d} (\nabla_x P)[f_s] \, dx \right) \right\|_{\mathrm{L}^2(d\gamma_\infty)} \leq \left\| v \int_{\mathbb{R}^d} (\nabla_x P)[f_s] \, dx \right\|_{\mathrm{L}^2(d\gamma_\infty)} e^{-\lambda_m (t-s)}.$$

As in the case $j = 0$, we deduce from Corollary 8 that

$$\begin{aligned} \left\| e^{\mathsf{L}(t-s)} \left(v \int_{\mathbb{R}^d} (\nabla_x P)[f_s] \, dx \right) \right\|_{\mathrm{L}^2((1+|v|^q)\, d\gamma_k)} &\leq \left\| v \int_{\mathbb{R}^d} (\nabla_x P)[f_s] \, dx \right\|_{\mathrm{L}^2(d\gamma_{k+q})} e^{-\lambda(t-s)} \\ &\lesssim_{q,k} \| f_0 \|_{\mathrm{L}^2(d\gamma_{k+q+2(j+1)}; \mathrm{L}^1((1+|x|^j)\, dx))} (1+s)^j e^{-\lambda t}. \end{aligned}$$

Moreover, since

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} f_0(x, v) P(x) \, dx \, dv = 0,$$

for the same reasons we also have

$$\| e^{\mathsf{L}t} P[f](0, \cdot) \|_{\mathrm{L}^2(d\gamma_{k+q})} \leq \| P[f_0] \|_{\mathrm{L}^2((1+|v|^q)\, d\gamma_k)} e^{-\lambda t}$$

for some $\lambda > 0$. We deduce from Duhamel's formula that

$$\begin{aligned} &\| P[f] \|_{\mathrm{L}^2(d\gamma_{k+q})} \\ &\quad \lesssim \| e^{\mathsf{L}t} P[f](0, \cdot) \|_{\mathrm{L}^2(d\gamma_{k+q})} + \int_0^t \left\| e^{-\mathsf{L}(t-s)} \left(v \int_{\mathbb{R}^d} \nabla_x P[f_s] \, dx \right) \right\|_{\mathrm{L}^2(d\gamma_{k+q})} ds \\ &\quad \lesssim_k \| f_0 \|_{\mathrm{L}^2(d\gamma_{k+q}; \mathrm{L}^1((1+|x|^{j+1})\, dx))} e^{-\lambda t} + \int_0^t (1+s)^j e^{-\lambda t} \| f_0 \|_{\mathrm{L}^2(d\gamma_{k+q+2(j+1)}; \mathrm{L}^1((1+|x|^j)\, dx))} \, ds \\ &\quad \lesssim_k \| f_0 \|_{\mathrm{L}^2(d\gamma_{k+q+2(j+1)}; \mathrm{L}^1((1+|x|^{j+1})\, dx))} (1+t)^{j+1} e^{-\lambda t}, \end{aligned}$$

which proves the induction.

Step 3: improved decay of f. Let us choose some $t_0 > 0$. In order to estimate

$$\| f(t, \cdot, \cdot) \|^2_{\mathrm{L}^2(dx\, d\gamma_k)} = \| e^{(\mathsf{L}-\mathsf{T})t} f_0 \|^2_{\mathrm{L}^2(dx\, d\gamma_k)},$$

we compute its evolution on $(0, 2t_0)$ and split the interval on $(0, t_0)$ and $(t_0, 2t_0)$ using the semigroup property

$$\| e^{(\mathsf{L}-\mathsf{T})(2t_0)} f_0 \|^2_{\mathrm{L}^2(dx\, d\gamma_k)} = \| e^{(\mathsf{L}-\mathsf{T})t_0} (e^{(\mathsf{L}-\mathsf{T})t_0} f_0) \|^2_{\mathrm{L}^2(dx\, d\gamma_k)}.$$

Up to the end of this section, $\mathsf{T} = v \cdot \nabla_x$ denotes the transport operator in position and velocity variables. We decompose $f_{t_0} = e^{(\mathsf{L}-\mathsf{T})t_0} f_0$ into

$$f_{t_0} = \left(\sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} X^\alpha [f_{t_0}] \partial^\alpha \varphi \right) + g_0, \quad \text{with } g_0 := f_{t_0} - \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} X^\alpha [f_{t_0}] \partial^\alpha \varphi,$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_i, \ldots, \alpha_d) \in \mathbb{N}^d$ is a multi-index such that $|\alpha| = \sum_{i=1}^d \alpha_i \leq \ell$ and φ is given by

$$\varphi(x) := (2\pi)^{-\frac{d}{2}} e^{-\frac{|x|^2}{2}} \quad \text{for all } x \in \mathbb{R}^d.$$

Here we use the notation $\partial^\alpha \varphi = \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \cdots \partial_{x_d}^{\alpha_d} \varphi$ and $X^\alpha = \prod_i^n X_i^{\alpha_i}$. According to (22), we know that

$$\|X^\alpha[f_{t_0}]\|_{\mathrm{L}^2(d\gamma_k)} \lesssim_j \|f_0\|_{\mathrm{L}^2(d\gamma_{k+2j};\mathrm{L}^1((1+|x|^j)\,dx))} (1+t_0)^j e^{-\lambda t_0},$$

so that, by considering the evolution of the first term on $(t_0, 2t_0)$, we obtain

$$\left\| e^{(\mathsf{L}-\mathsf{T})t_0} \left(\sum_{|\alpha| \le \ell} \frac{1}{\alpha!} X^\alpha[f_{t_0}] \partial^\alpha \varphi \right) \right\|_{\mathrm{L}^2(dx\,d\gamma_k)} \lesssim \sum_{|\alpha| \le \ell} \|X^\alpha[f_{t_0}]\|_{\mathrm{L}^2(d\gamma_k)} \|\partial^\alpha \varphi\|_{\mathrm{L}^2(dx)} \lesssim e^{-\frac{\lambda}{2} t_0}. \tag{24}$$

Next, let us consider the second term and define, on $t + t_0 \in (t_0, 2t_0)$, the function

$$g := f_{t+t_0} - \sum_{|\alpha| \le \ell} \frac{1}{\alpha!} X^\alpha[f_{t+t_0}] \partial^\alpha \varphi.$$

With initial datum g_0, it solves on $(0, t_0)$ the equation

$$\begin{aligned}
\partial_t g &= \partial_t f_{t+t_0} - \sum_{|\alpha| \le \ell} \frac{1}{\alpha!} \partial_t (X^\alpha[f_{t+t_0}])\, \partial^\alpha \varphi \\
&= (\mathsf{L}-\mathsf{T})(f_{t+t_0}) - \mathsf{L}\left(\sum_{|\alpha| \le \ell} \frac{1}{\alpha!} X^\alpha[f_{t+t_0}]\, \partial^\alpha \varphi \right) + \sum_{|\alpha| \le \ell} \frac{1}{\alpha!} \left(\int_{\mathbb{R}^d} (v \cdot \nabla_x x^\alpha) f_{t+t_0}\, dx \right) \partial^\alpha \varphi \\
&= (\mathsf{L}-\mathsf{T})(g) - \mathsf{T}\left(\sum_{|\alpha| \le \ell} \frac{1}{\alpha!} X^\alpha[f_{t+t_0}]\, \partial^\alpha \varphi \right) + \sum_{|\alpha| \le \ell} \frac{1}{\alpha!} \left(\int_{\mathbb{R}^d} (v \cdot \nabla_x x^\alpha) f_{t+t_0}\, dx \right) \partial^\alpha \varphi \\
&= (\mathsf{L}-\mathsf{T})(g) + v \sum_{|\alpha| \le \ell} \frac{1}{\alpha!} (\nabla_x X^\alpha[f]\, \partial^\alpha \varphi - X^\alpha[f_{t+t_0}] \nabla_x (\partial^\alpha \varphi)),
\end{aligned}$$

where $\alpha! = \prod_{i=1}^d \alpha_i!$ is associated with the multi-index $\alpha = (\alpha_i)_{i=1}^d$ and

$$\nabla_x X^\alpha[f] = (\partial_{x_i} X^\alpha[f])_{i=1}^d := \left(\int_{\mathbb{R}^d} \partial_{x_i} x^\alpha f\, dx \right)_{i=1}^d = \left(\int_{\mathbb{R}^d} \alpha_i x^{\alpha_{\wedge i}} f\, dx \right)_{i=1}^d.$$

Here the notation $\alpha_{\wedge i}$ denotes the multi-index $(\alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \alpha_i - 1, \alpha_{i+1}, \ldots, \alpha_d)$ with the convention that $X^{\alpha_{\wedge i}} \equiv 0$ if $\alpha_i = 0$. We also define the opposite transformation

$$\alpha_{\vee i} := (\alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \alpha_i + 1, \alpha_{i+1}, \ldots, \alpha_d)$$

so that $\partial_{x_i}(\partial^\alpha \varphi) = \partial^{\alpha_{\vee i}} \varphi$. Let us consider the last term and start with the case $d = 1$. In that case,

$$\begin{aligned}
v \sum_{|\alpha| \le \ell} & \frac{1}{\alpha!} (\nabla_x X^\alpha[f]\, \partial^\alpha \varphi - X^\alpha[f_{t+t_0}] \nabla_x (\partial^\alpha \varphi)) \\
&= v_1 \sum_{\alpha_1 = 0}^{\ell} \frac{1}{\alpha_1!} \left(\left(\int_{\mathbb{R}} (\alpha_1 x^{\alpha_1 - 1}) f_{t+t_0}\, dx \right) \partial x_1^{\alpha_1} \varphi - \left(\int_{\mathbb{R}} x^{\alpha_1} f_{t+t_0}\, dx \right) \partial x_1^{\alpha_1 + 1} \varphi \right) \\
&= -\frac{v_1}{\ell!} \left(\int_{\mathbb{R}} x^\ell f_{t+t_0}\, dx \right) \partial x_1^{\ell+1} \varphi
\end{aligned}$$

because it is a telescoping sum. We adopt the convention that $\alpha! = 1$ if $\alpha_i \le 0$ for some $i = 1, 2, \ldots, d$. The same property holds in higher dimensions:

$$\sum_{|\alpha|\le\ell} \frac{1}{\alpha!}(\partial_{x_i} X^\alpha[f]\partial^\alpha\varphi - X^\alpha[f_{t+t_0}]\partial_{x_i}(\partial^\alpha\varphi))$$
$$= \sum_{|\alpha|\le\ell}\left(\frac{1}{\alpha_{\wedge i}!}X^{\alpha_{\wedge i}}[f]\partial^\alpha\varphi - \frac{1}{\alpha!}X^\alpha[f_{t+t_0}]\partial^{\alpha\vee i}\varphi\right) = -\sum_{|\alpha|=\ell}\frac{1}{\alpha!}X^\alpha[f_{t+t_0}]\,\partial_{x_i}(\partial^\alpha\varphi).$$

We deduce that

$$\partial_t g = (\mathsf{L}-\mathsf{T})(g) - v\sum_{|\alpha|=\ell}\frac{1}{\alpha!}X^\alpha[f_{t+t_0}]\nabla_x(\partial^\alpha\varphi).$$

Duhamel's formula in Fourier variables gives

$$\hat g(t_0,\xi,v) = e^{(\mathsf{L}-\mathsf{T})t_0}\hat g_0 - \int_0^{t_0} e^{(\mathsf{L}-\mathsf{T})(t_0-s)}\left(v\sum_{|\alpha|=\ell}\frac{1}{\alpha!}X^\alpha[f_{s+t_0}]\widehat{\nabla_x(\partial^\alpha\varphi)}\right)ds$$

up to a straightforward abuse of notation. Hence

$$\|\hat g(t_0,\xi,\cdot)\|_{\mathrm{L}^2(d\gamma_k)}$$
$$\lesssim e^{-\frac12\mu_\xi t_0}\|\hat g_0(\xi,\cdot)\|_{\mathrm{L}^2(d\gamma_k)} + \int_0^{t_0} e^{-\frac{\mu_\xi}{2}(t_0-s)}\sum_{|\alpha|=\ell}\frac{1}{\alpha!}\|X^\alpha[f_{s+t_0}]\|_{\mathrm{L}^2(|v|^2\,d\gamma_k)}|\widehat{\nabla_x(\partial^\alpha\varphi)}|\,ds.$$

Recall that (22) gives

$$\|X^\alpha[f_{s+t_0}]\|_{\mathrm{L}^2(|v|^2\,d\gamma_k)} \lesssim_\ell \|f_0\|_{\mathrm{L}^2(d\gamma_{k+2\ell+2};\mathrm{L}^1((1+|x|^\ell)\,dx))}e^{-\frac{\lambda}{2}s}.$$

On the other hand we use $|\widehat{\nabla_x(\partial^\alpha\varphi)}| \le |\xi|^{\ell+1}|\hat\varphi|$ and observe that

$$|\hat g_0(\xi,v)| \lesssim |\xi|^{\ell+1}\|g_0(\cdot,v)\|_{\mathrm{L}^1(|x|^\ell\,dx)} \quad \text{for all } (\xi,v)\in\mathbb{R}^d\times\mathbb{R}^d.$$

Collecting terms, we have

$$\|\hat g(t_0,\xi,\cdot)\|_{\mathrm{L}^2(d\gamma_k)} \lesssim e^{-\frac12\mu_\xi t_0}|\xi|^{\ell+1}\mathbf{1}_{|\xi|<1}\|g_0(\cdot,v)\|_{\mathrm{L}^2(d\gamma_k;\mathrm{L}^1(|x|^\ell\,dx))} + e^{-\frac12\mu_\xi t_0}\mathbf{1}_{|\xi|\ge1}\|\hat g_0(\xi,\cdot)\|_{\mathrm{L}^2(d\gamma_k)}$$
$$+|\xi|^{\ell+1}|\hat\varphi(\xi)|\,\|f_0\|_{\mathrm{L}^2(d\gamma_{k+2\ell+2};\mathrm{L}^1((1+|x|^\ell)\,dx))}\int_0^{t_0} e^{-\frac{\mu_\xi}{2}(t_0-s)}e^{-\frac{\lambda}{2}s}\,ds.$$

We know from (10) that $\mu_\xi = \Lambda|\xi|^2/(1+|\xi|^2)$ so that $\mu_\xi \ge \frac12\Lambda|\xi|^2$ if $|\xi|<1$ and $\mu_\xi \ge \frac12\Lambda$ if $|\xi|\ge1$. Hence, for any $t_0 \ge 1$,

$$\|e^{-\frac12\mu_\xi t_0}|\xi|^{\ell+1}\mathbf{1}_{|\xi|<1}\|_{\mathrm{L}^2(d\xi)} \le \left(\int_{\mathbb{R}^d} e^{-\frac{\Lambda}{2}|\xi|^2 t_0}|\xi|^{2(\ell+1)}\,d\xi\right)^{\frac12} \lesssim t_0^{-(1+\ell+\frac{d}{2})},$$

$$\int_{|\xi|\ge1} e^{-\mu_\xi t_0}\|\hat g_0(\xi,\cdot)\|^2_{\mathrm{L}^2(d\gamma_k)}\,d\xi \lesssim e^{-\frac{\Lambda}{2}t_0}\|g_0\|^2_{\mathrm{L}^2(dx\,d\gamma_k)}$$

by Plancherel's formula. We conclude by observing that

$$\int_{|\xi|\le 1} |\xi|^{\ell+1} |\hat{\varphi}(\xi)| \int_0^{t_0} e^{-\frac{\mu_\xi}{2}(t_0-s)} e^{-\frac{\lambda}{2}s}\, ds\, d\xi \le \|\varphi\|_{\mathrm{L}^1(dx)} \int_0^{t_0} \left(\int_{|\xi|\le 1} |\xi|^{\ell+1} e^{-\frac{\Lambda}{2}|\xi|^2(t_0-s)}\, d\xi \right) e^{-\frac{\lambda}{2}s}\, ds$$
$$\lesssim t_0^{-(1+\ell+\frac{d}{2})},$$

$$\int_{|\xi|\ge 1} |\xi|^{\ell+1} |\hat{\varphi}(\xi)| \int_0^{t_0} e^{-\frac{\mu_\xi}{2}(t_0-s)} e^{-\frac{\lambda}{2}s}\, ds\, d\xi \lesssim \|\, |\xi|^{\ell+1} \hat{\varphi}(\xi)\|_{\mathrm{L}^1(d\xi)} t_0 e^{-\frac{1}{4}\min\{\Lambda, 2\lambda\} t_0}.$$

Altogether, we obtain

$$\|g(t_0,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\, d\gamma_k)} = \|\hat{g}(t_0,\cdot,\cdot)\|^2_{\mathrm{L}^2(d\xi\, d\gamma_k)} \lesssim t_0^{-(1+\ell+\frac{d}{2})}.$$

The decay result of Theorem 3 is then obtained by writing

$$\|f_{2t_0}\|^2_{\mathrm{L}^2(dx\, d\gamma_k)} \lesssim \|g(t_0,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\, d\gamma_k)} + \left\| e^{(\mathsf{L}-\mathsf{T})t_0} \left(\sum_{|\alpha|\le \ell} \frac{1}{\alpha!} X^\alpha [f_{t_0}]\, \partial^\alpha \varphi \right) \right\|_{\mathrm{L}^2(dx\, d\gamma_k)}$$

and using (24) for any $t_0 \ge 1$, with $t = 2t_0$. For $t \le 2$, the estimate of Theorem 3 is straightforward by Corollary 8, which concludes the proof.

Appendix A: An explicit computation of Green's function for the kinetic Fokker–Planck equation and consequences

In the whole-space case, when M is the normalized Gaussian function, let us consider the kinetic Fokker–Planck equation of Case (a)

$$\partial_t f + v \cdot \nabla_x f = \nabla_v \cdot (v f + \nabla_v f) \tag{25}$$

on $(0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d \ni (t,x,v)$. The characteristics associated with the equations

$$\frac{dx}{dt} = v, \quad \frac{dv}{dt} = -v$$

suggest to change variables and consider the distribution function g such that

$$f(t,x,v) = e^{dt} g(t, x + (1-e^t)v, e^t v) \quad \text{for all } (t,x,v) \in (0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d.$$

The kinetic Fokker–Planck equation is changed into a heat equation in both variables x and v with t dependent coefficients, which can be written as

$$\partial_t g = \nabla \cdot \dot{\mathscr{D}}\, \nabla g, \tag{26}$$

where $\nabla g = (\nabla_v g, \nabla_x g)$ and $\dot{\mathscr{D}}$ is the t-derivative of the block matrix

$$\mathscr{D} = \frac{1}{2} \begin{pmatrix} \mathsf{a}\,\mathrm{Id} & \mathsf{b}\,\mathrm{Id} \\ \mathsf{b}\,\mathrm{Id} & \mathsf{c}\,\mathrm{Id} \end{pmatrix},$$

with $\mathsf{a} = e^{2t} - 1$, $\mathsf{b} = 2e^t - 1 - e^{2t}$, and $\mathsf{c} = e^{2t} - 4e^t + 2t + 3$. Here Id is the identity matrix on $\mathbb{R}^d$. We observe that $\dot{\mathscr{D}}$ is degenerate: it is nonnegative but its lowest eigenvalue is 0. However, the change of variables allows the computation of a Green's function.

Lemma 11. *The Green's function of (26) is given for any* $(t, x, v) \in (0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$ *by*

$$G(t, x, v) = \frac{1}{(2\pi(\mathsf{ac} - \mathsf{b}^2))^{\frac{d}{2}}} \exp\left(-\frac{\mathsf{a}|x|^2 - 2\mathsf{b}x \cdot v + \mathsf{c}|v|^2}{2(\mathsf{ac} - \mathsf{b}^2)}\right).$$

The method is standard and goes back to [Kolmogoroff 1934] (also see [Ilin and Hasminskii 1964; Hörmander 1967; Victory and O'Dwyer 1990; Bouchut 1993]).

Proof. By a Fourier transformation in x and v, with associated variables ξ and η, we find that

$$\log C - \log \widehat{G}(t, \xi, \eta) = (\eta, \xi) \cdot \mathscr{D}(\eta, \xi) = \tfrac{1}{2}(\mathsf{a}|\eta|^2 + 2\mathsf{b}\eta \cdot \xi + \mathsf{c}|\xi|^2) = \tfrac{1}{2}\mathsf{a}\left|\eta + \tfrac{\mathsf{b}}{\mathsf{a}}\xi\right|^2 + \tfrac{1}{2}\mathsf{A}|\xi|^2, \quad \mathsf{A} = \mathsf{c} - \tfrac{\mathsf{b}^2}{\mathsf{a}},$$

for some constant $C > 0$ which is determined by the mass normalization condition

$$\|G(t, \cdot, \cdot)\|_{\mathrm{L}^1(\mathbb{R}^d \times \mathbb{R}^d)} = 1.$$

Let us take the inverse Fourier transform with respect to η,

$$(2\pi)^{-d} \int_{\mathbb{R}^d} e^{iv \cdot \eta} \widehat{G}(t, \xi, \eta)\, d\eta = \frac{C}{(2\pi\mathsf{a})^{\frac{d}{2}}} e^{-\frac{|v|^2}{2\mathsf{a}} - i\frac{\mathsf{b}}{\mathsf{a}} v \cdot \xi} e^{-\frac{1}{2}\mathsf{A}|\xi|^2} = \frac{C}{(2\pi\mathsf{a})^d} e^{-\frac{|v|^2}{2\mathsf{a}}} e^{-\frac{1}{2}\mathsf{A}\left|\xi + i\frac{\mathsf{b}}{\mathsf{aA}} v\right|^2 - \frac{\mathsf{b}^2}{2\mathsf{a}^2\mathsf{A}}|v|^2},$$

and then the inverse Fourier transform with respect to ξ, so that we obtain

$$G(t, x, v) = \frac{C}{(2\pi\mathsf{a})^{\frac{d}{2}} (2\pi\mathsf{A})^{\frac{d}{2}}} e^{-\left(1 + \frac{\mathsf{b}^2}{\mathsf{aA}}\right)\frac{|v|^2}{2\mathsf{a}}} e^{-\frac{|x|^2}{2\mathsf{A}}} e^{\frac{\mathsf{b}}{\mathsf{aA}} x \cdot v} = \frac{C}{(4\pi^2\mathsf{aA})^{\frac{d}{2}}} e^{-\frac{1}{2\mathsf{A}}\left|x - \frac{\mathsf{b}}{\mathsf{a}} v\right|^2} e^{-\frac{|v|^2}{2\mathsf{a}}}.$$

It is easy to check that $C = 1$. □

Let us consider a solution g of (26) with initial datum $g_0 \in \mathrm{L}^1(\mathbb{R}^d \times \mathbb{R}^d)$. From the representation

$$g(t, \cdot, \cdot) = G(t, \cdot, \cdot) *_{x,v} g_0,$$

we obtain the estimate

$$\|g(t, \cdot, \cdot)\|_{\mathrm{L}^\infty(\mathbb{R}^d \times \mathbb{R}^d)} \le \|G(t, \cdot, \cdot)\|_{\mathrm{L}^\infty(\mathbb{R}^d \times \mathbb{R}^d)} \|g_0\|_{\mathrm{L}^1(\mathbb{R}^d \times \mathbb{R}^d)} = \frac{\|g_0\|_{\mathrm{L}^1(\mathbb{R}^d \times \mathbb{R}^d)}}{(8\pi^2)^{\frac{d}{2}}} t^{-\frac{d}{2}} e^{-dt} (1 + O(t^{-1}))$$

as $t \to \infty$. As a consequence, we obtain that the solution of (25) with a nonnegative initial datum f_0 satisfies

$$\|f(t, \cdot, \cdot)\|_{\mathrm{L}^\infty(\mathbb{R}^d \times \mathbb{R}^d)} = \frac{\|f_0\|_{\mathrm{L}^1(\mathbb{R}^d \times \mathbb{R}^d)}}{(8\pi^2 t)^{\frac{d}{2}}} (1 + o(1)) \quad \text{as } t \to \infty.$$

Using the simple Hölder interpolation inequality

$$\|f\|_{\mathrm{L}^p(\mathbb{R}^d \times \mathbb{R}^d)} \le \|f\|_{\mathrm{L}^1(\mathbb{R}^d \times \mathbb{R}^d)}^{\frac{1}{p}} \|f\|_{\mathrm{L}^\infty(\mathbb{R}^d \times \mathbb{R}^d)}^{1 - \frac{1}{p}},$$

we obtain the following decay result.

Corollary 12. *If f is a solution of (25) with a nonnegative initial datum $f_0 \in \mathrm{L}^1(\mathbb{R}^d \times \mathbb{R}^d)$, then for any $p \in (1, \infty]$ we have the decay estimate*

$$\|f(t,\cdot,\cdot)\|_{\mathrm{L}^p(\mathbb{R}^d\times\mathbb{R}^d)} \le \frac{\|f_0\|_{\mathrm{L}^1(\mathbb{R}^d\times\mathbb{R}^d)}}{(8\pi^2 t)^{\frac{d}{2}(1-\frac{1}{p})}}(1+o(1)) \quad \text{as } t \to \infty.$$

By taking $f_0(x,v) = G(1,x,v)$, it is moreover straightforward to check that this estimate is optimal. With $p = 2$, this also proves that the decay rate obtained in Theorem 1 for the Fokker–Planck operator, i.e., Case (a), is the optimal one because, again with $f_0(x,v) = G(1,x,v)$, we observe that

$$\|f(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,d\gamma_k)} = e^{dt}\|G(t,\cdot,\cdot)\|^2_{\mathrm{L}^2(dx\,dv)} = O(t^{-\frac{d}{2}}) \quad \text{as } t \to +\infty.$$

Appendix B: Consistency with the decay rates of the heat equation

In the whole-space case, the abstract approach of [Dolbeault et al. 2015] is inspired by the diffusion limit of (1). We consider the scaled equation

$$\varepsilon\frac{dF}{dt} + \mathsf{T}F = \frac{1}{\varepsilon}\mathsf{L}F, \tag{27}$$

which formally corresponds to a parabolic rescaling given by $t \mapsto \varepsilon^2 t$ and $x \mapsto \varepsilon x$, and investigate the limit as $\varepsilon \to 0_+$. Let us check that the rates are asymptotically independent of ε and consistent with those of the heat equation.

B.1. ***Mode-by-mode hypocoercivity.*** It is straightforward to check that in the estimate (7) for λ, the gap constant λ_m has to be replaced by λ_m/ε, while, with the notation of Proposition 4, C_M can be replaced by C_M/ε for $\varepsilon < 1$. In the asymptotic regime as $\varepsilon \to 0_+$, we obtain

$$\varepsilon\frac{d}{dt}\mathsf{H}[F] \le -\mathsf{D}[F] \le -\frac{\lambda_M}{3(1+\lambda_M)}\frac{\lambda_m\lambda_M\varepsilon}{(1+\lambda_M)C_M^2}\mathsf{D}[F],$$

which proves that the estimate of Proposition 4 becomes

$$\lambda \ge \frac{\lambda_m\lambda_M^2}{3(1+\lambda_M)^2C_M^2}.$$

We observe that this rate is independent of ε.

B.2. ***Decay rates based on Nash's inequality in the whole-space case.*** In the proof of Theorem 1, $\bar\sigma$ has to be replaced by $\bar\sigma/\varepsilon$ and in the limit as $\varepsilon \to 0_+$, we get that $\mathsf{b} \sim 4\bar\sigma/\varepsilon$ and (16) is satisfied with

$$4\mathsf{a} = \delta \sim \frac{\lambda_m}{8\bar\sigma^2}\varepsilon.$$

Hence (18) asymptotically becomes, as $\varepsilon \to 0_+$,

$$-\frac{d}{dt}\mathsf{H}[f] \ge \frac{\lambda_m}{4\bar\sigma^2}\mathsf{c}\left(\frac{2}{1+\delta}\mathsf{H}[f]\right)^{1+\frac{2}{d}},$$

which again gives a rate of decay which is independent of ε. The algebraic decay rate in Theorem 1 is the one of the heat equation on $\mathbb{R}^d$ and it is independent of ε in the limit as $\varepsilon \to 0_+$.

B.3. ***Decay rates in the whole-space case for distribution functions with moment cancellations.*** The improved rate of Theorem 2 is consistent with a parabolic rescaling: if f solves (1), then $f^\varepsilon(t,x,v) = \varepsilon^{-d} f(\varepsilon^{-2}t, \varepsilon^{-1}x, v)$ solves (27). With the notation of Section 6.1, let $g^\varepsilon = f^\varepsilon - f_\bullet^\varepsilon \varphi(\cdot/\varepsilon)$, with $\varphi^\varepsilon = \varepsilon^{-d}\varphi(\cdot/\varepsilon)$. The Fourier transform of g^ε solves

$$\varepsilon^2\,\partial_t \hat g^\varepsilon + \varepsilon \mathsf{T} \hat g^\varepsilon = \mathsf{L} \hat g^\varepsilon - \varepsilon f_\bullet^\varepsilon \mathsf{T} \hat\varphi^\varepsilon .$$

The decay rate λ in (20) becomes λ/ε^2 and the decay rate of the semigroup generated by $\mathsf{L} - \varepsilon\mathsf{T}$ is, with the notation of Corollary 5, $\mu_{\varepsilon\xi}$. Moreover, Λ in (10) is given by $\Lambda = \frac13 \min\{1,\Theta\}$ for any $\varepsilon > 0$, small enough. Duhamel's formula (21) has to be replaced by

$$\|\hat g^\varepsilon(t,\xi,\cdot)\|_{\mathrm{L}^2(d\gamma_k)} \le C e^{-\frac{\mu_{\varepsilon\xi}}{2\varepsilon^2}t} \|\hat g_0^\varepsilon(\xi,\cdot)\|_{\mathrm{L}^2(d\gamma_k)} + C\int_0^t e^{-\frac{\mu_{\varepsilon\xi}}{2\varepsilon^2}(t-s)} \|f_\bullet^\varepsilon(s,\cdot)\|_{\mathrm{L}^2(|v|^2\,d\gamma_k)} |\varepsilon\xi|\,|\hat\varphi(\varepsilon\xi)|\,ds.$$

Using

$$\lim_{\varepsilon\to 0_+} \frac{\mu_{\varepsilon\xi}}{\varepsilon^2} = \lim_{\varepsilon\to 0_+} \frac{\Lambda|\xi|^2}{1+\varepsilon^2|\xi|^2} = \Lambda|\xi|^2,$$

a computation similar to the one of Section 6.1 shows that the first term of the right-hand side is estimated by

$$\begin{aligned}\int_{\mathbb{R}^d} e^{-\frac{\mu_{\varepsilon\xi}}{\varepsilon^2}t} \|\hat g_0^\varepsilon(\xi,\cdot)\|^2_{\mathrm{L}^2(d\gamma_k)}\,d\xi &= \int_{|\xi|\le\frac1\varepsilon} e^{-\frac{\mu_{\varepsilon\xi}}{\varepsilon^2}t} \|\hat g_0^\varepsilon(\xi,\cdot)\|^2_{\mathrm{L}^2(d\gamma_k)}\,d\xi + \int_{|\xi|>\frac1\varepsilon} e^{-\frac{\mu_{\varepsilon\xi}}{\varepsilon^2}t} \|\hat g_0^\varepsilon(\xi,\cdot)\|^2_{\mathrm{L}^2(d\gamma_k)}\,d\xi\\ &\le \|g_0^\varepsilon\|^2_{\mathrm{L}^2(d\gamma_k;\mathrm{L}^1(|x|\,dx))} \int_{\mathbb{R}^d} |\xi|^2 e^{-\frac{\Lambda}{2}|\xi|^2 t}\,d\xi + \|g_0^\varepsilon\|^2_{\mathrm{L}^2(dx\,d\gamma_k)} e^{-\frac{\Lambda}{2}\frac{t}{\varepsilon^2}},\end{aligned}$$

while the square of the second term is bounded by

$$\begin{aligned}&\|f_\bullet^\varepsilon(t=0,\cdot)\|^2_{\mathrm{L}^2(|v|^2\,d\gamma_k)} \int_{\mathbb{R}^d} |\varepsilon\xi|^2 |\hat\varphi(\varepsilon\xi)|^2 \left(\int_0^{\varepsilon^{-2}t} e^{-\frac12\mu_{\varepsilon\xi}(\varepsilon^{-2}t-s)} e^{-\frac12\lambda s}\,ds\right)^2 d\xi\\ &\qquad\le \|f_0\|^2_{\mathrm{L}^2(|v|^2\,d\gamma_k;\mathrm{L}^1(dx))} \left(C_1 \frac{\varepsilon^{d+1}}{t^{\frac d2+1}} + \frac{C_2}{\varepsilon^3} e^{-\min\{\frac{\Lambda}{2},\lambda\}\frac{t}{\varepsilon^2}}\right).\end{aligned}$$

By collecting all terms and using Plancherel's formula, we conclude that the rate of convergence of Theorem 2 applied to the solution of (27) is independent of ε. We also notice that the scaled spatial density $\rho_{f^\varepsilon} = \int_{\mathbb{R}^d} f^\varepsilon\,dv$ satisfies

$$\|\rho_{f^\varepsilon}(t,\cdot)\|^2_{\mathrm{L}^2(dx)} \le \frac{\mathscr{C}_0}{(1+t)^{1+\frac d2}} \quad \text{for all } t \ge 0$$

for some positive constant $\mathscr{C}_0$ which depends on f_0 but is independent of ε. This is the decay of the heat equation with an initial datum of zero average.

Similar estimates can be obtained in the framework of Theorem 3.

Acknowledgments

The French National Research Agency (ANR) provided partial support of Bouin and Dolbeault through Projects EFI (ANR-17-CE40-0030) and Kibord (ANR-13-BS01-0004) and of Dolbeaut through Project STAB (ANR-12-BS01-0019). Schmeiser was supported by the Austrian Science Foundation (grants F65 and W1245), by the Fondation Sciences Mathématiques de Paris, and by Paris Science et Lettres. Mouhot and Bouin acknowledge partial funding by the ERC grants MATKIT 2011-2016 and MAFRAN 2017-2022. Moreover Mouhot and Schmeiser are very grateful for the hospitality at Université Paris-Dauphine.

References

[Achleitner et al. 2016] F. Achleitner, A. Arnold, and E. A. Carlen, "On linear hypocoercive BGK models", pp. 1–37 in *From particle systems to partial differential equations, III*, edited by P. Gonçalves and A. J. Soares, Springer Proc. Math. Stat. **162**, Springer, 2016. MR Zbl

[Bakry et al. 2008] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin, "A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case", *Electron. Commun. Probab.* **13** (2008), 60–66. MR Zbl

[Bartier et al. 2011] J.-P. Bartier, A. Blanchet, J. Dolbeault, and M. Escobedo, "Improved intermediate asymptotics for the heat equation", *Appl. Math. Lett.* **24**:1 (2011), 76–81. MR Zbl

[Beckner 1989] W. Beckner, "A generalized Poincaré inequality for Gaussian measures", *Proc. Amer. Math. Soc.* **105**:2 (1989), 397–400. MR Zbl

[Bhatnagar et al. 1954] P. L. Bhatnagar, E. P. Gross, and M. Krook, "A model for collision processes in gases, I: Small amplitude processes in charged and neutral one-component systems", *Phys. Rev.* **94**:3 (1954), 511–525. Zbl

[Bouchut 1993] F. Bouchut, "Existence and uniqueness of a global smooth solution for the Vlasov–Poisson–Fokker–Planck system in three dimensions", *J. Funct. Anal.* **111**:1 (1993), 239–258. MR Zbl

[Bouin et al. 2017] E. Bouin, F. Hoffmann, and C. Mouhot, "Exponential decay to equilibrium for a fiber lay-down process on a moving conveyor belt", *SIAM J. Math. Anal.* **49**:4 (2017), 3233–3251. MR Zbl

[Cáceres et al. 2003] M. J. Cáceres, J. A. Carrillo, and T. Goudon, "Equilibration rate for the linear inhomogeneous relaxation-time Boltzmann equation for charged particles", *Comm. Partial Differential Equations* **28**:5-6 (2003), 969–989. MR Zbl

[Degond et al. 2000] P. Degond, T. Goudon, and F. Poupaud, "Diffusion limit for nonhomogeneous and non-micro-reversible processes", *Indiana Univ. Math. J.* **49**:3 (2000), 1175–1198. MR Zbl

[Dolbeault et al. 2009] J. Dolbeault, C. Mouhot, and C. Schmeiser, "Hypocoercivity for kinetic equations with linear relaxation terms", *C. R. Math. Acad. Sci. Paris* **347**:9-10 (2009), 511–516. MR Zbl

[Dolbeault et al. 2015] J. Dolbeault, C. Mouhot, and C. Schmeiser, "Hypocoercivity for linear kinetic equations conserving mass", *Trans. Amer. Math. Soc.* **367**:6 (2015), 3807–3828. MR Zbl

[Duan 2011] R. Duan, "Hypocoercivity of linear degenerately dissipative kinetic equations", *Nonlinearity* **24**:8 (2011), 2165–2189. MR Zbl

[Eckmann and Hairer 2003] J.-P. Eckmann and M. Hairer, "Spectral properties of hypoelliptic operators", *Comm. Math. Phys.* **235**:2 (2003), 233–253. MR Zbl

[Evans 2017] J. Evans, "Hypocoercivity in Phi-entropy for the linear relaxation Boltzmann equation on the torus", preprint, 2017. arXiv

[Glassey 1996] R. T. Glassey, *The Cauchy problem in kinetic theory*, SIAM, Philadelphia, PA, 1996. MR Zbl

[Gualdani et al. 2017] M. P. Gualdani, S. Mischler, and C. Mouhot, *Factorization of non-symmetric operators and exponential H-theorem*, Mém. Soc. Math. Fr. (N.S.) **153**, Société Mathématique de France, Paris, 2017. MR Zbl

[Han-Kwan and Léautaud 2015] D. Han-Kwan and M. Léautaud, "Geometric analysis of the linear Boltzmann equation, I: Trend to equilibrium", *Ann. PDE* **1**:1 (2015), art. id. 3. MR Zbl

[Hérau 2006] F. Hérau, "Hypocoercivity and exponential time decay for the linear inhomogeneous relaxation Boltzmann equation", *Asymptot. Anal.* **46**:3-4 (2006), 349–359. MR Zbl

[Hérau and Nier 2004] F. Hérau and F. Nier, "Isotropic hypoellipticity and trend to equilibrium for the Fokker–Planck equation with a high-degree potential", *Arch. Ration. Mech. Anal.* **171**:2 (2004), 151–218. MR Zbl

[Hörmander 1967] L. Hörmander, "Hypoelliptic second order differential equations", *Acta Math.* **119** (1967), 147–171. MR Zbl

[Iacobucci et al. 2019] A. Iacobucci, S. Olla, and G. Stoltz, "Convergence rates for nonequilibrium Langevin dynamics", *Ann. Math. Qué.* **43**:1 (2019), 73–98. MR Zbl

[Ilin and Hasminskii 1964] A. M. Ilin and R. Z. Hasminskii, "On the equations of Brownian motion", *Teor. Verojatnost. i Primenen.* **9** (1964), 466–491. In Russian; translated in *Th. Prob. Appl.* **9**:3 (1964), 421–444. MR Zbl

[Kavian and Mischler 2015] O. Kavian and S. Mischler, "The Fokker–Planck equation with subcritical confinement force", preprint, 2015. arXiv

[Kawashima 1990] S. Kawashima, "The Boltzmann equation and thirteen moments", *Japan J. Appl. Math.* **7**:2 (1990), 301–320. MR Zbl

[Kolmogoroff 1934] A. Kolmogoroff, "Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)", *Ann. of Math.* (2) **35**:1 (1934), 116–117. MR Zbl

[Mischler and Mouhot 2016] S. Mischler and C. Mouhot, "Exponential stability of slowly decaying solutions to the kinetic-Fokker–Planck equation", *Arch. Ration. Mech. Anal.* **221**:2 (2016), 677–723. MR Zbl

[Monmarché 2017] P. Monmarché, "A note on Fisher information hypocoercive decay for the linear Boltzmann equation", preprint, 2017. arXiv

[Mouhot and Neumann 2006] C. Mouhot and L. Neumann, "Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus", *Nonlinearity* **19**:4 (2006), 969–998. MR Zbl

[Nash 1958] J. Nash, "Continuity of solutions of parabolic and elliptic equations", *Amer. J. Math.* **80** (1958), 931–954. MR Zbl

[Reed and Simon 1980] M. Reed and B. Simon, *Methods of modern mathematical physics, I: Functional analysis*, 2nd ed., Academic Press, New York, 1980. MR Zbl

[Vázquez 2017] J. L. Vázquez, "Asymptotic behaviour methods for the heat equation: convergence to the Gaussian", preprint, 2017. arXiv

[Victory and O'Dwyer 1990] H. D. Victory, Jr. and B. P. O'Dwyer, "On classical solutions of Vlasov–Poisson Fokker–Planck systems", *Indiana Univ. Math. J.* **39**:1 (1990), 105–156. MR Zbl

[Villani 2006] C. Villani, "Hypocoercive diffusion operators", pp. 473–498 in *International Congress of Mathematicians*, vol. III, edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR Zbl

[Villani 2009] C. Villani, *Hypocoercivity*, Mem. Amer. Math. Soc. **950**, Amer. Math. Soc., Providence, RI, 2009. MR Zbl

Received 7 Nov 2018. Revised 20 Sep 2019. Accepted 24 Nov 2019.

EMERIC BOUIN: bouin@ceremade.dauphine.fr
Ceremade (CNRS UMR 7534), PSL University, Université Paris-Dauphine, Paris, France

JEAN DOLBEAULT: dolbeaul@ceremade.dauphine.fr
Ceremade (CNRS UMR 7534), PSL University, Université Paris-Dauphine, Paris, France

STÉPHANE MISCHLER: mischler@ceremade.dauphine.fr
Ceremade (CNRS UMR 7534), PSL University, Université Paris-Dauphine, Paris, France

CLÉMENT MOUHOT: c.mouhot@dpmms.cam.ac.uk
DPMMS, Centre for Mathematical Sciences, University of Cambridge, Cambridge, United Kingdom

CHRISTIAN SCHMEISER: christian.schmeiser@univie.ac.at
Fakultät für Mathematik, Universität Wien, Wien, Austria

PURE and APPLIED
ANALYSIS

Vol. 2, No. 2, 2020
https://doi.org/10.2140/paa.2020.2.233

THE HYPERBOLIC YANG–MILLS EQUATION IN THE CALORIC GAUGE: LOCAL WELL-POSEDNESS AND CONTROL OF ENERGY-DISPERSED SOLUTIONS

SUNG-JIN OH AND DANIEL TATARU

This is the second part in a four-paper sequence, which establishes the threshold conjecture and the soliton bubbling vs. scattering dichotomy for the hyperbolic Yang–Mills equation in the (4+1)-dimensional space-time. This paper provides the key gauge-dependent analysis of the hyperbolic Yang–Mills equation.

We consider topologically trivial solutions in the caloric gauge, which was defined in the first paper of the sequence using the Yang–Mills heat flow. In this gauge, we establish a strong form of local well-posedness, where the time of existence is bounded from below by the energy concentration scale. Moreover, we show that regularity and dispersive properties of the solution persist as long as energy dispersion is small. We also observe that fixed-time regularity (but not dispersive) properties in the caloric gauge may be transferred to the temporal gauge without any loss, proving as a consequence small-data global well-posedness in the temporal gauge.

We use the results in this paper in subsequent papers to prove the sharp threshold theorem in caloric gauge in the trivial topological class, and the dichotomy theorem in arbitrary topological classes.

1. Introduction

In this paper, along with the companion papers [Oh and Tataru 2017a; 2017b; 2019a], we consider the hyperbolic Yang–Mills equation in the (4+1)-dimensional Minkowski space with a compact semisimple structure group.

MSC2010: 35L70, 70S15.
Keywords: hyperbolic Yang-Mills, energy critical, caloric gauge, small energy dispersion, regularity, scattering.

In [Oh and Tataru 2017b], we defined the notion of *caloric gauge* with the help of the Yang–Mills heat flow on $\mathbb{R}^4$, and showed that every subthreshold connection admits a caloric gauge representative (see Section 1B below for a review). The first main result of the present paper (Theorem 1.13) is a strong form of local well-posedness of the hyperbolic Yang–Mills equation in the manifold of caloric gauge connections, where the time of existence is estimated from below by the scale of energy concentration. The second main result (Theorem 1.16) asserts that regularity and dispersive behaviors persist as long as a certain quantity called *energy dispersion*, which measures a certain type of nondispersive concentration, remains small.

While the caloric gauge reveals the fine cancellation structure of the Yang–Mills equation, and is thus suitable for dispersive analysis at low regularity, it has the drawback that causality is lost. As a remedy, we also show that regularity (but not dispersive) properties in the caloric gauge may be transferred to the temporal gauge. As a corollary, we also obtain small-data global well-posedness of the hyperbolic Yang–Mills equation in the temporal gauge (Theorem 1.18).

In the subsequent papers in the sequence [Oh and Tataru 2017a; 2019a], we use the results proved in this paper to establish the threshold theorem (i.e., global well-posedness and scattering for subthreshold data) in the caloric gauge, as well as the soliton bubbling vs. scattering dichotomy theorem for general finite-energy solutions, formulated in a more gauge-covariant fashion. An overview of the entire series is provided in [Oh and Tataru 2019b].

1A. ***Hyperbolic Yang–Mills equation on*** $\mathbb{R}^{1+4}$**.** Our set-up is as follows. Let $\boldsymbol{G}$ be a compact noncommutative Lie group and $\mathfrak{g}$ its associated Lie algebra. We denote by $\mathrm{Ad}(O)X = OXO^{-1}$ the adjoint (or conjugation) action of $\boldsymbol{G}$ on $\mathfrak{g}$ and by $\mathrm{ad}(X)Y = [X, Y]$ the Lie bracket on $\mathfrak{g}$. We use the notation $\langle X, Y\rangle$ for a bi-invariant inner product on $\mathfrak{g}$,

$$\langle [X, Y], Z\rangle = \langle X, [Y, Z]\rangle, \quad X, Y, Z \in \mathfrak{g},$$

or equivalently

$$\langle X, Y\rangle = \langle \mathrm{Ad}(O)X, \mathrm{Ad}(O)Y\rangle, \quad X, Y \in \mathfrak{g},\ O \in \boldsymbol{G}.$$

If $\boldsymbol{G}$ is semisimple then one can take $\langle X, Y\rangle = -\mathrm{tr}(\mathrm{ad}(X)\,\mathrm{ad}(Y))$, i.e., the negative of the Killing form on $\mathfrak{g}$, which is then positive definite, However, a bi-invariant inner product on $\mathfrak{g}$ exists for any compact Lie group $\boldsymbol{G}$.

Let $\mathbb{R}^{1+4}$ be the (4+1)-dimensional Minkowski space equipped with the Minkowski metric, which takes the form $\mathrm{diag}(-1, +1, \dots, +1)$ in the rectangular coordinates $(x^0, x^1, \dots, x^4)$. The coordinate x^0 serves the role of time, and we will often write $x^0 = t$. Throughout this paper, we will use the standard convention for raising or lowering indices using the Minkowski metric, and summing up repeated upper and lower indices.

Our objects of study are connection 1-forms A on $\mathbb{R}^{1+4}$ taking values in the Lie algebra $\mathfrak{g}$. They define covariant differentiation operators $\boldsymbol{D}_\mu = \boldsymbol{D}^{(A)}_\mu = \partial_\mu + A_\mu$ (in coordinates) acting on sections of any vector bundle with structure group $\boldsymbol{G}$. The commutator $\boldsymbol{D}_\mu \boldsymbol{D}_\nu - \boldsymbol{D}_\nu \boldsymbol{D}_\mu$ yields the curvature 2-form $F_{\mu\nu} = F[A]_{\mu\nu}$, which is given in terms of A_μ by the formula

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu].$$

Given a $\boldsymbol{G}$-valued function O on $\mathbb{R}^{1+4}$, we introduce the notation

$$O_{;\mu} = \partial_\mu O O^{-1}.$$

The pointwise action of O on the vector bundle induces a gauge transformation for A and F, namely

$$A_\mu \mapsto O A_\mu O^{-1} - \partial_\mu O O^{-1} = \mathrm{Ad}(O) A_\mu - O_{;\mu}, \quad F_{\mu\nu} \mapsto O F_{\mu\nu} O^{-1} = \mathrm{Ad}(O) F_{\mu\nu}.$$

In view of this transformation property, F may be viewed as a 2-form taking values in the $\boldsymbol{G}$-vector bundle with fiber $\mathfrak{g}$, where $\boldsymbol{G}$ acts on $\mathfrak{g}$ by the adjoint action (geometrically, the adjoint vector bundle). Thus the covariant derivative $\boldsymbol{D}_\mu$ acts on F by

$$\boldsymbol{D}_\mu F_{\alpha\beta} = (\partial_\mu + \mathrm{ad}(A_\mu)) F_{\alpha\beta} = \partial_\mu F_{\alpha\beta} + [A_\mu, F_{\alpha\beta}].$$

The *hyperbolic Yang–Mills equation* on $\mathbb{R}^{1+4}$ is the Euler–Lagrange equation associated with the formal Lagrangian action functional

$$\mathcal{L}(A) = \frac{1}{2} \int_{\mathbb{R}^{1+4}} \langle F_{\alpha\beta}, F^{\alpha\beta} \rangle \, dx \, dt,$$

which takes the form

$$\boldsymbol{D}^\alpha F_{\alpha\beta} = 0. \tag{1-1}$$

Clearly, (1-1) is invariant under gauge transformations. This equation possesses a conserved energy, given by

$$\mathcal{E}_{\{t\}\times\mathbb{R}^4}(A) = \int_{\{t\}\times\mathbb{R}^4} \sum_{\alpha<\beta} |F_{\alpha\beta}|^2 \, dx. \tag{1-2}$$

Furthermore, both the equation (1-1) and the energy (1-2) are invariant under the scaling

$$A(t, x) \mapsto \lambda A(\lambda t, \lambda x) \quad (\lambda > 0).$$

Hence, the hyperbolic Yang–Mills equation is energy critical in dimension $(4+1)$, which is the reason why we focus on this dimension in the present series of papers.

We are interested in the initial value problem for (1-1). For this purpose, we first formulate a gauge-covariant notion of an initial data set. We say that a pair (a, e) of a connection 1-form a and a $\mathfrak{g}$-valued 1-form e on $\mathbb{R}^4$ is an initial data set for a solution A to (1-1) if

$$(A_j, F_{0j}) \restriction_{\{t=0\}} = (a_j, e_j).$$

Here and throughout this paper, roman letter indices stand for the spatial coordinates $x^1, \dots, x^4$. Note that (1-1) with $\beta = 0$ imposes the condition that

$$\boldsymbol{D}^j e_j = \partial^j e_j + [a^j, e_j] = 0. \tag{1-3}$$

This equation is the *Gauss* (or the *constraint*) *equation* for (1-1).

It turns out that (1-3) characterizes precisely those pairs (a, e) which can arise as an initial data set. Thus we make the following definition:

Definition 1.1. (1) A *regular initial data set* for the hyperbolic Yang–Mills equation is a pair $(a,e) \in H^N_{\mathrm{loc}} \times H^{N-1}$ $(N \geq 2)$ which has finite energy (i.e., $F[a] \in L^2$) and satisfies the constraint equation (1-3).
(2) A *finite-energy initial data set* is a pair $(a,e) \in \dot{H}^1_{\mathrm{loc}} \times L^2$ which has finite energy (i.e., $F[a] \in L^2$) and satisfies the constraint equation (1-3).

In this paper, we make an additional assumption that a decays suitably at infinity:

$$a \in \dot{H}^1. \tag{1-4}$$

This assumption turns out to be equivalent to the requirement that a is topologically trivial [Oh and Tataru 2019a]. As this property is conserved under any continuous evolution in time, this is the natural setting for scattering and thus for the threshold conjecture for (1-1), which is one main subject of the final paper [Oh and Tataru 2017a] of the series.

The hyperbolic Yang–Mills equation (1-1), when naively viewed as an evolution equation for A, fails to be locally well-posed; to restore (at least formally) well-posedness, we need to fix the gauge invariance.

There are several classical interesting gauge choices which can be made here, for instance the Coulomb gauge $\partial^j A_j = 0$, the temporal gauge $A_0 = 0$ and the Lorenz gauge $\partial^\alpha A_\alpha = 0$. For a more detailed discussion and comparison of these gauges we refer the reader to our first article [Oh and Tataru 2017b].

However, the main gauge choice we use in this paper is the so-called *caloric gauge*, which was defined in the first paper of the series [Oh and Tataru 2017b] with the help of a parabolic analogue of (1-1), namely the *Yang–Mills heat flow*. This is the subject of our next discussion.

1B. ***Yang–Mills heat flow and the caloric gauge.*** Let a be a connection 1-form on $\mathbb{R}^4$ (in short, a spatial connection). We say that a connection $A = A(x,s)$ on $\mathbb{R}^4 \times J$ (where J is a subinterval of $[0,\infty)$) is a (covariant) *Yang–Mills heat flow* development of a if it solves

$$F_{sj} = \boldsymbol{D}^\ell F_{\ell j}, \quad A(s=0) = a. \tag{1-5}$$

This equation is invariant under gauge transformations on $\mathbb{R}^4 \times J$. Under the *local caloric gauge* condition

$$A_s = 0, \tag{1-6}$$

the forward-in-s initial value problem for (1-5) is locally well-posed [Oh and Tataru 2017b, Theorem 2.7] in $\dot{H}^1$. We remark that the evolution (1-5) under the gauge (1-6) is precisely the gradient flow for the (spatial) energy

$$\mathcal{E}_e(a) = \frac{1}{2}\int_{\mathbb{R}^4} \langle F_{jk}[a], F^{jk}[a]\rangle \, dx = \int_{\mathbb{R}^4} \sum_{j<k} |F_{jk}[a]|^2 \, dx.$$

The key controlling norm for the Yang–Mills heat flow in the local caloric gauge is $\|F\|_{L^3_s(J;L^3)}$, which is both scale- and gauge-invariant.

Theorem 1.2 [Oh and Tataru 2017b]. *Consider a Yang–Mills heat flow $A \in C_s(J;\dot{H}^1)$ in the local caloric gauge satisfying*

$$\|F\|_{L^3_s(J;L^3)} \leq \mathcal{Q} < \infty. \tag{1-7}$$

When $J = [0, s_0)$ *for* $s_0 < \infty$, *A can be extended past* s_0 *as a (well-posed) Yang–Mills heat flow. When* $J = [0, \infty)$, *the solution has the property that the limit*

$$\lim_{s\to\infty} A(s) = a_\infty$$

exists in $\dot{H}^1$. *The limiting connection is flat* ($F[a_\infty] = 0$) *and the map* $a \mapsto a_\infty$ *is locally Lipschitz in* $\dot{H}^1$, H^N ($N \geq 1$) *and* $\dot{H}^1 \cap \dot{H}^N$ ($N \geq 2$). *Denoting by* $O(a)$ *a gauge transformation satisfying* $O^{-1}\partial_j O = a_\infty$, *the map* $a \mapsto O(a)$ *is continuous from* $\dot{H}^1$ *to* $\dot{H}^2$ *up to constant conjugations.*

In the case when the Yang–Mills heat flow with initial data a admits a global solution with finite L^3 norm for the curvature as in (1-7), we define the *caloric size* $\mathcal{Q}(a)$ of a as

$$\mathcal{Q}(a) = \|F\|^3_{L^3_s(\mathbb{R}^+;L^3)}. \tag{1-8}$$

We note that this is a gauge-invariant quantity.

Remark 1.3. Here we need to clarify the topology on the (nonlinear) space of gauge transformations. We will say that a sequence $O^{(n)}$ converges to O if there exists a sequence $\widetilde{O}^{(n)}$ of gauge transformations so that $\widetilde{O}^{(n)}(O^{(n)})^{-1}$ are constant and so that we have

- pointwise convergence,[1]
$$d(\widetilde{O}^{(n)}, O) \to 0 \quad \text{in } L^2_{\mathrm{loc}},$$
- convergence of derivatives,
$$\widetilde{O}^{(n)}_{;x} \to O_{;x} \quad \text{in } \dot{H}^1.$$

A simple but important case in which (1-7) holds with $J = [0, \infty)$ is when the initial energy $\mathcal{E}_e(a)$ is sufficiently small. The same conclusion holds as long as $\mathcal{E}_e(a)$ is below any nontrivial connection $a \in \dot{H}^1$ satisfying the harmonic Yang–Mills equation

$$\boldsymbol{D}^\ell F_{\ell j} = 0. \tag{1-9}$$

The above assertion is closely related to the topological class of connections. Relaxing the requirement $a \in \dot{H}^1$ to $a \in H^1_{\mathrm{loc}}$ allows also topologically nontrivial initial data sets, in which case the ground state energy

$$E_{\mathrm{GS}} = \inf\{\mathcal{E}_e(a) : a \in H^1_{\mathrm{loc}} \text{ is nontrivial and solves (1-9)}\} \tag{1-10}$$

is nonzero, and the minimum is attained for a special class of solutions called instantons. However, within the trivial topological class we have

$$2E_{\mathrm{GS}} \leq \inf\{\mathcal{E}_e(a) : a \in \dot{H}^1 \text{ is nontrivial and solves (1-9)}\}. \tag{1-11}$$

We further remark that in order for a connection a to have $\mathcal{Q}(a)$ finite, it must be topologically trivial. Because of this, the present paper is limited to topologically trivial connections, which are simply defined

[1] The functions $O^{(n)}$ are uniformly bounded in BMO so this property essentially provides the additional information that in some sense the local averages converge as well.

by the requirement that $a \in \dot{H}^1$ in a suitable gauge. For an extended discussion and further references we refer the reader to our next article in the series [Oh and Tataru 2019a].

In view of this discussion, the following result is natural:

Theorem 1.4 (threshold theorem for the Yang–Mills heat flow on $\mathbb{R}^4$ [Oh and Tataru 2017b]). *Assume that a is topologically trivial and that*

$$\mathcal{E}_e(a) < 2E_{\mathrm{GS}}.$$

Then the solution to (1-5) exists globally on $[0, \infty)$. *Moreover, there exists a nondecreasing function* $\mathcal{Q}(\cdot) : [0, 2E_{\mathrm{GS}}) \to [0, \infty)$ *such that*

$$\mathcal{Q}(a) \leq \mathcal{Q}(\mathcal{E}_e(a)).$$

We now return to the discussion of an arbitrary (not necessarily subthreshold) spatial connection a, whose Yang–Mills heat flow development satisfies (1-7) with $J = [0, \infty)$. Since the limiting connection a_∞ is flat, it must be gauge equivalent to the zero connection. This motivates the following definition of the *caloric gauge*:

Definition 1.5 (caloric gauge). We say that a connection $a_j \in \dot{H}^1$ is *caloric* if $J = [0, \infty)$ and a_∞ in Theorem 1.2 is equal to zero. We denote the set of all such connections by $\mathcal{C}$. More quantitatively, we denote by $\mathcal{C}_Q$ the set of all caloric connections whose Yang–Mills heat flow development satisfies

$$\mathcal{Q}(a) \leq Q. \tag{1-12}$$

Given a connection $a \in \dot{H}^1$ satisfying (1-7) with $J = [0, \infty)$, note that

$$\mathrm{Cal}(a)_j = \mathrm{Ad}(O(a))a_j - O(a)_{;j}$$

is its caloric representative, which is unique up to constant conjugations.

To solve the Yang–Mills equation in the caloric gauge, we need to view the family $\mathcal{C}$ of the caloric gauge connections as an infinite-dimensional manifold. Here the $\dot{H}^1$ topology is no longer sufficient, so we introduce the slightly stronger topology

$$\boldsymbol{H} = \{a \in \dot{H}^1 : \|a\|_{\boldsymbol{H}} < \infty\}, \quad \text{where } \|a\|_{\boldsymbol{H}} := \|a\|_{\dot{H}^1} + \sum_j \|P_j(\partial^\ell a_\ell)\|_{L^2}.$$

Here, $\{P_j\}$ refer to the standard Littlewood–Paley projections to dyadic frequency annuli on $\mathbb{R}^4$. It turns out that every caloric connection belongs to $\boldsymbol{H}$, which reflects the fact, to be discussed in Section 3 in greater detail, that caloric connections satisfy a nonlinear form of the Coulomb gauge condition. Moreover, the following theorem holds.

Theorem 1.6. (1) *For a connection $a \in \mathcal{C}$ with energy $\mathcal{E}$ and caloric size $\mathcal{Q}$ we have*

$$\|a\|_{\boldsymbol{H}} \lesssim_{\mathcal{E},\mathcal{Q}} 1.$$

(2) *Consider a connection $a \in \boldsymbol{H}$ (not necessarily caloric) satisfying (1-12). Then $O(a)$ in Theorem 1.2 may be uniquely fixed by imposing $\lim_{|x|\to\infty} O(a) = I$. Such a map $a \mapsto O(a)$ is locally C^1 from $\boldsymbol{H}$ to $\dot{H}^2 \cap C^0$, and also from H^N to $\dot{H}^2 \cap \dot{H}^{N+1}$* $(N \geq 2)$.

Essentially as a corollary, we have:

Theorem 1.7. *The set $\mathcal{C}$ is an infinite-dimensional C^1 submanifold of $\boldsymbol{H}$.*

The spatial components of finite-energy Yang–Mills waves will be continuous functions of time which take values into $\mathcal{C}$. They are however not C^1 in time; instead their time derivative will merely belong to L^2. Because of this, we need to take the closure of its tangent space $T\mathcal{C}$ (which a priori is a closed subspace of $\boldsymbol{H}$) in L^2. This is denoted by $T_a^{L^2}\mathcal{C}$. It is also convenient to have a direct way of characterizing this space; that is naturally done via the linearization of (1-5):

Definition 1.8. For a caloric gauge connection $a \in \mathcal{C}$, we say that $L^2 \ni b \in T_a^{L^2}\mathcal{C}$ if and only if the solution to the linearized local caloric gauge Yang–Mills heat flow equation

$$\partial_s B_k = [B^j, F_{kj}] + \boldsymbol{D}^j(\boldsymbol{D}_k B_j - \boldsymbol{D}_j B_k), \quad B_k(s=0) = b_k, \tag{1-13}$$

(where $\boldsymbol{D} = \boldsymbol{D}^{(a)}$) satisfies

$$\lim_{s\to\infty} B(s) = 0.$$

We say that $(a,b) \in T^{L^2}\mathcal{C}_{\mathcal{Q}}$ if $a \in \mathcal{C}_{\mathcal{Q}}$ and $b \in T_a^{L^2}\mathcal{C}$, and we say that $(a,b) \in T^{L^2}\mathcal{C}$ if $a \in \mathcal{C}$ and $b \in T_a^{L^2}\mathcal{C}$.

A key property of the tangent space $T_a^{L^2}\mathcal{C}$ is the following nonlinear div-curl-type decomposition:

Theorem 1.9. *Let $a \in \mathcal{C}_{\mathcal{Q}}$ with energy $\mathcal{E}$. Then for each $e \in L^2$ there exists a unique decomposition*

$$e = b - \boldsymbol{D}^{(a)} a_0, \quad b \in T_a^{L^2}\mathcal{C},\ a_0 \in \dot{H}^1, \tag{1-14}$$

with the corresponding bound

$$\|b\|_{L^2} + \|a_0\|_{\dot{H}^1} \lesssim_{\mathcal{E},\mathcal{Q}} \|e\|_{L^2}. \tag{1-15}$$

A hyperbolic Yang–Mill connection consists not only of spatial components (the sole subject of discussion so far), but also of a temporal component. As in the Coulomb gauge, we will consider the spatial components of the connection as the dynamic variables, which satisfy a system of wave equations. The temporal components, on the other hand, will be viewed as an auxiliary variable determined from the spatial components. This point of view motivates the following definition.

Definition 1.10 (initial data in the caloric gauge). An initial data for the Yang–Mills equation in the caloric gauge is a pair (a,b) where $(a,b) \in T^{L^2}\mathcal{C}$.

The notion of covariant Yang–Mills initial data (Definition 1.1) is connected to the preceding definition by the following result proved in [Oh and Tataru 2017b] (which motivates the notation in Theorem 1.9):

Theorem 1.11. (1) *Given any Yang–Mills initial data pair $(a,e) \in \dot{H}^1 \times L^2$ such that the Yang–Mills heat flow development of a satisfies (1-12), there exists a caloric gauge Yang–Mills data $(\tilde{a}, b) \in T^{L^2}\mathcal{C}$ and $a_0 \in \dot{H}^1$, so that the initial data pair $(\tilde{a}, \tilde{e})$ is gauge equivalent to (a,e), where*

$$\tilde{e}_k = b_k - \boldsymbol{D}_k^{(\tilde{a})} a_0.$$

In addition, $(\tilde{a}, b)$ *and* a_0 *are unique up to constant conjugations*, *and depend continuously on* (a, e) *in the corresponding quotient topology. Further*, *the map* $(a, e) \mapsto (\tilde{a}, b)$ *is locally* C^1 *in the stronger topology*[2] $\boldsymbol{H} \times L^2 \to \boldsymbol{H} \times L^2$, *as well as in more regular spaces* $H^N \times H^{N-1} \to H^N \times H^{N-1}$ $(N \geq 2)$.

(2) *Given any caloric gauge data* $(a, b) \in T^{L^2}\mathcal{C}$, *there exists a unique* $a_0 \in \dot{H}^1$, *with Lipschitz dependence on* $(a, b) \in \dot{H}^1 \times L^2$, *so that*

$$e_k = b_k - \boldsymbol{D}_k^{(a)} a_0$$

satisfies the constraint equation (1-3). *Further*, *the map* $(a, b) \to a_0$ *is also Lipschitz from* $H^N \times H^{N-1}$ *to* H^N *for* $N \geq 3$.

Remark 1.12. The caloric gauge just described is a global version of a local caloric gauge previously introduced by Oh [2014; 2015], and is based on an idea by Tao [2004] in his study of the energy-critical wave maps into the hyperbolic space [Tao 2008a; 2008b; 2008c; 2009a; 2009b].

1C. *The main results.* The first main result is a strong gauge-dependent local well-posedness theorem for the Yang–Mills equation as an evolution in the manifold of caloric connections. To state this result, we define the *energy concentration scale* r_c of a Yang–Mills initial data set (a, e) with threshold ε_* (or the ε_*-energy concentration scale) to be

$$r_c^{\varepsilon_*} = r_c^{\varepsilon_*}[a, e] = \sup\{r : \mathcal{E}_{B_r}(a, e) \leq \varepsilon_*^2\}.$$

Theorem 1.13 (local well-posedness in caloric gauge). *There exists a nonincreasing function* $\varepsilon_*(\mathcal{E}, \mathcal{Q}) > 0$ *and a nondecreasing function* $M_*(\mathcal{E}, \mathcal{Q}) > 0$ *such that the Yang–Mills equation in the caloric gauge is locally well-posed on the time interval of length* $r_c = r_c^{\varepsilon_*}(\mathcal{E}, \mathcal{Q})$ *for initial data* (a, e) *with energy* $\leq \mathcal{E}$ *and* $a \in \mathcal{C}_{\mathcal{Q}}$. *More precisely*, *the following statements hold*:

(1) (*regular data*) *Let* (a, e) *be a smooth initial data set with energy* $\leq \mathcal{E}$, *where* $a \in \mathcal{C}_{\mathcal{Q}}$. *Then there exists a unique smooth solution* $A_{t,x}$ *to the Yang–Mills equation in caloric gauge on* $I = [-r_c, r_c]$ *such that* $(A_j, F_{0j}) \restriction_{\{t=0\}} = (a_j, e_j)$.

(2) (*rough data*) *The data-to-solution map admits a continuous extension*

$$\mathcal{C} \times L^2 \ni (a, e) \mapsto (A_x, \partial_t A_x) \in C(I, T^{L^2}\mathcal{C})$$

in the class of initial data with energy $\leq \mathcal{E}$, $a \in \mathcal{C}_{\mathcal{Q}}$ *and energy concentration scale* $\geq r_c$.

(3) (*a priori bound*) *The solution defined as above obeys the a priori bound*

$$\|A_x\|_{S^1[I]} \leq M_*(\mathcal{E}, \mathcal{Q}).$$

(4) (*weak Lipschitz dependence*) *Let* $(a', e') \in \mathcal{C} \times L^2$ *be another initial data set with energy concentration scale* $\geq r_c$. *For* $\sigma < 1$ *close to* 1, *we have the global bound*

$$\|A_x - A'_x\|_{S^\sigma[I]} \lesssim_{M_*(\mathcal{E}, \mathcal{Q}), \sigma} \|(a, e) - (a', e')\|_{\dot{H}^\sigma \times \dot{H}^{\sigma-1}}.$$

[2] Here we impose again the condition $\lim_{|x| \to \infty} O(a) = I$ in order to fix the choice of $O(a)$.

The a priori bound (3) is highly gauge-dependent and has strong consequences. The S^1-norm, which is essentially the same as in [Krieger et al. 2015] and is recalled in Section 4A below, serves the role of a controlling (or scattering) norm for the Yang–Mills equation in the caloric gauge. As we will see in Section 5, finiteness of the S^1-norm implies fine properties of the solution itself, such as frequency envelope control, persistence of regularity, continuation and scattering towards endpoints of I, and also for those nearby, such as weak Lipschitz dependence and local-in-time continuous dependence.

Theorem 1.13 implies small energy global well-posedness in the caloric gauge, analogous to the similar Coulomb gauge result in [Krieger and Tataru 2017]:

Corollary 1.14. *If the energy of the initial data set is smaller than* $\varepsilon_*^2 := \min\{1, \varepsilon_*^2(1, \mathcal{Q}(1))\}$, *then the corresponding solution* $A_{t,x}$ *in the caloric gauge exists globally and obeys*

$$\|A_x\|_{S^1[(-\infty,\infty)]} \le M_*(\mathcal{E}).$$

Moreover, if the initial data set (a, e) has subthreshold energy, then by Theorem 1.4 we have $a \in \mathcal{C}_{\mathcal{Q}}$ with $\mathcal{Q} \le \mathcal{Q}(\mathcal{E})$. Therefore, we immediately obtain:

Corollary 1.15. *For initial data with subthreshold energy, the conclusions of Theorem 1.13 hold with* ε_*, M_* *and* r_c *depending only on the energy* $\mathcal{E}$.

The local well-posedness result (Theorem 1.13) provides a basic framework for considering dynamics of the Yang–Mills equation in the manifold of caloric connections $\mathcal{C}$. The second main result, which we now state, is a continuation/scattering criterion for this equation in terms of smallness of a quantity called *energy dispersion* (denoted by $\mathrm{ED}[I]$ below).

Theorem 1.16 (regularity and scattering of energy-dispersed YM solutions). *There exists a nonincreasing function* $\varepsilon(\mathcal{E}, \mathcal{Q}) > 0$ *and a nondecreasing function* $M(\mathcal{E}, \mathcal{Q})$ *such that if* $A_{t,x}$ *is a solution* (*in the sense of Theorem 1.13*) *to the Yang–Mills equation in caloric gauge on* I *with energy* $\le \mathcal{E}$ *and with initial caloric size* $\mathcal{Q}$ *that obeys*

$$\|F\|_{\mathrm{ED}[I]} = \sup_{k \in \mathbb{Z}} 2^{-2k} \|P_k F\|_{L^\infty(I \times \mathbb{R}^4)} \le \varepsilon(\mathcal{E}, \mathcal{Q}),$$

then it satisfies the a priori bound

$$\|A_x\|_{S^1[I]} \le M(\mathcal{E}, \mathcal{Q}),$$

as well as

$$\sup_{t \in I} \mathcal{Q}(A(0)) \ll 1.$$

By finiteness of the S^1-norm, $A_{t,x}$ may be continued as a solution to the Yang–Mills equation in the caloric gauge past finite endpoints of I, and scatters in some sense towards the infinite endpoints; see Remarks 5.2 and 5.3.

Remark 1.17. In contrast to Theorem 1.13, in Theorem 1.16 the dependence on $\mathcal{Q}$ is very mild. This feature is due to the fact that small energy dispersion, combined with the energy bound, implies that $\mathcal{Q}$ must be either very large or very small; see Lemma 5.10 below. In particular if $\mathcal{E}$ is subthreshold then the dependence on $\mathcal{Q}$ above can be omitted altogether.

While powerful conclusions about the solution (represented by the S^1-norm bound) can be made in the caloric gauge, it has the disadvantage that the causality (or the finite speed of propagation) property is lost. To remedy this, we also establish small-data well-posedness result in the temporal gauge $A_0 = 0$:

Theorem 1.18. *If the energy of the initial data set is smaller than ε_*^2 (as in Corollary 1.14), then the corresponding solution $(A_{t,x}, \partial_t A_{t,x})$ in the temporal gauge $A_0 = 0$ exists globally in $C_t(\mathbb{R}; \dot{H}^1 \times L^2)$. The solution is unique among the local-in-time limits of smooth solutions, and it depends continuously on data $(a, e) \in \dot{H}^1 \times L^2$.*

In fact, Theorem 1.18 is a consequence of Corollary 1.14, after the observation that the gauge transformation from the caloric gauge to the temporal gauge obeys optimal regularity bounds; see Theorem 5.1 (10) below. We note that the strong dispersive S^1-norm bound for A is generally lost in the temporal gauge, as some part of the solution is merely transported (instead of solving a wave equation).

Theorem 1.18 is used in the third paper [Oh and Tataru 2019a] of the sequence to establish the large-data local theory for the (4+1)-dimensional Yang–Mills equation in arbitrary topological classes. Then in the fourth paper [Oh and Tataru 2017a], this theory is put together with Theorems 1.13 and 1.16 to establish global well-posedness and scattering in the caloric gauge for data with subthreshold energy (often called the *threshold theorem* in the literature), as well as a bubbling vs. scattering dichotomy for arbitrary finite-energy solutions, formulated in a gauge-covariant sense.

Remark 1.19. Within the setup of this paper, one could in effect easily relax the hypothesis of the above theorem, and show that temporal gauge solutions exist for as long as caloric solutions exist. We do not pursue this, as our primary interest in terms of the temporal gauge is to use it for solutions which are not necessarily caloric. These matters are further discussed in our third and fourth papers [Oh and Tataru 2017a; 2019a].

The overall strategy for the proofs originated from the work of Sterbenz and the second author on the energy-critical wave maps [Sterbenz and Tataru 2010a; 2010b] and was adapted to the case of the energy-critical Maxwell–Klein–Gordon (MKG) equation, which is a simpler model for Yang–Mills, in our previous works [Oh and Tataru 2016a; 2016b; 2018]. We also note an alternative independent approach for the energy-critical wave maps [Krieger and Schlag 2012] and MKG [Krieger and Lührmann 2015] based on the Kenig–Merle method [2008; 2006]. A more extensive historical perspective is provided in the fourth paper [Oh and Tataru 2017a].

In [Oh and Tataru 2016b; 2018], the analogues of Theorems 1.13 and 1.16 (respectively) were proved using distinct strategies. However, here we derive both main results (see Section 7 for details) from the following single a priori estimate concerning regular solutions, whose proof is the central goal of this paper:

Theorem 1.20. *There exist nonincreasing functions $\varepsilon(\mathcal{E}, \mathcal{Q}), T(\mathcal{E}, \mathcal{Q}) > 0$ as well as a nondecreasing function $M(\mathcal{E}, \mathcal{Q})$ such that if $A_{t,x}$ is a regular solution to the Yang–Mills equation in caloric gauge on I with energy $\leq \mathcal{E}$ such that $A_x \in \mathcal{C}_{\mathcal{Q}}$ for all $t \in I$, and moreover*

$$\sup_{k \geq m} 2^{-2k} \| P_k F \|_{L^\infty(I \times \mathbb{R}^4)} \leq \varepsilon(\mathcal{E}, \mathcal{Q}) \quad \textit{and} \quad |I| \leq 2^{-m} T(\mathcal{E}, \mathcal{Q})$$

for some $m \in \mathbb{Z}$, then it satisfies the a priori bound

$$\|A_x\|_{S^1[I]} \le M(\mathcal{E}, \mathcal{Q}).$$

In words, for a regular solution with small energy dispersion only at certain frequency 2^m and above, an a priori S^1-norm bound holds on time intervals of the corresponding scale $O(2^{-m})$.

1D. *Overview of the paper.* Section 2: In this section, we collect some notation and conventions used throughout this paper for the reader's convenience. Some basic concepts, such as disposability, dyadic function spaces, frequency envelopes, etc., are also described.

After Section 2, the paper is organized into two tiers. The first tier consists of Sections 3 to 7, and its goal is to describe the large-scale proof of the main results, assuming the validity of certain linear and multilinear estimates collected in Section 4.

Section 3: Here, we recall from [Oh and Tataru 2017b] further results concerning the Yang–Mills heat flow and the caloric gauge. First, we state some quantitative bounds for the Yang–Mills heat flow and its linearization in the caloric gauge, using the language of frequency envelopes (Section 3A). Next, we derive the wave equation satisfied by A_x and $A_x(s)$ $(s > 0)$ in the caloric gauge (Section 3B). In this process we use the *dynamic Yang–Mills heat flow* (3-5), which is the Yang–Mills heat flow augmented with a heat evolution (in s) for the temporal component.

Section 4: We first describe the fine function space framework for analyzing the hyperbolic Yang–Mills equation in the caloric gauge (Section 4A). The main function spaces are identical to those in [Krieger et al. 2015; Oh and Tataru 2018; Krieger and Tataru 2017], which in turn have their roots in the works on wave maps [Tataru 2001; Tao 2001]. We also explain the three main sources of smallness in our analysis: divisibility, small energy dispersion and short time interval. Then we state the linear and multilinear estimates needed for the proof of the main theorems (Sections 4B and 4C); it is the goal of the second tier of the paper (described below) to prove them. The primary estimates here are the bilinear null form estimates, which in the context of our function spaces have their origin in [Krieger et al. 2015; Oh and Tataru 2018; Krieger and Tataru 2017]. The bilinear null structure of the Yang–Mills nonlinearities was first described in [Klainerman and Machedon 1994]; a secondary trilinear null structure, which also plays a role here, was discovered in [Machedon and Sterbenz 2004] in the (MKG) context.

Section 5: We prove a strong *structure theorem* for a solution to the hyperbolic Yang–Mills equation in the caloric gauge with finite S^1-norm (Section 5A). In particular, it reduces the tedious task of controlling various parts of a solution $A_{t,x}$ to proving a single S^1-norm bound for the spatial components A_x. We also consider the effect of small inhomogeneous energy dispersion on a correspondingly short time interval (Section 5B). The analysis is repeated for the dynamic Yang–Mills heat flow of a solution (Section 5C).

Section 6: We prove the central result, Theorem 1.20, by an induction-on-energy argument. The argument is similar to [Oh and Tataru 2018], which in turn was based on the work [Sterbenz and Tataru 2010a], with modifications to handle the low frequencies with possibly large energy dispersion with the short length of the time interval (see, in particular, scenario (1) in Section 6B).

Section 7: Here, we derive the main theorems stated in Section 1C from Theorem 1.20. The key point in the derivation of Theorem 1.13 is the simple fact that energy dispersion is small for frequencies above the inverse of the energy-concentration scale (Section 7B). Theorem 1.16 follows essentially by scaling (Section 7C).

The second tier consists of Sections 8 to 11. Here, we provide proofs of the estimates stated in Section 4.

Section 8: The goal of this section is to prove all multilinear estimates stated in Section 4. The proofs proceed in two stages: In the first stage, we assume global-in-time dyadic (in spatial frequency) estimates (Section 8B), and derive the interval-localized frequency envelope bounds stated in Section 4 (Section 8C). A key technical issue in interval localization is to deal with modulation projections, which are nonlocal in time. In the second stage, we establish the global-in-time dyadic estimates (Section 8D). Much is borrowed from the previous works [Krieger et al. 2015; Oh and Tataru 2018; Krieger and Tataru 2017].

Section 9: We begin this section by reducing the proof of the key linear estimates in Section 4 to construction of a parametrix for the paradifferential d'Alembertian $\Box + 2\sum_k \mathrm{ad}(P_{<k-\kappa} \boldsymbol{P}_\alpha A)\partial^\alpha P_k$ (Section 9A). As in [Krieger and Tataru 2017], the parametrix is constructed via conjugation of the free-wave propagator by a pseudodifferential renormalization operator. We define and state the key properties of the renormalization operator (Section 9C), and establish the desired estimates for the parametrix assuming these properties (Section 9D).

Section 10: Here, we prove the mapping properties of the renormalization operator claimed in Section 9. The key difference from [Krieger and Tataru 2017] lies in the source of smallness: whereas smallness of the S^1-norm of A was used in that paper, in this paper we rely instead on largeness of the frequency gap κ in the paradifferential d'Alembertian. The idea of exploiting a large frequency gap was used in [Sterbenz and Tataru 2010a; Oh and Tataru 2018].

Section 11: Finally, we estimate the error for conjugation of the paradifferential d'Alembertian by the renormalization operator claimed in Section 9, thereby completing our parametrix construction. One aspect of our proof that differs from the previous works [Sterbenz and Tataru 2010a; Oh and Tataru 2018] is that, in addition to the large frequency gap κ, we need to use smallness of a divisible norm (weaker than S^1) of A, which requires a careful interval localization procedure (Sections 11C and 11D).

2. Notation, conventions and other preliminaries

2A. *Notation and conventions.* Here we collect some notation and conventions used in this paper.

- The symbols $\lesssim$, $\gtrsim$, $\ll$ and $\gg$ are defined with their usual meanings, where the implicit constants in these notations are allowed to vary from line to line.

- By $A \lesssim_E B$ and $A \ll_E B$, we mean that $A \le C_E B$ and $A \le c_E B$, respectively, where $C_E = C_0(1+E)^{C_1}$ and $c_E = C_0^{-1}(1+E)^{-C_1}$ for some constants $C_0, C_1 > 0$ that are again allowed to vary from line to line.

- For $u \in \mathfrak{g}$ and $O \in \boldsymbol{G}$, define $\mathrm{ad}(u) = [u, \cdot\,]$ and $\mathrm{Ad}(O) = O(\cdot\,)O^{-1}$, both of which are in $\mathrm{End}(\mathfrak{g})$. Recall the minus Killing form, which is invariant under $\mathrm{Ad}(O)$ and $\mathrm{ad}(X)$. On $\mathfrak{g}$, define $|\cdot|_{\mathfrak{g}}$ on $\mathfrak{g}$ by the minus Killing form. On $\mathrm{End}(\mathfrak{g})$, use the induced metric $|a|_{\mathrm{End}(\mathfrak{g})} = \sup_{|u|_{\mathfrak{g}} \leq 1} |au|_{\mathfrak{g}}$. By Ad-invariance, $|\mathrm{Ad}(O)a|_{\mathrm{End}(\mathfrak{g})} = |a \,\mathrm{Ad}(O^{-1})|_{\mathrm{End}(\mathfrak{g})} = |a|_{\mathrm{End}(\mathfrak{g})}$.

- We use the notation $B_r(x)$ for the ball of radius r centered at x. We write $|\angle(\xi, \eta)|$ for the angular distance $|\xi/|\xi| - \eta/|\eta||$, and $|\angle(\mathcal{C}, \mathcal{C}')|$ for $\inf_{\xi \in \mathcal{C}, \, \eta \in \mathcal{C}'} |\angle(\xi, \eta)|$.

- We use the notation $\nabla = \partial_{t,x}$, $D_\mu = i^{-1}\partial_\mu$. Also, for D and A we often suppress the subscript x and write $D = D_x$ and $A = A_x$.

- We say that a multilinear operator $\mathcal{O}(u_1, \dots, u_m)$ is *disposable* if its kernel is translation-invariant and has mass $\lesssim 1$. In particular, we have

$$\|\mathcal{O}(u_1, \dots, u_m)\|_Y \lesssim \|u_1\|_{X_1} \cdots \|u_m\|_{X_m}$$

for any translation-invariant spaces $X_1, \dots, X_m, Y$ provided that a product estimate

$$\|u_1 \cdots u_m\|_Y \lesssim \|u_1\|_{X_1} \cdots \|u_m\|_{X_m}$$

holds for any functions $u_1 \in X_1, \dots, u_m \in X_m$.

- We often use the "duality" pairing

$$\iint u_0 \mathcal{O}(u_1, \dots, u_m) \, dx \, dt$$

so as to have symmetry among u_0 and the inputs. Indeed, we have

$$\iint u_0 \mathcal{O}(u_1, \dots, u_m) \, dx \, dt = \iint_{\Xi^0 + \Xi^1 + \dots + \Xi^m = 0} \mathcal{O}(\Xi^1, \dots, \Xi^m) \tilde{u}_0(\Xi^0) \tilde{u}_1(\Xi^1) \cdots \tilde{u}_m(\Xi^m) \, d\Xi \, dt.$$

- We define $\mathcal{O}^{*i}$ as

$$\iint u_0 \mathcal{O}^{*i}(u_1, \dots, u_i, \dots, u_m) \, dt \, dx = \iint u_i \mathcal{O}(u_1, \dots, \overbrace{u_0}^{i\text{-th entry}}, \dots, u_m) \, dt \, dx.$$

- By a *bilinear operator* (*of* $\mathfrak{g}$*-valued functions*) *with symbol* $m(\xi, \eta) = m^{\boldsymbol{ab}}(\xi, \eta)$ (which is a complex-valued 4×4-matrix), we mean an expression of the form

$$\mathfrak{L}(a, b) = \iint (m^{\boldsymbol{ab}}(\xi, \eta)[\hat{a}_{\boldsymbol{a}}(\xi), \hat{b}_{\boldsymbol{b}}(\eta)]) e^{i(\xi + \eta)\cdot x} \frac{d\xi \, d\eta}{(2\pi)^8}.$$

For a scalar-valued symbol $m(\xi, \eta)$, we implicitly associate the corresponding multiple of the identity $m^{\boldsymbol{ab}}(\xi, \eta) = m(\xi, \eta)\delta^{\boldsymbol{ab}}$.

If $\mathfrak{L}$ were symmetric, then the symbol $m(\xi, \eta)$ would be *antisymmetric in* ξ, η, in the sense that $m^{\boldsymbol{ab}}(\xi, \eta) = -m^{\boldsymbol{ba}}(\eta, \xi)$; this is due to the antisymmetry of the Lie bracket.

2B. ***Basic multipliers and function spaces.*** Here we provide the definitions of basic multipliers and function spaces. For the more elaborate frequency projections and function spaces for the hyperbolic Yang–Mills equation, see Section 4A.

• Given a function space X (on either $\mathbb{R}^d$ or $\mathbb{R}^{1+d}$), we define the space $\ell^p X$ by

$$\|u\|_{\ell^p X}^p = \sum_k \|P_k u\|_X^p$$

(with the usual modification for $p = \infty$), where P_k ($k \in \mathbb{Z}$) are the usual Littlewood–Paley projections to dyadic frequency annuli.

• For a spatial 1-form A, we define $\boldsymbol{P}A$ to be its *Leray projection*, i.e., the L^2-projection to divergence-free vector fields:

$$\boldsymbol{P}_j A = A_j + (-\Delta)^{-1}\partial_j \partial^\ell A_\ell.$$

We write $\boldsymbol{P}_j^\perp A = A_j - \boldsymbol{P}_j A$.

• For a space-time 1-form A_α, we introduce the notation $\boldsymbol{P}_\alpha A = (\boldsymbol{P}A)_\alpha$ by defining

$$\boldsymbol{P}_\alpha A = \begin{cases} \boldsymbol{P}_j A_x, & \alpha = j \in \{1, \dots, 4\}, \\ A_0, & \alpha = 0. \end{cases}$$

We also define $\boldsymbol{P}_\alpha^\perp A = (\boldsymbol{P}^\perp A)_\alpha = A_\alpha - \boldsymbol{P}_\alpha A$.

• We denote by $\dot{W}^{\sigma,p}$ the homogeneous L^p-Sobolev space with regularity σ. In the case $p = 2$, we simply write $\dot{H}^\sigma = \dot{W}^{\sigma,2}$.

• The mixed space-time norm $L_t^q \dot{W}_x^{\sigma,r}$ of functions on $\mathbb{R}^{1+d}$ is often abbreviated as $L^q \dot{W}^{\sigma,r}$.

2C. *Frequency envelopes.* To provide more accurate versions of many of our estimates and results we use the language of frequency envelopes.

Definition 2.1. Given a translation-invariant space of functions X, we say that a sequence c_k is a frequency envelope for a function $u \in X$ if

(i) the dyadic pieces of u satisfy

$$\|P_k u\|_X \leq c_k,$$

(ii) the sequence c_k is slowly varying,

$$2^{-\delta(j-k)} \lesssim \frac{c_k}{c_j} \lesssim 2^{\delta(j-k)}, \quad j > k.$$

Here δ is a small positive universal constant. For some of the results we need to relax the slowly varying property in a quantitative way. Fixing a universal small constant $0 < \varepsilon \ll 1$, we set:

Definition 2.2. Let $\sigma_1, \sigma_2 > 0$. A frequency envelope c_k is called $(-\sigma_1, \sigma_2)$-admissible if

$$2^{-\sigma_1(1-\varepsilon)(j-k)} \lesssim \frac{c_k}{c_j} \lesssim 2^{\sigma_2(1-\varepsilon)(j-k)}, \quad j > k.$$

When $\sigma_1 = \sigma_2$, we simply say that c_k is σ-admissible.

Another situation that will occur frequently is that where we have a reference frequency envelope c_k, and then a secondary envelope d_k describing properties which apply on a background controlled by c_k.

In this context the envelope d_k often cannot be chosen arbitrarily but instead must be in a constrained range depending on c_k. To address such matters we set:

Definition 2.3. We say that the envelope d_k is σ-*compatible* with c_k if we have

$$c_k \sum_{j<k} 2^{\sigma(1-\varepsilon)(j-k)} d_j \lesssim d_k.$$

We will often replace envelopes d_k which do not satisfy the above compatibility condition by slightly larger envelopes that do:

Lemma 2.4 [Oh and Tataru 2017b, Lemma 3.5]. *Assume that c_k and d_k are $(-\sigma_1, S)$ envelopes, and also that c_k is bounded. Then for $\tilde{\sigma} < \sigma(1-\varepsilon)$ the envelope*

$$e_k = d_k + c_k \sum_{j<k} 2^{\tilde{\sigma}(j-k)} d_j$$

is σ-compatible with c_k. The implicit constant in Definition 2.3 is bounded above by $1 + C_{\sigma(1-\varepsilon)-\tilde{\sigma}} \|c\|_{\ell^\infty}$.

Finally we need the following additional frequency envelope notation:

$$(c \cdot d)_k = c_k d_k, \qquad a_{\leq k} = \sum_{j \leq k} a_j, \qquad c_k^{[\sigma]} = \sup_{j<k} 2^{(1-\varepsilon)\sigma(j-k)} c_j \quad (\sigma > 0).$$

2D. *Global small constants.* In this paper, we use a string of global small constants $\delta_1, \dots, \delta_6, \delta_7$ with the hierarchy

$$0 < \delta_* = \delta_7 \ll \delta_6 \ll \delta_5 \ll \delta_4 \ll \delta_3 \ll \delta_2 \ll \delta_1 \ll \delta_0 \ll 1. \tag{2-1}$$

These are fixed from right to left, so that

$$\delta_{i+1} \ll \delta_i^{100}.$$

The role of each constant is roughly as follows:

- δ_0: for definition of functions spaces, such as Str^1 and b_0, b_1, p_0 in Section 4.
- δ_1: for all bounds from other papers, such as [Oh and Tataru 2017b; 2018; Krieger and Tataru 2017]; also for all dyadic gains in explicit nonlinearities (Section 8) and for energy dispersion gains in the Str^1 norm (4-21).
- δ_2: for energy dispersion, frequency gap and off-diagonal gains in Sections 4.
- δ_3: for frequency envelope admissibility range in Sections 4.
- δ_4: for energy dispersion and frequency gap gains in Sections 5.
- δ_5: for frequency envelope admissibility range in Sections 5.
- δ_6: for energy dispersion and frequency gap gains in Sections 6.
- δ_*: for frequency envelope admissibility range in Sections 6.

We use an additional set of small constants in our parametrix construction (Sections 9–11), which are fixed after δ_1 but before δ_2.

3. Yang–Mills heat flow and the caloric gauge

In this section, which is a continuation of Section 1B, we recall the results from the first paper [Oh and Tataru 2017b] that are needed in the present paper.

In Section 3A, we state quantitative bounds for the Yang–Mills heat flow (and its linearization) in the caloric gauge, using the language of frequency envelopes. Section 3B is concerned with the task of interpreting the hyperbolic Yang–Mills equation in the caloric gauge as a system of nonlinear wave equations for A_x.

3A. ***Frequency envelope bounds in the caloric gauge.*** We begin with frequency envelope bounds for the caloric gauge Yang–Mills heat flow and its linearization.

Proposition 3.1 [Oh and Tataru 2017b, Proposition 7.27]. *Let $(a,b) \in T^{L^2}\mathcal{C}_Q$ with $\mathcal{E} = \mathcal{E}_e(a)$, and let (A, B) be the solution to (1-5) and (1-13) with (a,b) as data. Let c_k be a $(-\delta_1, S)$-frequency envelope in $\dot{H}^1 \times L^2$ for (a,b), and let $c_k^{\sigma,p}$ be a $(-\delta_1, S)$-frequency envelope in $\dot{W}^{\sigma,p} \times \dot{W}^{\sigma-1,p}$ for (a,b) which is δ_1-compatible with c_k. Define*

$$\boldsymbol{A}(s) = A(s) - e^{s\Delta}a, \quad \boldsymbol{B}(s) = B(s) - e^{s\Delta}b. \tag{3-1}$$

Then the following properties hold:

(1) *We have*

$$\|P_k \boldsymbol{A}(s)\|_{\dot{H}^1} + \|P_k \boldsymbol{B}(s)\|_{L^2} \lesssim_{\mathcal{E},\mathcal{Q},N} \langle 2^{-2k}s^{-1}\rangle^{-\delta_1}\langle 2^{2k}s\rangle^{-N} c_k^2. \tag{3-2}$$

(2) *For (σ, p) and (σ_1, p_1) satisfying*

$$c_{\delta_1} \le \sigma \le \frac{4}{p} - c_{\delta_1}, \quad 2 + c_{\delta_1} \le p \le c_{\delta_1}^{-1}, \quad 0 \le \sigma_1 \le \sigma - c_{\delta_1}, \quad \frac{4}{p_1} - \sigma_1 = 2\Big(\frac{4}{p} - \sigma\Big), \tag{3-3}$$

we have

$$\|P_k \boldsymbol{A}(s)\|_{\dot{W}^{\sigma_1+1,p_1}} + \|P_k \boldsymbol{B}(s)\|_{\dot{W}^{\sigma_1,p_1}} \lesssim_{\mathcal{E},\mathcal{Q},N} \langle 2^{-2k}s^{-1}\rangle^{-\delta_1}\langle 2^{2k}s\rangle^{-N} (c_k^{\sigma,p})^2. \tag{3-4}$$

A central object of the remainder of this section is the *dynamic Yang–Mills heat flow* for space-time connections, which is an augmentation of (1-5) with an equation for the temporal component. More precisely, we say that a pair (A_0, A) of a $\mathfrak{g}$-valued function A_0 and a connection A on $\mathbb{R}^4 \times J$ (where J is a subinterval of $[0,\infty)$) is the dynamic Yang–Mills heat flow development of (a_0, a) if

$$F_{s\alpha} = \boldsymbol{D}^\ell F_{\ell\alpha}, \quad (A_0, A)(s=0) = (a_0, a). \tag{3-5}$$

This flow is well-defined as long as the spatial and s-components of A are well-defined as a solution to (1-5). In particular, if $a \in \mathcal{C}$, then (A_0, A) exists on $[0,\infty)$, $\lim_{s\to\infty} A_0 = 0$ in $\dot{H}^1$ and $\lim_{s\to\infty} F_{0j} = 0$ in L^2. Moreover, the following proposition holds.

Proposition 3.2 [Oh and Tataru 2017b, Propositions 7.7 and 8.9]. *Let $a \in \mathcal{C}_Q$ and $e \in L^2$ satisfy $\|(f,e)\|_{L^2}^2 \le \mathcal{E}$. Consider also $a_0 \in \dot{H}^1$ and $b \in T_a^{L^2}\mathcal{C}$ which obeys $e = b - \boldsymbol{D}a_0$ (see Theorem 1.9), and let (A_0, A) be a caloric gauge solution to (3-5) with data (a_0, a). Then the following properties hold.*

(1) *The spatial* 1*-form* $B_j(s) = F_{0j}(s) - \mathbf{D}_j A_0(s)$ *obeys the linearized Yang–Mills heat flow in the caloric gauge with* $B_j(0) = b_j$. *Moreover*,

$$\|A(s)\|_{\dot{H}^1} + \|B(s)\|_{L^2} \lesssim_{\mathcal{E},\mathcal{Q}} \|(f,e)\|_{L^2}. \tag{3-6}$$

(2) *Let* d_k *be a* δ_1*-frequency envelope for* (f,e) *in* $\dot{W}^{-2,\infty}$. *Then*

$$2^{-k}\|P_k A(s)\|_{L^\infty} + 2^{-2k}\|P_k B(s)\|_{L^\infty} \lesssim_{\mathcal{E},\mathcal{Q},N} \langle 2^{2k}s\rangle^{-N}(d_k)^{\frac{1}{2}}. \tag{3-7}$$

(3) *Let* c_k *be a* $(-\delta_1, S)$*-frequency envelope for* (a,b) *in* $\dot{H}^1 \times L^2$. *Then*

$$\|P_k \boldsymbol{A}(s)\|_{\dot{H}^1} + \|P_k \boldsymbol{B}(s)\|_{L^2} \lesssim_{\mathcal{E},\mathcal{Q},N} \langle 2^{-2k}s^{-1}\rangle^{-\delta_1}\langle 2^{2k}s\rangle^{-N}(d_k)^{\frac{1}{2}}c_k, \tag{3-8}$$

$$\|P_k \partial^j A_j(s)\|_{L^2} + \|P_k \partial^j B_j(s)\|_{\dot{H}^{-1}} \lesssim_{\mathcal{E},\mathcal{Q},N} \langle 2^{-2k}s^{-1}\rangle^{-\delta_1}\langle 2^{2k}s\rangle^{-N}(d_k)^{\frac{1}{2}}c_k, \tag{3-9}$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ *are as in* (3-1).

3B. *Wave equation for A in caloric gauge.* Here, and in the rest of this paper, we shift the notation and denote by $A_{t,x} = A_{t,x}(t,x)$, instead of (a_0, a), the space-time connection on $I \times \mathbb{R}^4$ (viewed as $\{s = 0\}$). For the spatial components, we omit the subscript x and write $A_x(t,x) = A(t,x)$. We write $A_{t,x,s}(s) = A_{t,x,s}(t,x,s)$ for the dynamic Yang–Mills heat flow of $A_{t,x}(t,x)$.

In this subsection, we recall from [Oh and Tataru 2017b] the interpretation of the hyperbolic Yang–Mills equations for a space-time connection $A_{t,x}$ in the caloric gauge as a hyperbolic evolution for the spatial components A augmented with nonlinear expressions of $\partial^\ell A_\ell$, A_0 and $\partial_0 A_0$ in terms of $(A, \partial_t A)$; see Theorem 3.5. An analogous hyperbolic equation holds for the dynamic Yang–Mills heat flow development $A_{t,x}(s)$ of $A_{t,x}$ in the caloric gauge, which may be thought of as a gauge-covariant regularization of A; see Theorem 3.6.

We present explicit expressions for the quadratic nonlinearities, for which we need to reveal the null structure in order to handle them, and state stronger bounds for the remaining higher order nonlinearities. For economy of notation in the latter task, we introduce the following definition:

Definition 3.3. Let X, Y be dyadic norms.

• A map $\boldsymbol{F} : X \to Y$ is said to be *envelope-preserving of order* $\geq n$ ($n \in \mathbb{N}$ with $n \geq 2$) if the following property holds: Let c be a $(-\delta_1, S)$-frequency envelope for a in X. Then

$$\|\boldsymbol{F}(a)\|_{Y_{(c^{[\delta_1]})^{n-1}c}} \lesssim_{\|a\|_X} 1.$$

• A map $\boldsymbol{F} : X \to Y$ is said to be *Lipschitz envelope-preserving of order* $\geq n$ if, in addition to being envelope-preserving of order $\geq n$, the following additional property holds: Let c be a common δ_1-frequency envelope for a_1 and a_2 in X, and let d be a δ_1-frequency envelope for $a_1 - a_2$ in X that is δ_1-compatible with c. Then

$$\|P_k(\boldsymbol{F}(a_1) - \boldsymbol{F}(a_2))\|_{Y_k} \lesssim_{\|a_1\|_X, \|a_2\|_X} c_k^{n-2} e_k,$$

where $e_k = d_k + c_k (c \cdot d)_{\leq k}$.

Remark 3.4. The modified envelope e appears since the maps $\boldsymbol{F}$ that arise below are defined on a nonlinear manifold, namely, spatial connections a on a time interval I such that $(a, \partial_t a)(t) \in T^{L^2}\mathcal{C}$ for each fixed time. We remark moreover that if the frequency envelopes c and d are ℓ^2-summable, which is usually the case in practice, then $\boldsymbol{F}(a)$ and $\boldsymbol{F}(a_1) - \boldsymbol{F}(a_2)$ belong to $\ell^1 Y$.

We also need to introduce the nonsharp Strichartz spaces Str and Str^1, which scale like $L^\infty L^2$ and $L^\infty \dot{H}^1$, respectively. We define

$$\|u\|_{\mathrm{Str}} = \sup\Big\{\|u\|_{L^p \dot{W}^{\sigma,q}} : \frac{1}{q} + \frac{4}{p} = 2,\ \delta_0 \leq \frac{1}{p} \leq \frac{1}{2} - \delta_0,\ \frac{2}{p} + \frac{3}{q} \leq \frac{3}{2} - \delta_0\Big\}, \tag{3-10}$$

as well as

$$\|u\|_{\mathrm{Str}^1} = \|\nabla u\|_{\mathrm{Str}}. \tag{3-11}$$

Conditions in (3-10) ensure that the (p, q, σ)'s are Strichartz exponents, but away from the sharp endpoints. These norms have two key properties:

- They are divisible in time, i.e., can be made small by subdividing the time interval.
- Saturating the associated Strichartz inequalities requires strong pointwise concentration (i.e., small energy dispersion).

In [Oh and Tataru 2017b], we have shown that the spatial components of the Yang–Mills equation $\boldsymbol{D}^\alpha F_{j\alpha} = 0$ ($j \in \{1, 2, 3, 4\}$) may be interpreted as a system of wave equation for the spatial components $A = A_x$, where the temporal component A_0 is determined in terms of $(A, \partial_t A)$, as follows:

Theorem 3.5 [Oh and Tataru 2017b, Theorem 9.1]. *Let $A_{t,x} = (A_0, A) \in C_t(I; \dot{H}^1 \times \mathcal{C}_Q)$ with $(\partial_t A_0, \partial_t A) \in C_t(I; L^2 \times T^{L^2}_{A(t)}\mathcal{C}_Q)$ be a solution to (1-1) with energy $\mathcal{E}$. Then its spatial components $A = A_x$ satisfy an equation of the form*

$$\Box_A A_j = \boldsymbol{P}_j[A, \partial_x A] + 2\Delta^{-1}\partial_j \boldsymbol{Q}(\partial^\alpha A, \partial_\alpha A) + R_j(A), \tag{3-12}$$

together with a compatibility condition

$$\partial^\ell A_\ell = \boldsymbol{DA}(A) := \boldsymbol{Q}(A, A) + \boldsymbol{DA}^3(A). \tag{3-13}$$

Moreover, the temporal component A_0 and its time derivative $\partial_t A_0$ admit the expressions

$$A_0 = \boldsymbol{A}_0(A) := \Delta^{-1}[A, \partial_t A] + 2\Delta^{-1}\boldsymbol{Q}(A, \partial_t A) + \boldsymbol{A}_0^3(A), \tag{3-14}$$

$$\partial_t A_0 = \boldsymbol{DA}_0(A) := -2\Delta^{-1}\boldsymbol{Q}(\partial_t A, \partial_t A) + \boldsymbol{DA}_0^3(A). \tag{3-15}$$

Here $\boldsymbol{P}$ is the Leray projector, and $\boldsymbol{Q}$ is a symmetric[3] *bilinear form with symbol*

$$\boldsymbol{Q}(\xi, \eta) = \frac{|\xi|^2 - |\eta|^2}{2(|\xi|^2 + |\eta|^2)}. \tag{3-16}$$

[3] Observe here that the symbol of $\boldsymbol{Q}$ is odd, but this is combined with the antisymmetry of the Lie brackets appearing in the bilinear form.

Moreover, $R_j(t)$, $\boldsymbol{DA}^3(t)$, $\boldsymbol{A}_0^3(t)$ *and* $\boldsymbol{DA}_0^3(t)$ *are uniquely determined by* $(A, \partial_t A)(t) \in T^{L^2}\mathcal{C}$, *and are Lipschitz envelope-preserving maps of order* ≥ 3 *on the spaces*

$$R_j(t) : \dot{H}^1 \to \dot{H}^{-1}, \tag{3-17}$$

$$\boldsymbol{DA}^3(t) : \dot{H}^1 \to L^2, \tag{3-18}$$

$$\boldsymbol{A}_0^3(t) : \dot{H}^1 \to \dot{H}^1, \tag{3-19}$$

$$\boldsymbol{DA}_0^3(t) : \dot{H}^1 \to L^2. \tag{3-20}$$

Finally, on any interval $I \subseteq \mathbb{R}$, R_j, $\boldsymbol{DA}^3$, $\boldsymbol{A}_0^3$ *and* $\boldsymbol{DA}_0^3$ *are Lipschitz envelope-preserving maps of order* ≥ 3 (*with bounds independent of* I) *on the spaces*

$$R_j : \mathrm{Str}^1[I] \to L^1 L^2 \cap L^2 \dot{H}^{-\frac{1}{2}}[I], \tag{3-21}$$

$$\boldsymbol{DA}^3 : \mathrm{Str}^1[I] \to L^1 \dot{H}^1 \cap L^2 \dot{H}^{\frac{1}{2}}[I], \tag{3-22}$$

$$\boldsymbol{A}_0^3 : \mathrm{Str}^1[I] \to L^1 \dot{H}^2 \cap L^2 \dot{H}^{\frac{3}{2}}[I], \tag{3-23}$$

$$\boldsymbol{DA}_0^3 : \mathrm{Str}^1[I] \to L^1 \dot{H}^1 \cap L^2 \dot{H}^{\frac{1}{2}}[I]. \tag{3-24}$$

All implicit constants depend on $\mathcal{Q}$ *and* $\mathcal{E}$.

Next, we consider the dynamic Yang–Mills heat flow $A_{t,x}(s)$ of $A_{t,x}$ in the caloric gauge. For $s > 0$, we have $\boldsymbol{D}^\beta F_{\alpha\beta}(s) = w_\alpha \neq 0$ in general. We expect the "heat-wave commutator" w_α (called the Yang–Mills tension field) to be concentrated primarily at frequency comparable to $s^{-1/2}$. Indeed, the following theorem holds.

Theorem 3.6 [Oh and Tataru 2017b, Theorem 9.3]. *Let* $A_{t,x} = (A_0, A) \in C_t(I; \dot{H}^1 \times \mathcal{C}_\mathcal{Q})$ *with* $(\partial_t A_0, \partial_t A) \in C_t(I; L^2 \times T^{L^2}_{A(t)} \mathcal{C}_\mathcal{Q})$ *be a solution to* (1-1) *with energy* $\mathcal{E}$. *Let* $A_{t,x}(s) = A_{t,x}(t, x, s)$ *be the dynamic Yang–Mills heat flow development of* $A_{t,x}$ *in the caloric gauge. Then the spatial components* $A(s) = A_x(s)$ *of* $A_{t,x}(s)$ *satisfy an equation of the form*

$$\Box_{A(s)} A_j(s) = \boldsymbol{P}_j[A(s), \partial_x A(s)] + 2\Delta^{-1}\partial_j \boldsymbol{Q}(\partial^\alpha A(s), \partial_\alpha A(s)) + R_j(A(s)) \\ + \boldsymbol{P}_j \boldsymbol{w}_x^2(\partial_t A, \partial_t A, s) + R_{j;s}(A), \tag{3-25}$$

together with the compatibility condition

$$\partial^\ell A_\ell(s) = \boldsymbol{DA}(A(s)). \tag{3-26}$$

Moreover, *the temporal component* $A_0(s)$ *and its time derivative* $\partial_t A_0(s)$ *admit the expansions*

$$A_0(s) = \boldsymbol{A}_0(A(s)) + \boldsymbol{A}_{0;s}(A) := \boldsymbol{A}_0(A(s)) + \Delta^{-1}\boldsymbol{w}_0^2(A, A, s) + \boldsymbol{A}_{0;s}^3(A), \tag{3-27}$$

$$\partial_t A_0(s) = \boldsymbol{DA}_0(A(s)) + \boldsymbol{DA}_{0;s}(A). \tag{3-28}$$

Here $\boldsymbol{P}$, $\boldsymbol{Q}$, R_j, $\boldsymbol{DA}$, $\boldsymbol{A}_0$ and $\boldsymbol{DA}_0$ are as before, and the $\boldsymbol{w}^2_\alpha$ are defined as

$$\boldsymbol{w}_0^2(A, B, s) = -2\boldsymbol{W}(\partial_t A, \Delta B, s), \tag{3-29}$$

$$\boldsymbol{w}_j^2(A, B, s) = -2\boldsymbol{W}(\partial_t A, \partial_j \partial_t B - 2\partial_x \partial_t B_j, s), \tag{3-30}$$

where $\boldsymbol{W}(\cdot, \cdot, s)$ is a bilinear form with symbol

$$\boldsymbol{W}(\xi, \eta, s) = -\frac{1}{2\xi \cdot \eta} e^{-s|\xi+\eta|^2}(1 - e^{2s(\xi\cdot\eta)}). \tag{3-31}$$

Moreover, $R_{j;s}(t)$, $\boldsymbol{A}^3_{0;s}(t)$ and $\boldsymbol{DA}_{0;s}(t)$ are uniquely determined by $(A, \partial_t A)(t) \in T^{L^2}\mathcal{C}$ for each $s > 0$, and satisfy the following properties:

- $R_{j;s}(t) : \dot{H}^1 \to \dot{H}^{-1}$ *is a Lipschitz map with output concentrated at frequency $s^{-1/2}$. More precisely,*

$$(1 - s\Delta)^N R_{j;s}(t) : \dot{H}^1 \to 2^{-\delta_1 k(s)} \dot{H}^{-1-\delta_1}. \tag{3-32}$$

- $\boldsymbol{A}^3_{0;s}(t) : \dot{H}^1 \to \dot{H}^1$ *is a Lipschitz map with output concentrated at frequency $s^{-1/2}$; i.e.,*

$$(1 - s\Delta)^N \boldsymbol{A}^3_{0;s}(t) : \dot{H}^1 \to 2^{-\delta_1 k(s)} \dot{H}^{1-\delta_1}. \tag{3-33}$$

- $\boldsymbol{DA}_{0;s}(t) : \dot{H}^1 \to L^2$ *is a Lipschitz map with output concentrated at frequency $s^{-1/2}$; i.e.,*

$$(1 - s\Delta)^N \boldsymbol{DA}_{0;s}(t) : \dot{H}^1 \to 2^{-\delta_1 k(s)} \dot{H}^{-\delta_1}. \tag{3-34}$$

Finally, on any time interval $I \subseteq \mathbb{R}$ (with bounds independent of I), $R_{j;s}$, $\boldsymbol{A}^3_{0;s}$ and $\boldsymbol{DA}_{0;s}$ satisfy the following properties:

- $R_{j;s} : \mathrm{Str}^1[I] \to L^1L^2 \cap L^2\dot{H}^{-1/2}[I]$ *is a Lipschitz map with output concentrated at frequency $s^{-1/2}$; i.e.,*

$$(1 - s\Delta)^N R_{j;s} : \mathrm{Str}^1[I] \to 2^{-\delta_1 k(s)}(L^1\dot{H}^{-\delta_1} \cap L^2\dot{H}^{-\frac{1}{2}-\delta_1})[I]. \tag{3-35}$$

- $\boldsymbol{A}^3_{0;s} : \mathrm{Str}^1[I] \to L^1\dot{H}^2 \cap L^2\dot{H}^{3/2}[I]$ *is a Lipschitz map with output concentrated at frequency $s^{-1/2}$; i.e.,*

$$(1 - s\Delta)^N \boldsymbol{A}^3_{0;s} : \mathrm{Str}^1[I] \to 2^{-\delta_1 k(s)}(L^1\dot{H}^{2-\delta_1} \cap L^2\dot{H}^{\frac{3}{2}-\delta_1})[I]. \tag{3-36}$$

- $\boldsymbol{DA}_{0;s} : \mathrm{Str}^1[I] \to L^2\dot{H}^{1/2}[I]$ *is a Lipschitz map with output concentrated at frequency $s^{-1/2}$; i.e.,*

$$(1 - s\Delta)^N \boldsymbol{DA}_{0;s} : \mathrm{Str}^1[I] \to 2^{-\delta_1 k(s)} L^2\dot{H}^{\frac{1}{2}-\delta_1}[I]. \tag{3-37}$$

All implicit constants depend on $\mathcal{Q}$ and $\mathcal{E}$.

Remark 3.7. Some notable features of Theorem 3.6 are as follows:

- Compared with the prior result, here we have additional contributions $R_{k;s}$, $\boldsymbol{A}_{0;s}$ and $\boldsymbol{DA}_{0;s}$ as well as the $\boldsymbol{w}$ terms. These have the downside that they depend on A and $\partial_t A$ at $s = 0$ rather than $A(s)$ and $\partial_t A(s)$. The redeeming feature is that these terms will not only be small due to the energy dispersion, but also, critically, concentrated at frequency $s^{-1/2}$.
- The other change here is due to the inhomogeneous terms $\boldsymbol{w}^2_\alpha$; these are matched in the $A_k(s)$ and the $A_0(s)$ equations, and will interact in the trilinear analysis (see Proposition 4.29 below).
- For the new error terms here we do not need to worry about difference bounds; see Section 6 below.

4. Summary of function spaces and estimates

In this section, we summarize the properties of the function spaces and the estimates needed to analyze the hyperbolic Yang–Mills equation in the caloric gauge, as given by Theorems 3.5 and 3.6.

4A. ***Function spaces.*** The aim of this subsection is to give precise definitions of the fine functions spaces used to analyze caloric Yang–Mills waves.

4A1. *Frequency projections.* We start with a brief discussion of various frequency projections. Let $m_0 : \mathbb{R} \to \mathbb{R}$ be a smooth nonnegative even bump function supported on $\{x \in \mathbb{R} : |x| \in (2^{-1}, 2^2)\}$ such that $\{m_k = m_0(\cdot/2^k)\}_{k\in\mathbb{Z}}$ is a partition of unity on $\mathbb{R}$. For $k \in \mathbb{Z}$, recall that P_k was defined as the multiplier on $\mathbb{R}^4$ with symbol $P_k(\xi) = m_k(|\xi|)$. Given $j \in \mathbb{Z}$ and a sign $\pm$, we introduce the modulation projections $Q_j^{\pm}$ and Q_j, which are multipliers on $\mathbb{R}^{1+4}$ with symbols

$$Q_j^{\pm}(\tau,\xi) = m_j(\tau \mp |\xi|), \quad Q_j(\tau,\xi) = m_j(|\tau| - |\xi|).$$

We also define $Q^{\pm}_{<j}$, $Q^{\pm}_{\geq j}$, $Q_{<j}$, $Q_{\geq j}$ etc. in the obvious manner. To connect $Q_j^{\pm}$ with Q_j, we introduce the sharp time-frequency cutoffs $Q^{\pm}$, which are multipliers on $\mathbb{R}^{1+4}$ with symbols

$$Q^{\pm}(\tau,\xi) = \chi_{(0,\infty)}(\pm\tau).$$

Note that $P_k Q^{\pm} Q_j = P_k Q_j^{\pm}$ for $j < k$.

For $\ell \in -\mathbb{N}$, consider a collection of directions $\omega \in \mathbb{S}^3 \subseteq \mathbb{R}^4$, which are maximally separated with distance $\simeq 2^{\ell}$. To each such an ω, we associate a smooth cutoff function m_ℓ^ω supported on a cap of radius $\simeq 2^\ell$ centered at ω, with the property that $\sum_\omega m_\omega = 1$. Let P_ℓ^ω be the multiplier on $\mathbb{R}^4$ with symbol

$$P_\ell^\omega(\xi) = m_\ell^\omega\left(\frac{\xi}{|\xi|}\right).$$

Given $k' \in \mathbb{Z}$ and $\ell' \in -\mathbb{N}$, consider rectangular boxes $\mathcal{C}_{k'}(\ell')$ of dimensions $2^{k'} \times (2^{k'+\ell'})^3$ (where the $2^{k'}$-side lies along the radial direction), which cover $\mathbb{R}^4 \setminus \{|x| \lesssim 2^{k'}\}$ and have finite overlap with each other. Let $m_{\mathcal{C}_{k'}(\ell')}$ b a partition of unity adapted to $\{\mathcal{C}_{k'}(\ell')\}$, and we define the multiplier $P_{\mathcal{C}_{k'}(\ell')}$ on $\mathbb{R}^4$ with symbol

$$P_{\mathcal{C}_{k'}(\ell')}(\xi) = m_{\mathcal{C}_{k'}(\ell')}(\xi).$$

For convenience, when $k' = k$, we choose the covering and the partition of unity so that $P_k P_\ell^\omega = P_k P_{\mathcal{C}_k(\ell)}$.

We now discuss the boundedness properties of the frequency projections. For any $k \in \mathbb{Z}$, let $P_{k/<k}$ denote one of the dyadic frequency projections $\{P_k, P_{<k}\}$. Let $Q^{\square}_{j/<j}$ denote one of the modulation projections $Q_j^{\pm}$, $Q^{\pm}_{<j}$, Q_j or $Q_{<j}$. Let ω be an angular sector of size $\simeq 2^\ell$ ($\ell \in -\mathbb{N}$), and $\mathcal{C}$ a rectangular box of the form $\mathcal{C}_{k'}(\ell')$ ($k' \in \mathbb{Z}$, $\ell' \in -\mathbb{N}$). Then the following statements hold:

- The multipliers $P_{k/<k}$, $P_{k/<k} P_\ell^\omega$ and $P_{\mathcal{C}}$ are disposable.
- The multiplier $P_{k/<k} Q^{\square}_{j/<j}$ is disposable if $j \geq k + O(1)$; see [Tao 2001, Lemma 3]. For general $j, k \in \mathbb{Z}$, it is straightforward to check that $P_{k/<k} Q^{\square}_{j/<j}$ has a kernel with mass $O(2^{4(k-j)_+})$.

- The multiplier $P_{k/<k} Q^{\square}_{j/<j}$ is bounded on $L^p L^2$ for any $1 \le p \le \infty$; see [Tao 2001, Lemma 4].
- The multiplier $P_{k/<k} P^{\omega}_{\ell} Q^{\square}_{j/<j}$ is disposable if $j \ge k + 2\ell + O(1)$; see [Tao 2001, Lemma 6].

4A2. *Function spaces on the whole space-time.* Here, we define the global-in-time function spaces used in this work. Unless otherwise stated, all spaces below are defined for functions on $\mathbb{R}^{1+4}$. We remark that all of them are translation-invariant.

We first define the space $X^{\sigma,b}_r$, equipped with the norm

$$\|u\|^2_{X^{\sigma,b}_r} = \sum_k 2^{2\sigma k} \Big(\sum_j (2^{bj} \|P_k Q_j u\|_{L^2 L^2})^r \Big)^{\frac{2}{r}}$$

when $1 \le r < \infty$. As usual, we replace the ℓ^r-sum by the supremum in j when $r = \infty$. The spaces $X^{\sigma,b}_{\pm,r}$ are defined similarly, with Q_j replaced by $Q^{\pm}_j$.

We are now ready to introduce the function spaces in earnest, which are all defined in terms of (semi-)norms.

Core nonlinearity norm N. We define

$$N = L^1 L^2 + X^{0,-\frac{1}{2}}_1.$$

This norm scales like $L^1 L^2$. We also define $N_{\pm} = L^1 L^2 + X^{0,-1/2}_{\pm,1}$. Note that $N = N_+ \cap N_-$. Moreover, we have the embeddings

$$X^{0,-\frac{1}{2}}_1 \subseteq N \subseteq X^{0,-\frac{1}{2}}_{\infty}, \quad X^{0,-\frac{1}{2}}_{\pm,1} \subseteq N \subseteq X^{0,-\frac{1}{2}}_{\pm,\infty}.$$

The inclusions on the left are obvious, whereas the inclusions on the right follow from Bernstein in time. We omit the proofs.

Core solution norm S. We define

$$\|u\|^2_S = \sum_k \|u\|^2_{S_k}, \quad S_k = S^{\mathrm{str}}_k \cap X^{0,\frac{1}{2}}_{\infty} \cap S^{\mathrm{ang}}_k \cap S^{sq}_k,$$

where S^{sq}_k is related to square function bounds,

$$\|u\|_{S^{sq}_k} = 2^{-\frac{3}{10}k} \|u\|_{L^{10/3}_x L^2_t}$$

and S^{str}_k and S^{ang}_k are essentially as in [Krieger et al. 2015, equations (6)–(8)]:

$$\|u\|_{S^{\mathrm{str}}_k} = \sup_{(p,q):\frac{1}{p}+\frac{3}{2q}\le\frac{3}{4}} 2^{-(2-\frac{1}{p}-\frac{4}{q})k} \|u\|_{L^p L^q},$$

$$\|u\|^2_{S^{\mathrm{ang}}_k} = \sup_{\ell<0} \sum_{\omega} \|P^{\omega}_{\ell} Q_{<k+2\ell} u\|^2_{S^{\omega}_k(\ell)},$$

$$\begin{aligned}\|u\|^2_{S^{\omega}_k(\ell)} = \|u\|^2_{S^{\mathrm{str}}_k} + 2^{-2k}\|u\|^2_{NE} + 2^{-3k} \sum_{\pm} \|Q^{\pm} u\|^2_{PW^{\mp}_{\omega}(\ell)} \\ + \sup_{\substack{k' \le k,\ \ell' \le 0 \\ k+2\ell \le k'+\ell' \le k+\ell}} \sum_{\mathcal{C}_{k'}(\ell')} \Big(\|P_{\mathcal{C}_{k'}(\ell')} u\|^2_{S^{\mathrm{str}}_k} + 2^{-2k} \|P_{\mathcal{C}_{k'}(\ell')} u\|^2_{NE} \\ + 2^{-2k'-k} 2^{-\ell'} \|P_{\mathcal{C}_{k'}(\ell')} u\|^2_{L^2 L^\infty} \\ + 2^{-3(k'+\ell')} \sum_{\pm} \|Q^{\pm} P_{\mathcal{C}_{k'}(\ell')} u\|^2_{PW^{\mp}_{\omega}(\ell)} \Big).\end{aligned}$$

Here, the NE and $PW^{\mp}_{\omega}(\ell)$ are the *null frame spaces* [Tataru 2001; Tao 2001], defined by

$$\|u\|_{PW^{\mp}_{\omega}(\ell)} = \inf_{u=\int u^{\omega'}} \int_{|\omega-\omega'|\le 2^{\ell}} \|u^{\omega'}\|_{L^2_{\pm\omega'} L^{\infty}_{(\pm\omega')^{\perp}}}\, d\omega',$$
$$\|u\|_{NE} = \sup_{\omega} \|\not\nabla_{\omega} u\|_{L^{\infty}_{\omega} L^2_{\omega^{\perp}}},$$

where the L^q_{ω} norm is with respect to the variable $t^{\pm}_{\omega} = t \pm \omega\cdot x$, the $L^r_{\omega^{\perp}}$ norm is defined on each $\{t^{\pm}_{\omega} = \text{const}\}$, and $\not\nabla_{\omega}$ denotes the tangential derivatives to $\{t^{\pm}_{\omega} = \text{const}\}$.

In the last two lines of the definition of $S^{\omega}_k(\ell)$, the restrictions $k' \le k,\ \ell' \le 0$ and $k'+\ell' \le k+\ell$ ensure that rectangular boxes of the form $\mathcal{C}_{k'}(\ell')$ fit in the frequency support of P^{ω}_{ℓ}. The restriction $k+2\ell \le k'+\ell'$ is imposed by the main parametrix estimate (see Section 10H or [Krieger et al. 2015, Section 11]), to ensure square-summability in $\mathcal{C}_{k'}(\ell')$.

The null frame spaces in $S^{\omega}_k(\ell)$ allow one to exploit transversality in frequency space, and play an important role in the proof of the trilinear null form estimate; see [Krieger et al. 2015, equations (136)–(138)] and Proposition 8.18 below. On the other hand, the L^2L^{∞}-norm for $P_{\mathcal{C}_{k'}(\ell')}u$ allows us to gain the dimensions of $\mathcal{C}_{k'}(\ell')$.

Remark 4.1. For the reader who is familiar with the function space framework in [Krieger et al. 2015], we point out that our $S^{\omega}_k(\ell)$ is slightly stronger than that in [loc. cit.]. More precisely, instead of $2^{-k'-(1/2)k}2^{-(1/2)\ell'}\|P_{\mathcal{C}_{k'}(\ell')}u\|_{L^2L^{\infty}}$ as in our definition, it is $2^{-k'-(1/2)k}\|P_{\mathcal{C}_{k'}(\ell')}u\|_{L^2L^{\infty}}$ in [loc. cit.]. However, we note that the extra factor $2^{-(1/2)\ell'}$ is actually present in the main parametrix estimate in [loc. cit., Subsection 11.3].

Remark 4.2. The square function norm S^{sq}_k is new here in the structure of the S norm. It plays no role in the study of the solutions for the hyperbolic Yang–Mills equation in the caloric gauge, i.e., in Theorems 1.13 and 1.16. Instead, it is only needed in order to justify the transition to the temporal gauge in Theorem 1.18.

This norm scales like $L^{\infty}L^2$. Moreover, it obeys the embeddings

$$P_k X^{0,\frac{1}{2}}_1 \subseteq S_k, \quad S_k \subseteq X^{0,\frac{1}{2}}_{\infty}.$$

For $k, k' \in \mathbb{Z}$ satisfying $k' \le k$ and $\ell' < -5$, we define

$$\|u\|^2_{S_k[\mathcal{C}_{k'}(\ell')]} = 2^{-\frac{5}{3}k}\|u\|^2_{L^2L^6} + 2^{-2k'-k}2^{-\ell'}\|u\|^2_{L^2L^{\infty}} + \sup_{j:|j-(k'+2\ell')|\le 5}\Big(\|Q_{<j}u\|^2_{L^{\infty}L^2} + 2^{-2k}\|Q_{<j}u\|^2_{NE} + 2^{-3(k'+\ell')}\sum_{\pm}\|Q^{\pm}_{<j}u\|^2_{PW^{\mp}_{\omega}(\frac{j-k}{2})}\Big).$$

The virtue of this norm is that it is square-summable in boxes of the form $\mathcal{C}_{k'}(\ell')$:

Lemma 4.3. *For any k, k', ℓ' such that $k' \le k$ and $\ell' \le 0$, we have*

$$\sum_{\mathcal{C}\in\{\mathcal{C}_{k'}(\ell')\}} \|P_{\mathcal{C}}u\|^2_{S_k[\mathcal{C}_{k'}(\ell')]} \lesssim \|u\|^2_{S_k}. \tag{4-1}$$

Proof. The desired square-summability estimates for the $L^\infty L^2$, NE and $PW_\omega^\mp$ components follow immediately from the definition of $S_k^{\mathrm{ang}} \supseteq S_k$. For the $L^2 L^6$ and $L^2 L^\infty$ components, we write

$$u = Q_{<k'+2\ell'} u + Q_{\geq k'+2\ell'} u.$$

For the former we use S_k^{ang}, and for the latter we simply note that, by Bernstein,

$$2^{-\frac{5}{6}k} \| Q_{\geq k'+2\ell'} P_{\mathcal{C}_{k'}(\ell')} u\|_{L^2 L^6} + 2^{-k'-\frac{1}{2}k} 2^{-\frac{1}{2}\ell'} \| Q_{\geq k'+2\ell'} P_{\mathcal{C}_{k'}(\ell')} u\|_{L^2 L^\infty} \lesssim \| P_{\mathcal{C}_{k'}(\ell')} u\|_{X_\infty^{0,1/2}},$$

which is clearly square-summable. □

Sharp solution norm $S^\sharp$. We define

$$\|u\|_{S_k^\sharp} = 2^{-k} (\|\nabla u\|_{L^\infty L^2} + \|\Box u\|_N),$$
$$\|u\|_{(S_\pm^\sharp)_k} = \|u\|_{L^\infty L^2} + \|(D_t \mp |D|) u\|_{N_\pm},$$

both of which scale like $L^\infty L^2$. These norms are used in the parametrix construction in Section 9.

Remark 4.4. Again for the reader familiar with [Krieger et al. 2015], we note that our definition of $S_k^\sharp$ differs from that in [loc. cit.] by a factor of 2^k (in [loc. cit.], $S_k^\sharp$ scales like $L^\infty \dot{H}^1$).

Scattering (*or controlling*) *norm* S^1. Given any $\sigma \in \mathbb{R}$, we define $S^\sigma = \ell^2 S^\sigma$, i.e.,

$$\|u\|_{S^\sigma}^2 = \sum_k \|P_k u\|_{S_k^\sigma}^2, \quad \|u\|_{S_k^\sigma} = 2^{(\sigma-1)k} (\|\nabla u\|_S + \|\Box u\|_{L^2 \dot{H}^{-1/2}}). \tag{4-2}$$

This norm scales like $L^\infty \dot{H}^\sigma$. The norm S^1 will be the main scattering (or controlling) norm, in the sense that finiteness of this norm for a caloric Yang–Mills wave would imply finer properties of the solution itself and those nearby (see Theorem 5.1 below).

$X_r^{\sigma,b,p}$*-type norms.* To close the estimates for caloric Yang–Mills waves, we need norms which give additional control[4] off the characteristic cone (i.e., "high" modulation regime). We use an $L^p L^{p'}$ generalization of the usual $L^2 L^2$-based $X^{\sigma,b}$-norm, defined as follows: for $\sigma, b \in \mathbb{R}$, $1 \leq p, r < \infty$, let

$$\|u\|_{(X_r^{\sigma,b,p})_k} = 2^{\sigma k} \left(\sum_j \left(2^{bj} \left(\sum_\omega \|P_k Q_j P_{\frac{j-k}{2}}^\omega u\|_{L^p L^{p'}}^2 \right)^{\frac{1}{2}} \right)^r \right)^{\frac{1}{r}}, \tag{4-3}$$

where $p' = \frac{p}{p-1}$ is the dual Lebesgue exponent of p. The cases $p = \infty$ and $r = \infty$ are defined in the obvious manner. We also define the dyadic norm $(X_{\pm,r}^{\sigma,b,p})_k$ by replacing Q_j by $Q_j^\pm$ in the above definition.

When $p = 2$, by orthogonality we have

$$\|u\|_{(X_r^{\sigma,b,2})_k} = 2^{\sigma k} \left(\sum_j (2^{bj} \|P_k Q_j u\|_{L^2 L^2})^r \right)^{\frac{1}{r}}.$$

[4]In particular, with ℓ^1-summability in dyadic frequencies.

Analogous identities hold for $X^{\sigma,b,2}_{\pm,r}$. To be consistent with the usual notation, we will often omit the exponents p and r when they are equal to 2, i.e., $X^{\sigma,b}_r = X^{\sigma,b,2}_r$, $X^{\sigma,b} = X^{\sigma,b,2}_2$, $X^{\sigma,b}_{\pm,r} = X^{\sigma,b,2}_{\pm,r}$ and $X^{\sigma,b}_{\pm} = X^{\sigma,b,2}_{\pm,2}$.

Before we introduce the specific norms we use, for logical clarity, we first fix the parameters that will be used. We introduce b_0, b_1 and p_0, which are smaller than but close to $\frac{1}{4}$, $\frac{1}{2}$ and ∞, respectively. More precisely, we fix

$$b_0 = \frac{1}{4} - \delta_0, \quad b_1 = \frac{1}{2} - 10\delta_0, \quad 1 - \frac{1}{p_0} = 5\delta_0,$$

so that

$$0 < \frac{1}{4} - b_0 < \frac{1}{48}, \quad 2\Big(\frac{1}{4} - b_0\Big) < 1 - \frac{1}{p_0} < \frac{1}{24}, \tag{4-4}$$

$$\frac{1}{4} < b_1 < \frac{1}{2} - \Big(1 - \frac{1}{p_0}\Big). \tag{4-5}$$

We define

$$\|f\|_{\Box Z^1_k} = \|Q_{<k+C} f\|_{X_1^{-5/4-b_0,-3/4+b_0,1}},$$

$$\|u\|_{Z^1_k} = \|\Box u\|_{\Box Z^1_k} = \|Q_{<k+C} u\|_{X_1^{-1/4-b_0,1/4+b_0,1}}.$$

Note that the Z^1_k-norm scales like $L^\infty \dot{H}^1$. As in [Krieger et al. 2015; Krieger and Tataru 2017], this norm is used as an auxiliary device to control the bulk of nonlinearities (i.e., the part where the secondary null structure is *not* necessary) when reiterating the Yang–Mills equations; see the proofs of Propositions 4.23–4.29 in Section 8.

Remark 4.5. The Z^1-norm used in [Krieger et al. 2015] corresponds to the case $b_0 = 0$. Therefore, our Z^1-norm is *weaker* than the Z^1-norm in [loc. cit.]. This modification is made to handle the contribution of $\Box^{-1} \boldsymbol{P}[A^\alpha, \partial_\alpha A]$ in the reiteration procedure; see Proposition 4.22.

Next, we also define

$$\|f\|_{(\Box Z^1_{p_0})_k} = \|Q_{<k+C} f\|_{X_\infty^{3/2-3/p_0+(1/4-b_0)\theta_0,-1/2-(1/4-b_0)\theta_0,p_0}},$$

where $\theta_0 = 2\big(\frac{1}{p_0} - \frac{1}{2}\big)$, as well as the intermediate norm

$$\|f\|_{(\Box \tilde{Z}^1_{p_0})_k} = \|Q_{<k+C} f\|_{X_1^{5/4-3/p_0+(1/4-b_0)\theta_0,-1/4-(1/4-b_0)\theta_0,p_0}}.$$

These norms scale like $L^1 L^2$. Clearly, $(\Box Z^1_{p_0})_k \subseteq (\Box \tilde{Z}^1_{p_0})_k$. Given any caloric Yang–Mills wave A with a finite S^1-norm, we will put $\Box \boldsymbol{P} A$ in $\ell^1 \Box \tilde{Z}^1_{p_0}$ and $\Box \boldsymbol{P} A \in \ell^1 \Box Z^1_{p_0}$; see Proposition 5.4.

Note that the following embeddings hold:

$$P_k Q_j L^1 L^2 \subseteq 2^{\frac{1}{4}(j-k)} \Box Z^1_k, \tag{4-6}$$

$$X_\infty^{0,-\frac{1}{2}} \cap \Box Z^1_k \subseteq (\Box Z^1_{p_0})_k \subseteq (\Box \tilde{Z}^1_{p_0})_k. \tag{4-7}$$

Estimate (4-6) follows from Bernstein, whereas the first embedding in (4-7) follows by a simple interpolation argument. We omit the straightforward proofs.

Finally, as in [Krieger and Tataru 2017], we also need to use the function space

$$\ell^1 X^{-\frac{1}{2}+b_1,-b_1},$$

which also scales like L^1L^2. Given any caloric Yang–Mills wave A with a finite S^1-norm, we will be able to place $\Box \boldsymbol{P}A$ in $\ell^1 X^{-1/2+b_1,-b_1}$. This bound, in turn, is used crucially in the parametrix construction.

High-modulation norms $\underline{X}^1$ and $\widetilde{X}^1$ for 1-forms. In our analysis below, we need to use different high-modulation norms for the Leray projection $\boldsymbol{P}A$ than for the general components of a caloric Yang–Mills wave. Hence it is convenient to define norms for 1-forms with this distinction built in.

Let A and G be spatial 1-forms on $\mathbb{R}^{1+4}$. We define

$$\|G\|_{\Box \underline{X}^1_k} = \|G\|_{L^2\dot{H}^{-1/2}} + \|G\|_{L^{9/5}\dot{H}^{-4/9}} + \|\boldsymbol{P}G\|_{(\Box Z^1_{p_0})_k}.$$

For any $\sigma \in \mathbb{R}$, we define

$$\|G\|_{\Box \underline{X}^\sigma_k} = 2^{(\sigma-1)k}\|G\|_{\Box \underline{X}^1_k}, \quad \|A\|_{\underline{X}^\sigma_k} = \|\Box A\|_{\Box \underline{X}^\sigma_k}.$$

Similarly, we define

$$\|G\|_{\Box \widetilde{X}^1_k} = \|G\|_{L^2\dot{H}^{-1/2}} + \|G\|_{L^{9/5}\dot{H}^{-4/9}} + \|\boldsymbol{P}G\|_{(\Box \widetilde{Z}^1_{p_0})_k},$$

as well as $\Box\widetilde{X}^\sigma_k$ and $\widetilde{X}^\sigma_k$. Given any caloric Yang–Mills wave A with a finite S^1-norm, we will place $\Box A$ successively in $\ell^1\Box\widetilde{X}^1$ and $\Box A \in \ell^1\Box\underline{X}^1$; see Proposition 5.4.

We have the embeddings

$$P_k(L^1L^2 \cap L^2\dot{H}^{-\frac{1}{2}}) \subseteq (\Box\underline{X}^1)_k \subseteq (\Box\widetilde{X}^1)_k.$$

Since $L^1L^2 \subseteq N$, it follows that

$$\|G\|_{N\cap\Box\underline{X}^1} \lesssim \|G\|_{L^1L^2\cap L^2\dot{H}^{-1/2}}. \tag{4-8}$$

Strengthened solution norm $\underline{S}^1$. Putting together S^1 and $\underline{X}^1$, for a 1-form A on $\mathbb{R}^{1+4}$, we define

$$\|A\|_{\underline{S}^\sigma_k} = \|A\|_{S^\sigma_k} + \|\Box A\|_{\Box\underline{X}^\sigma_k}.$$

Core elliptic norm Y. We return to functions u on $\mathbb{R}^{1+4}$. We define

$$\|u\|_{Y_k} = \|u\|_{L^2\dot{H}^{1/2}} + \|u\|_{L^{p_0}\dot{W}^{2-3/p_0,p_0'}},$$

where p_0 was fixed in (4-4) above. This norm scales like $L^\infty L^2$.

Main elliptic norm Y^1. For $\sigma \in \mathbb{R}$, we define

$$\|u\|^2_{Y^\sigma} = \sum_k \|P_k u\|^2_{Y^\sigma_k}, \quad \|u\|_{Y^\sigma_k} = 2^{\sigma k}(\|u\|_{Y_k} + 2^{-k}\|\partial_t u\|_{L^2\dot{H}^{1/2}}).$$

This norm scales like $L^\infty\dot{H}^\sigma$. We will put the elliptic components A_0 and $\boldsymbol{P}^\perp A = \Delta^{-1}\partial_x\partial^\ell A_\ell$ of a caloric Yang–Mills wave in Y^1.

4A3. *Interval localization and extension.* So far, the function spaces have been defined over the whole space-time $\mathbb{R}^{1+4}$. In our analysis, we also need to consider localization of these spaces on finite time intervals. We use the same set-up as [Oh and Tataru 2018; Krieger and Tataru 2017].

For most of our function spaces (with the important exceptions of $Z^1_{p_0}$, $\widetilde{Z}^1_{p_0}$, $\underline{X}^1$ and $\widetilde{X}^1$; see below), we take a simple route and define the interval-localized counterparts by restriction. In particular, given a time interval $I \subseteq \mathbb{R}$, we define

$$\|u\|_{S^\sigma[I]} = \inf_{\tilde{u} \in S^\sigma : u = \tilde{u}\restriction_I} \|\tilde{u}\|_{S^\sigma}, \quad \|u\|_{S[I]} = \inf_{\tilde{u} \in S : u = \tilde{u}\restriction_I} \|\tilde{u}\|_{S}, \quad \|f\|_{N[I]} = \inf_{\tilde{f} \in N : f = \tilde{f}\restriction_I} \|\tilde{f}\|_{N}. \tag{4-9}$$

An important technical question then is that of finding a common extension procedure outside I which preserves these norms. The following proposition provides an answer.

Proposition 4.6. *Let I be a time interval.*

(1) *Let χ_I be the characteristic function of I. Then we have the bounds*

$$\|\chi_I u\|_S \lesssim \|u\|_S, \quad \|\chi_I f\|_N \lesssim \|f\|_N. \tag{4-10}$$

For a fixed function f on $\mathbb{R}^{1+4}$, the norms $\|\chi_I f\|_N$ and $\|f\|_{N[I]}$ are also continuous as a function of the endpoints of I. We also have the linear estimates

$$\|\nabla u\|_{S[I]} \lesssim \|\nabla u(0)\|_{L^2} + \|\Box u\|_{N[I]}, \tag{4-11}$$

$$\|u\|_{S^1[I]} \lesssim \|\nabla u(0)\|_{L^2} + \|\Box u\|_{N \cap L^2 \dot{H}^{-1/2}[I]}. \tag{4-12}$$

(2) *Consider any partition $I = \bigcup_k I_k$. Then the N and $L^2\dot{H}^{-1/2}$ are interval square divisible, i.e.,*

$$\sum_k \|f\|^2_{N[I_k]} \lesssim \|f\|^2_{N[I]}, \quad \sum_k \|f\|^2_{L^2\dot{H}^{-1/2}[I_k]} \lesssim \|f\|^2_{L^2\dot{H}^{-1/2}[I]}, \tag{4-13}$$

and the S and S^1 are interval square summable, i.e.,

$$\|u\|^2_{S[I]} \lesssim \sum_k \|u\|^2_{S[I_k]}, \quad \|u\|^2_{S^1[I]} \lesssim \sum_k \|u\|^2_{S^1[I_k]}. \tag{4-14}$$

For a proof, we refer to [Oh and Tataru 2018, Proposition 3.3].

Remark 4.7. As a consequence of part (1), up to equivalent norms, we can replace the arbitrary extension in (4-9) by the zero extension in the case of S and N, and by the homogeneous waves with $(\phi, \partial_t \phi)$ at each endpoint as data outside I in the case of S^1.

The elliptic norms Y and Y^1 only involve spatial multipliers and norms of the form $L^p L^q$, so their interval-localization $Y[I]$ and $Y^1[I]$ are obviously defined (either by restriction, or using the $L^p L^q[I]$-norm; both are equivalent). In particular, in the case of Y, observe that

$$\|u\|_{Y[I]} = \|\chi_I u\|_Y \leq \|u\|_Y,$$

so the zero extension can be used.

On the other hand, given a function u on I, we *directly* define the $\|u\|_{(Z^1_{p_0})_k[I]}$ (resp. $\|u\|_{(\tilde{Z}_{p_0})^1_k[I]}$) to be $\|u^{\mathrm{ext}}\|_{(Z^1_{p_0})_k[I]}$ (resp. $\|u^{\mathrm{ext}}\|_{(\tilde{Z}_{p_0})^1_k[I]}$), where u^{ext} is the extension of u outside I by homogeneous waves. Equivalently, for $(\Box Z^1_{p_0})_k$ and $(\Box \widetilde{Z}^1_{p_0})_k$, we define

$$\|f\|_{(\Box Z^1_{p_0})_k[I]} = \|\chi_I f\|_{(\Box Z^1_{p_0})_k}, \quad \|f\|_{(\Box \tilde{Z}^1_{p_0})_k[I]} = \|\chi_I f\|_{(\Box Z^1_{p_0})_k}.$$

Accordingly, we define

$$\|G\|_{\Box \underline{X}^1_k[I]} = \|G\|_{L^2 \dot{H}^{-1/2}[I]} + \|G\|_{L^{9/5}\dot{H}^{-4/9}[I]} + \|\chi_I \boldsymbol{P} G\|_{(\Box Z^1_{p_0})_k}, \quad \|A\|_{\underline{X}^1_k[I]} = \|\Box A\|_{\Box \underline{X}^1_k[I]},$$

and similarly for $\Box \widetilde{X}^1[I]$ and $\widetilde{X}^1[I]$.

The advantage of this definition is clear: We may thus use a common extension procedure (namely, by homogeneous waves) for S^1 and $\underline{X}^1$. The price we pay is that in estimating the $\Box Z^1_{p_0}$- and the $\Box \widetilde{Z}^1_{p_0}$-norms, we need to carefully absorb the sharp time cutoff χ_I.

4A4. *Sources of smallness: divisibility, energy dispersion and short time interval.* In this work, we rely on several sources of smallness for analysis of caloric Yang–Mills waves.

One important source of smallness is divisibility, which refers to the property of a norm on an interval that it can be made arbitrarily small by splitting the interval into a controlled number of pieces. Unfortunately, our main function space $S^1[I]$ is far from satisfying such a property (see, however, Theorem 5.1(6) below), which causes considerable difficulty. Our workaround, as in [Oh and Tataru 2018], is to utilize a weaker yet divisible norm

$$\|u\|_{DS^1[I]} = \||D|^{-\frac{5}{6}} \nabla u\|_{L^2 L^6[I]} + \|\nabla u\|_{\mathrm{Str}^0[I]} + \|\Box u\|_{L^2 \dot{H}^{-1/2}[I]}. \tag{4-15}$$

Another important source of smallness is *energy dispersion*:

Definition 4.8. Given any $m \in \mathbb{Z}$, we define the *energy dispersion below scale* 2^{-m} (or above frequency 2^m) of u of orders 0 and 1 to be, respectively,

$$\|u\|_{\mathrm{ED}_{\geq m}[I]} := \sup_{k \in \mathbb{Z}} 2^{-\delta_2 (m-k)_+} 2^{-2k} \|P_k u\|_{L^\infty L^\infty[I]}, \tag{4-16}$$

$$\|u\|_{\mathrm{ED}^1_{\geq m}[I]} := \sup_{k \in \mathbb{Z}} 2^{-\delta_2 (m-k)_+} 2^{-2k} \|\nabla P_k u\|_{L^\infty L^\infty[I]}. \tag{4-17}$$

The quantity $\|\cdot\|_{\mathrm{ED}_{\geq m}[I]}$ (resp. $\|\cdot\|_{\mathrm{ED}^1_{\geq m}[I]}$) is used at the level of the curvature F (resp. the connection A). As we work mostly at the level of the connection, unless stated otherwise, by energy dispersion we usually refer to the order-1 case.

Clearly, $\mathrm{ED}^1_{\geq m}[I]$ fails to be useful at frequencies below $O(2^m)$. In this regime, we exploit instead the *length* $|I|$ of the time interval as a source of smallness. Due to the scaling property of $\Box$, we must require $2^m |I|$ to be sufficiently small. To conveniently pack together the previous two concepts, we introduce the notion of an (ε, M)*-energy-dispersed function on an interval*.

Definition 4.9 ((ε, M)-energy-dispersed function on an interval). Let I be a time interval, and let $u \in S^1[I]$. We will say that the pair (u, I) is (ε, M)*-energy-dispersed* if there exists some $m \in \mathbb{Z}$ and $M > 0$ such that the following properties hold:

- (S^1-norm bound)

$$\|u\|_{S^1[I]} \leq M. \tag{4-18}$$

- (small energy dispersion)

$$\|u\|_{\mathrm{ED}^1_{\geq m}[I]} \leq \varepsilon M. \tag{4-19}$$

- (high-modulation bound)

$$\|\Box u\|_{L^2 \dot{H}^{-1/2}[I]} \leq \varepsilon M. \tag{4-20}$$

- (short time interval) $|I| \leq \varepsilon 2^{-m}$.

Observe (by interpolation) that if (u, I) is (ε, M)-energy-dispersed, then

$$\sup_k \|P_k u\|_{\mathrm{Str}^1[I]} \leq C \varepsilon^{\delta_1} M. \tag{4-21}$$

Finally, we state a proposition showing how the norms $DS^1[I]$ and $\mathrm{ED}^1_{\geq m}[I]$ behave under the extension procedure described above. Given an interval I, we denote by χ_I^k a generalized cutoff function adapted to the scale 2^{-k}:

$$\chi_I^k(t) = (1 + 2^k \operatorname{dist}(t, I))^{-N}, \tag{4-22}$$

where N is a sufficiently large number. Let us recall [Oh and Tataru 2018, Proposition 3.4]:[5]

Proposition 4.10. *Let $k \in \mathbb{Z}$, $\kappa \geq 0$ and I be a time interval such that $|I| \geq 2^{-k-\kappa}$. Consider a function u_I on I localized at frequency 2^k, and denote by u_I^{ext} its extension outside I as homogeneous waves. Then we have*

$$2^{-k} \|\chi_I^k \nabla u_I^{\mathrm{ext}}\|_{L^q L^r} \lesssim_N 2^{C\kappa} (\|u_I\|_{L^q L^r[I]} + 2^{(\frac{1}{2} - \frac{1}{q} - \frac{4}{r})} \|\Box u_I\|_{L^2 L^2[I]}), \tag{4-23}$$

$$2^{-2k} \|\chi_I^k \nabla u_I^{\mathrm{ext}}\|_{L^\infty L^\infty} \lesssim_N 2^{-2k} \|\nabla u_I\|_{L^\infty L^\infty[I]}, \tag{4-24}$$

where (q, r) is any pair of admissible Strichartz exponents on $\mathbb{R}^{1+4}$.

Remark 4.11. Since $2^{-k}[\chi_I^k, \nabla] = 2^{-k}(\nabla \chi_I^k)$ is simply multiplication by another generalized cutoff function adapted to the frequency scale 2^k, the conclusions of Proposition 4.10 also hold with $\chi_I^k 2^{-k} \nabla u_I^{\mathrm{ext}}$ replaced by $2^{-k} \nabla (\chi_I^k u_I^{\mathrm{ext}})$ on the left-hand sides.

4B. ***Estimates for quadratic nonlinearities.*** Here we state estimates for the quadratic nonlinearities in Theorems 3.5 and 3.6. All estimates stated here are proved in Section 8C.

Throughout this and the next subsections, we will denote by A a $\mathfrak{g}$-valued *spatial* 1-form $A = A_j \, dx^j$ on $I \times \mathbb{R}^4$ for some time interval I. To denote a $\mathfrak{g}$-valued *space-time* 1-form, we use the notation $A_{t,x} = A_\alpha \, dx^\alpha$. We will use B (resp. $B_{t,x}$) to denote[6] another $\mathfrak{g}$-valued spatial (resp. space-time) 1-form on $I \times \mathbb{R}^4$. Unless otherwise stated, all frequency envelopes will be assumed to be δ_3-admissible.

[5] To be pedantic, [Oh and Tataru 2018, Proposition 3.4] only corresponds to the case $\kappa = 0$. However, the required modification of the proof is straightforward.

[6] Note that this convention is different from [Oh and Tataru 2017b] and Section 3, where B was reserved for caloric gauge-linearized Yang–Mills heat flows.

We begin with the quadratic nonlinearities in the equations for A_0, $\partial_t A_0$ and $\partial^\ell A_\ell$. We introduce the notation

$$\mathcal{M}_0^2(A, B) = [A_\ell, \partial_t B^\ell], \tag{4-25}$$

$$\mathcal{D}\mathcal{M}_0^2(A, B) = -2\boldsymbol{Q}(\partial_t A, \partial_t B). \tag{4-26}$$

These are the main quadratic nonlinearities in the ΔA_0 and $\Delta \partial_t A_0$ equations, respectively. The estimates that we need for these nonlinearities are as follows.

Proposition 4.12. *We have the fixed-time bounds*

$$\| |D|^{-1} \mathcal{M}_0^2(A, B)(t) \|_{L^2_{cd}} \lesssim \|A(t)\|_{\dot{H}^1_c} \|\partial_t B(t)\|_{L^2_d}, \tag{4-27}$$

$$\| |D|^{-2} \mathcal{D}\mathcal{M}_0^2(A, B)(t) \|_{L^2_{cd}} \lesssim \|\partial_t A(t)\|_{L^2_c} \|\partial_t B(t)\|_{L^2_d}, \tag{4-28}$$

and the space-time bounds

$$\| |D|^{-1} \mathcal{M}_0^2(A, B) \|_{Y_{cd}[I]} \lesssim \|A\|_{S^1_c[I]} \|B\|_{S^1_d[I]}, \tag{4-29}$$

$$\| |D|^{-1} \mathcal{M}_0^2(A, B) \|_{L^2 \dot{H}^{1/2}_{cd}[I]} + \| |D|^{-2} \mathcal{D}\mathcal{M}_0^2(A, B) \|_{L^2 \dot{H}^{1/2}_{cd}[I]} \lesssim \|A\|_{\mathrm{Str}^1_c[I]} \|B\|_{\mathrm{Str}^1_d[I]}. \tag{4-30}$$

Moreover, for any $\kappa > 0$, *the nonlinearity* $\mathcal{M}_0^2(A, B)$ *admits the splitting*

$$\mathcal{M}_0^2(A, B) = \mathcal{M}^{\kappa,2}_{0,\mathrm{small}}(A, B) + \mathcal{M}^{\kappa,2}_{0,\mathrm{large}}(A, B),$$

where the small part obeys the improved bound

$$\| |D|^{-1} \mathcal{M}^{\kappa,2}_{0,\mathrm{small}}(A, B) \|_{Y_{cd}[I]} \lesssim 2^{-\delta_2 \kappa} \|A\|_{S^1_c[I]} \|B\|_{S^1_d[I]}, \tag{4-31}$$

and the large part is bounded by divisible norms of A *and* B:

$$\| |D|^{-1} \mathcal{M}^{\kappa,2}_{0,\mathrm{large}}(A, B) \|_{Y_{cd}[I]} \lesssim 2^{C\kappa} \|A\|_{DS^1_c[I]} \|B\|_{DS^1_d[I]}. \tag{4-32}$$

Finally, if either

$$\|A\|_{S^1_c[I]} \leq 1 \quad \textit{and} \quad (B, I) \textit{ is } (\varepsilon, M)\textit{-energy-dispersed, or}$$

$$\|B\|_{S^1_c[I]} \leq 1 \quad \textit{and} \quad (A, I) \textit{ is } (\varepsilon, M)\textit{-energy-dispersed,}$$

then we have

$$\| |D|^{-1} \mathcal{M}_0^2(A, B) \|_{Y_c[I]} \lesssim \varepsilon^{\delta_2} M, \tag{4-33}$$

$$\| |D|^{-2} \mathcal{D}\mathcal{M}_0^2(A, B) \|_{L^2 \dot{H}^{1/2}_c[I]} \lesssim \varepsilon^{\delta_2} M. \tag{4-34}$$

The remaining quadratic nonlinearities in the equations for A_0 and $\partial^\ell A_\ell$ involve $\boldsymbol{Q}$, and they obey simpler estimates.

Proposition 4.13. *For* $\sigma = 0$ *or* 1, *we have the fixed-time bound*

$$\| |D|^{-\sigma} \boldsymbol{Q}(A, \partial_t^\sigma B)(t) \|_{L^2_{cd}} \lesssim \|A(t)\|_{\dot{H}^1_c} \|\partial_t^\sigma B(t)\|_{\dot{H}^{1-\sigma}_d} \tag{4-35}$$

and the space-time bounds

$$\| |D|^{-\sigma} \boldsymbol{Q}(A, \partial_t^{\sigma} B)\|_{L^2 \dot{H}^{1/2}_{cd}[I]} \lesssim \|A\|_{\mathrm{Str}^1_c[I]} \|B\|_{\mathrm{Str}^1_d[I]}, \tag{4-36}$$

$$\| |D|^{-\sigma} \boldsymbol{Q}(A, \partial_t^{\sigma} B)\|_{Y_{cd}[I]} + \| |D|^{-\sigma-1} \boldsymbol{Q}(A, \partial_t^{\sigma} B)\|_{L^1 L^{\infty}_{cd}[I]} \lesssim \|A\|_{DS^1_c[I]} \|B\|_{DS^1_d[I]}. \tag{4-37}$$

Finally, if either

$$\|A\|_{S^1_c[I]} \le 1 \quad \textit{and} \quad (B, I) \textit{ is } (\varepsilon, M)\textit{-energy-dispersed, or}$$
$$\|B\|_{S^1_c[I]} \le 1 \quad \textit{and} \quad (A, I) \textit{ is } (\varepsilon, M)\textit{-energy-dispersed,}$$

then

$$\| |D|^{-\sigma} \boldsymbol{Q}(A, \partial_t^{\sigma} B)\|_{Y_c[I]} \lesssim \varepsilon^{\delta_2} M. \tag{4-38}$$

Also for the quadratic part $\boldsymbol{A}_0^2$ of A_0, given by

$$\boldsymbol{A}_0^2(A, A) = \Delta^{-1}([A, \partial_t A]) + 2Q(A, \partial_t A),$$

we have the following additional property, which will be used in the proof of Theorem 1.18:

Proposition 4.14. *For the quadratic form $\boldsymbol{A}_0^2$ we have*

$$\| |D|^2 \boldsymbol{A}_0^2(A, B)\|_{(L^2_x L^1_t)_{cd}[I]} \lesssim \|\nabla A\|_{S^{sq}_c} \|\nabla B\|_{S^{sq}_d}. \tag{4-39}$$

For the quadratic nonlinearity in the $\Box_A A_j$ equation, we introduce the notation

$$\boldsymbol{P}_j \mathcal{M}^2(A, B) = \boldsymbol{P}_j [A_\ell, \partial_x B^\ell],$$
$$\boldsymbol{P}_j^{\perp} \mathcal{M}^2(A, B) = 2\Delta^{-1} \partial_j \boldsymbol{Q}(\partial^\alpha A, \partial_\alpha A),$$

so that (3-12) becomes

$$\Box_A A_j = \boldsymbol{P}_j \mathcal{M}(A, A) + \boldsymbol{P}_j^{\perp} \mathcal{M}(A, A) + R_j(A, \partial_t A).$$

Proposition 4.15. *We have the fixed-time bounds*

$$\|\boldsymbol{P} \mathcal{M}^2(A, B)(t)\|_{\dot{H}^{-1}_{cd}} \lesssim \|A(t)\|_{\dot{H}^1_c} \|B(t)\|_{\dot{H}^1_d}, \tag{4-40}$$

$$\|\boldsymbol{P}^{\perp} \mathcal{M}^2(A, B)(t)\|_{\dot{H}^{-1}_{cd}} \lesssim \|\nabla A(t)\|_{L^2_c} \|\nabla B(t)\|_{L^2_d}, \tag{4-41}$$

and space-time bounds

$$\|\boldsymbol{P} \mathcal{M}^2(A, B)\|_{(N \cap \Box \underline{X}^1)_{cd}[I]} \lesssim \|A\|_{S^1_c[I]} \|B\|_{S^1_d[I]}, \tag{4-42}$$

$$\|\boldsymbol{P}^{\perp} \mathcal{M}^2(A, B)\|_{(N \cap \Box \underline{X}^1)_{cd}[I]} \lesssim \|A\|_{S^1_c[I]} \|B\|_{S^1_d[I]}. \tag{4-43}$$

In particular, the $L^2 \dot{H}^{-1/2}$-norms are bounded by the Str^1*-norms of A and B:*

$$\|\boldsymbol{P} \mathcal{M}^2(A, B)\|_{L^2 \dot{H}^{-1/2}_{cd}[I]} \lesssim \|A\|_{\mathrm{Str}^1_c[I]} \|B\|_{\mathrm{Str}^1_d[I]}, \tag{4-44}$$

$$\|\boldsymbol{P}^{\perp} \mathcal{M}^2(A, B)\|_{L^2 \dot{H}^{-1/2}_{cd}[I]} \lesssim \|A\|_{\mathrm{Str}^1_c[I]} \|B\|_{\mathrm{Str}^1_d[I]}. \tag{4-45}$$

Moreover, for any $\kappa > 0$, the terms $\boldsymbol{P}_j \mathcal{M}^2(A, B)$ and $\boldsymbol{P}_j^{\perp} \mathcal{M}^2(A, B)$ admit the splittings

$$\boldsymbol{P}_j \mathcal{M}^2(A, B) = \boldsymbol{P}_j \mathcal{M}^{\kappa,2}_{\text{small}}(A, B) + \boldsymbol{P}_j \mathcal{M}^{\kappa,2}_{\text{large}}(A, B),$$
$$\boldsymbol{P}_j^{\perp} \mathcal{M}^2(A, B) = \boldsymbol{P}_j^{\perp} \mathcal{M}^{\kappa,2}_{\text{small}}(A, B) + \boldsymbol{P}_j^{\perp} \mathcal{M}^{\kappa,2}_{\text{large}}(A, B),$$

so that the N-norms of the small parts obey the improved bounds

$$\|\boldsymbol{P} \mathcal{M}^{\kappa,2}_{\text{small}}(A, B)\|_{N_{cd}[I]} \lesssim 2^{-\delta_2 \kappa} \|A\|_{S^1_c[I]} \|B\|_{S^1_d[I]}, \tag{4-46}$$
$$\|\boldsymbol{P}^{\perp} \mathcal{M}^{\kappa,2}_{\text{small}}(A, B)\|_{N_{cd}[I]} \lesssim 2^{-\delta_2 \kappa} \|A\|_{S^1_c[I]} \|B\|_{S^1_d[I]}, \tag{4-47}$$

and those of the large parts are bounded by divisible norms of A and B:

$$\|\boldsymbol{P} \mathcal{M}^{\kappa,2}_{\text{large}}(A, B)\|_{N_{cd}[I]} \lesssim 2^{C\kappa} \|A\|_{DS^1_c[I]} \|B\|_{DS^1_d[I]}, \tag{4-48}$$
$$\|\boldsymbol{P}^{\perp} \mathcal{M}^{\kappa,2}_{\text{large}}(A, B)\|_{N_{cd}[I]} \lesssim 2^{C\kappa} \|A\|_{DS^1_c[I]} \|B\|_{DS^1_d[I]}. \tag{4-49}$$

Finally, if either

$$\|A\|_{\underline{S}^1_c[I]} \le 1 \quad \textit{and} \quad (B, I) \textit{ is } (\varepsilon, M)\textit{-energy-dispersed, or}$$
$$\|B\|_{\underline{S}^1_c[I]} \le 1 \quad \textit{and} \quad (A, I) \textit{ is } (\varepsilon, M)\textit{-energy-dispersed,}$$

then

$$\|\boldsymbol{P} \mathcal{M}^2(A, B)\|_{(N \cap L^2 \dot{H}^{-1/2})_c[I]} \lesssim \varepsilon^{\delta_2} M, \tag{4-50}$$
$$\|\boldsymbol{P}^{\perp} \mathcal{M}^2(A, B)\|_{(N \cap L^2 \dot{H}^{-1/2})_c[I]} \lesssim \varepsilon^{\delta_2} M. \tag{4-51}$$

We end this subsection with bilinear estimates for $\boldsymbol{w}_0^2$ and $\boldsymbol{w}_x^2$, which arise in the equation for a dynamic Yang–Mills heat flow of a caloric Yang–Mills wave.

Proposition 4.16. *For any $s > 0$, we have the fixed-time bound*

$$\||D|^{-1} P_k \boldsymbol{w}_0^2(A, B, s)(t)\|_{L^2} \lesssim \langle 2^{2k} s\rangle^{-10} \langle 2^{-2k} s^{-1}\rangle^{-\delta_2} c_k d_k \|\partial_t A(t)\|_{L^2_c} \|B(t)\|_{\dot{H}^1_d} \tag{4-52}$$

and the space-time bounds

$$\||D|^{-1} P_k \boldsymbol{w}_0^2(A, B, s)\|_{L^2 \dot{H}^{1/2}[I]} \lesssim \langle 2^{2k} s\rangle^{-10} \langle 2^{-2k} s^{-1}\rangle^{-\delta_2} c_k d_k \|A\|_{\mathrm{Str}^1_c[I]} \|B\|_{\mathrm{Str}^1_d[I]}, \tag{4-53}$$
$$\||D|^{-1} P_k \boldsymbol{w}_0^2(A, B, s)\|_{Y[I]} \lesssim \langle 2^{2k} s\rangle^{-10} \langle 2^{-2k} s^{-1}\rangle^{-\delta_2} c_k d_k \|A\|_{S^1_c[I]} \|B\|_{S^1_d[I]}. \tag{4-54}$$

Moreover, if (B, I) is (ε, M)-energy-dispersed, then

$$\||D|^{-1} P_k \boldsymbol{w}_0^2(A, B, s)\|_{Y[I]} \lesssim \varepsilon^{\delta_2} \langle 2^{2k} s\rangle^{-10} \langle 2^{-2k} s^{-1}\rangle^{-\delta_2} c_k \|A\|_{S^1_c[I]} M. \tag{4-55}$$

Proposition 4.17. *For any $s > 0$, we have the fixed-time bound*

$$\|P_k \boldsymbol{P} \boldsymbol{w}_x^2(A, B, s)(t)\|_{\dot{H}^{-1}} \lesssim \langle 2^{2k} s\rangle^{-10} \langle 2^{-2k} s^{-1}\rangle^{-\delta_2} c_k d_k \|\nabla A(t)\|_{L^2_c} \|\nabla B(t)\|_{L^2_d} \tag{4-56}$$

and the space-time bounds

$$\begin{aligned}&\|P_k \boldsymbol{P}\boldsymbol{w}_x^2(A,B,s)\|_{L^2\dot{H}^{-1/2}[I]}\\&\qquad\lesssim \langle 2^{2k}s\rangle^{-10}\langle 2^{-2k}s^{-1}\rangle^{-\delta_2}c_k d_k \|(\nabla A, \nabla \boldsymbol{P}^{\perp}A)\|_{(\mathrm{Str}^0\times L^2\dot{H}^{1/2})_c[I]}\|B\|_{\mathrm{Str}^1_d[I]},\end{aligned}\tag{4-57}$$

$$\begin{aligned}&\|P_k \boldsymbol{P}\boldsymbol{w}_x^2(A,B,s)\|_{N\cap\Box\underline{X}^1[I]}\\&\qquad\lesssim \langle 2^{2k}s\rangle^{-10}\langle 2^{-2k}s^{-1}\rangle^{-\delta_2}c_k d_k \|(A, \boldsymbol{P}^{\perp}A)\|_{(S^1\times Y^1)_c[I]}\|B\|_{S^1_d[I]}.\end{aligned}\tag{4-58}$$

Moreover, if (B,I) is (ε,M)-energy-dispersed, then

$$\begin{aligned}&\|P_k \boldsymbol{P}\boldsymbol{w}_x^2(A,B,s)\|_{N\cap L^2\dot{H}^{-1/2}[I]}\\&\qquad\lesssim \langle 2^{2k}s\rangle^{-10}\langle 2^{-2k}s^{-1}\rangle^{-\delta_2}c_k(\varepsilon^{\delta_2}\|A\|_{\underline{S}^1_c[I]}+\|\nabla\boldsymbol{P}^{\perp}A\|_{L^2\dot{H}^{1/2}_c[I]})M.\end{aligned}\tag{4-59}$$

4C. *Estimates for the covariant wave operator.* We now state estimates concerning the covariant wave operator $\Box_A$. All estimates stated here without proofs are proved in Section 8C, with the exceptions of Theorem 4.24 and Proposition 4.25, which are proved in Section 9.

We begin by expanding $\Box_A B$ to

$$\Box_A B = \Box B + 2[A_\alpha, \partial^\alpha B] + [\partial^\alpha A_\alpha, B] + [A^\alpha, [A_\alpha, B]].$$

We have the following simple fixed-time estimates for $\Box_A - \Box$.

Proposition 4.18. *For any $\alpha,\beta,\gamma\in\{0,1,\dots,4\}$, we have the fixed-time bounds*

$$\|[A_\alpha,\partial^\alpha B](t)\|_{\dot{H}^{-1}_{cd}} \lesssim \|(A_0,A)(t)\|_{\dot{H}^1_c}\|\nabla B(t)\|_{L^2_d},\tag{4-60}$$

$$\|[\partial^\alpha A_\alpha, B](t)\|_{\dot{H}^{-1}_{cd}} \lesssim (\|A(t)\|_{\dot{H}^1_c}+\|\partial_t A_0(t)\|_{L^2_c})\|B(t)\|_{\dot{H}^1_d},\tag{4-61}$$

$$\|[A^{(1)}_\alpha,[A^{(2)\alpha},B]](t)\|_{\dot{H}^{-1}_{cde}} \lesssim \|(A^{(1)}_0,A^{(1)})(t)\|_{\dot{H}^1_c}\|(A^{(2)}_0,A^{(2)})(t)\|_{\dot{H}^1_d}\|B(t)\|_{\dot{H}^1_e}\tag{4-62}$$

and the space-time bounds

$$\|[A_\ell,\partial^\ell B]\|_{L^2\dot{H}^{-1/2}_{cd}[I]} \lesssim \|A\|_{\mathrm{Str}^1_c[I]}\|B\|_{\mathrm{Str}^1_d[I]},\tag{4-63}$$

$$\|[A_0,\partial_0 B]\|_{L^2\dot{H}^{-1/2}_{cd}[I]} \lesssim \|\nabla A_0\|_{L^2\dot{H}^{1/2}_c[I]}\|B\|_{\mathrm{Str}^1_d[I]},\tag{4-64}$$

$$\|[\partial^\alpha A_\alpha,B]\|_{L^2\dot{H}^{-1/2}_{cd}[I]} \lesssim \|(\nabla A_0,\nabla\boldsymbol{P}^{\perp}A)\|_{L^2\dot{H}^{1/2}_c[I]}\|B\|_{\mathrm{Str}^1_d[I]},\tag{4-65}$$

$$\begin{aligned}\|[A^{(1)}_\alpha,[A^{(2)\alpha},B]](t)\|_{L^2\dot{H}^{-1/2}_{cde}[I]} &\lesssim \|(\nabla A^{(1)}_0,\nabla A^{(1)})(t)\|_{L^2\dot{H}^{1/2}\times\mathrm{Str}^0_c[I]}\\&\qquad\times\|(\nabla A^{(2)}_0,\nabla A^{(2)})(t)\|_{L^2\dot{H}^{1/2}\times\mathrm{Str}^0_c[I]}\|B\|_{\mathrm{Str}^1_e[I]}.\end{aligned}\tag{4-66}$$

In order to proceed, we recall the notation $\boldsymbol{P}_\alpha A = (\boldsymbol{P}A)_\alpha$ for a space-time 1-form $A_{t,x}$:

$$\boldsymbol{P}_\alpha A = \begin{cases}\boldsymbol{P}_j A_x, & \alpha = j\in\{1,\dots,4\},\\ A_0, & \alpha=0.\end{cases}$$

We also write $\boldsymbol{P}^{\perp}_\alpha A = (\boldsymbol{P}^{\perp}A)_\alpha = A_\alpha - \boldsymbol{P}_\alpha A$.

Given a parameter $\kappa \in \mathbb{N}$, we furthermore decompose $2[A_\alpha, \partial^\alpha B]$ so that

$$\begin{aligned}\Box_A B &= \Box B + 2[A_\alpha, \partial^\alpha B] + \mathrm{Rem}^3_A B \\ &= \Box B + \mathrm{Diff}^\kappa_{\boldsymbol{P}A} B + \mathrm{Diff}^\kappa_{\boldsymbol{P}^\perp A} B + \mathrm{Rem}^{\kappa,2}_A B + \mathrm{Rem}^3_A B,\end{aligned} \tag{4-67}$$

where[7]

$$\mathrm{Diff}^\kappa_{\boldsymbol{P}A} = \sum_k 2[P_{<k-\kappa} \boldsymbol{P}_\alpha A, \partial^\alpha P_k B], \tag{4-68}$$

$$\mathrm{Diff}^\kappa_{\boldsymbol{P}^\perp A} = \sum_k 2[P_{<k-\kappa} \boldsymbol{P}^\perp_\alpha A, \partial^\alpha P_k B], \tag{4-69}$$

$$\mathrm{Rem}^{\kappa,2}_A = \sum_k 2[P_{\geq k-\kappa} A_\alpha, \partial^\alpha P_k B], \tag{4-70}$$

$$\mathrm{Rem}^3_A B = [\partial^\alpha A_\alpha, B] + [A^\alpha, [A_\alpha, B]]. \tag{4-71}$$

We now turn to the bounds for each part of the decomposition (4-67). For a fixed $B \in S^1[I]$, we introduce the nonlinear maps

$$\mathrm{Rem}^3(A)B = -[\boldsymbol{DA}_0(A), B] + [\boldsymbol{DA}(A), B] - [\boldsymbol{A}_0(A), [\boldsymbol{A}_0(A), B]] + [A^\ell, [A_\ell, B]], \tag{4-72}$$

$$\mathrm{Rem}^3_s(A)B = -[\boldsymbol{DA}_{0;s}(A), B] - [\boldsymbol{A}_{0;s}(A), [\boldsymbol{A}_{0;s}(A), B]], \tag{4-73}$$

defined for spatial connections A on I such that $(A, \partial_t A)(t) \in T^{L^2}\mathcal{C}$ for each fixed time $t \in I$. In view of Theorems 3.5 and 3.6, for a caloric Yang–Mills wave A we have

$$\begin{aligned}\mathrm{Rem}^3_A B &= \mathrm{Rem}^3(A)B, \\ \mathrm{Rem}^3_{A(s)} B &= \mathrm{Rem}^3(A(s))B + \mathrm{Rem}^3_s(A)B.\end{aligned}$$

The nonlinear maps $\mathrm{Rem}^3(A)B$ and $\mathrm{Rem}^3_s(A)B$ are well-behaved:

Proposition 4.19. *Suppose that $A(t) \in \mathcal{C}_\mathcal{Q}$ for every $t \in I$. Then the following properties hold with bounds depending on $\mathcal{Q}$, but otherwise independent of I:*

- *Let c and d be $(-\delta_2, S)$-frequency envelopes for A and B in $\mathrm{Str}^1[I]$, respectively. Then*

$$\| P_k (\mathrm{Rem}^3(A)B) \|_{L^1 L^2 \cap L^2 \dot{H}^{-1/2}[I]} \lesssim_{\mathcal{Q}, \|A\|_{\mathrm{Str}^1[I]}} (c^{[\delta_2]}_k)^2 d_k + c_k c^{[\delta_2]}_k d^{[\delta_2]}_k. \tag{4-74}$$

- *For a fixed $A \in \mathrm{Str}^1[I]$, $\mathrm{Rem}^3(A)B$ is linear in B. On the other hand, for a fixed B with $\|B\|_{\mathrm{Str}^1[I]} \leq 1$, $\mathrm{Rem}^3(\cdot)B : \mathrm{Str}^1[I] \to L^1 L^2 \cap L^2 \dot{H}^{-1/2}[I]$ is Lipschitz envelope-preserving.*
- *For a fixed $A \in \mathrm{Str}^1[I]$, $\mathrm{Rem}^3_s(A)B$ is linear in B. On the other hand, for a fixed $B \in S^1[I]$ with $\|B\|_{\mathrm{Str}^1[I]} \leq 1$, $\mathrm{Rem}^3_s(A)B$ is a Lipschitz map*

$$\mathrm{Rem}^3_s(A)B : \mathrm{Str}^1[I] \to L^1 L^2 \cap L^2 \dot{H}^{-\frac{1}{2}}[I] \tag{4-75}$$

[7] Although the definition depends on the whole space-time connection $A_{t,x}$, we deviate from our convention and simply write $\mathrm{Diff}^\kappa_{\boldsymbol{P}A}$, $\mathrm{Diff}^\kappa_{\boldsymbol{P}^\perp A}$, $\mathrm{Rem}^{\kappa,2}_A$ etc. to avoid cluttered notation.

with output concentrated at frequency $s^{-1/2}$,

$$(1-s\Delta)^N \mathrm{Rem}^3_s(A)B : \mathrm{Str}^1[I] \to 2^{-\delta_2 k(s)} L^1 \dot{H}^{-\delta_2} \cap L^2 \dot{H}^{-\frac{1}{2}-\delta_2}[I]. \tag{4-76}$$

Next, we consider the term

$$2[A_\alpha, \partial^\alpha B] = \mathrm{Diff}^\kappa_{\boldsymbol{P}A} B + \mathrm{Diff}^\kappa_{\boldsymbol{P}^\perp A} B + \mathrm{Rem}^{\kappa,2}_A B.$$

We begin with $\mathrm{Rem}^{\kappa,2}_A B$, which obeys analogous bounds as $\boldsymbol{P}\mathcal{M}^2(A,B)$ and $\boldsymbol{P}^\perp\mathcal{M}^2(A,B)$ (see Proposition 4.15).

Proposition 4.20. *For any* $\kappa > 0$, *the term* $\mathrm{Rem}^{\kappa,2}_A B$ *obeys the bound*

$$\|\mathrm{Rem}^{\kappa,2}_A B\|_{(N\cap\Box\underline{X}^1)_{cd}[I]} \lesssim 2^{C\kappa}(\|A\|_{S^1_c[I]} + \|(\boldsymbol{P}^\perp A, A_0)\|_{Y^1_c[I]})\|B\|_{S^1_d[I]}. \tag{4-77}$$

In particular, its $L^2\dot{H}^{-1/2}$*-norm is bounded by*

$$\|\mathrm{Rem}^{\kappa,2}_A B\|_{L^2\dot{H}^{-1/2}_{cd}[I]} \lesssim (\|A\|_{\mathrm{Str}^1_c[I]} + \|(\nabla\boldsymbol{P}^\perp A, \nabla A_0)\|_{(L^2\dot{H}^{1/2})_c[I]})\|B\|_{\mathrm{Str}^1_d[I]}. \tag{4-78}$$

Furthermore, $\mathrm{Rem}^{\kappa,2}_A B$ *admits the splitting*

$$\mathrm{Rem}^{\kappa,2}_A B = \mathrm{Rem}^{\kappa,2}_{A,\mathrm{small}} B + \mathrm{Rem}^{\kappa,2}_{A,\mathrm{large}} B$$

so that the N*-norm of the small part obeys the improved bound*

$$\|\mathrm{Rem}^{\kappa,2}_{A,\mathrm{small}} B\|_{N_{cd}[I]} \lesssim 2^{-\delta_2\kappa}\|A\|_{S^1_c[I]}\|B\|_{S^1_d[I]}, \tag{4-79}$$

and that of the large part is bounded by a divisible norm of (A_0, A):

$$\|\mathrm{Rem}^{\kappa,2}_{A,\mathrm{large}} B\|_{N_{cd}[I]} \lesssim 2^{C\kappa}(\|A\|_{DS^1_c[I]} + \|(\nabla\boldsymbol{P}^\perp A, \nabla A_0)\|_{(L^2\dot{H}^{1/2})_c[I]})\|B\|_{S^1_d[I]}. \tag{4-80}$$

Finally, if (B, I) *is* (ε, M)*-energy-dispersed, then*

$$\|\mathrm{Rem}^{\kappa,2}_A B\|_{(N\cap L^2\dot{H}^{-1/2})_c[I]} \lesssim (2^{-\delta_2\kappa} + 2^{C\kappa}\varepsilon^{\delta_2})\|A\|_{\underline{S}^1_c[I]}M + 2^{C\kappa}\|(\nabla\boldsymbol{P}^\perp A, \nabla A_0)\|_{(L^2\dot{H}^{1/2})_c[I]}M. \tag{4-81}$$

It remains to consider the paradifferential terms. The term $\mathrm{Diff}^\kappa_{\boldsymbol{P}^\perp A} B$ can be handled using the following estimate, in combination with (3-22) and Proposition 4.12:

Proposition 4.21. *For any* $\kappa > 0$, *we have*

$$\|\mathrm{Diff}^\kappa_{\boldsymbol{P}^\perp A} B\|_{(X^{-1/2+b_1,-b_1}\cap\Box\underline{X}^1)_{cd}[I]} \lesssim \|\boldsymbol{P}^\perp A\|_{Y^1_c[I]}\|B\|_{S^1_d[I]}. \tag{4-82}$$

Moreover, we have

$$\|\mathrm{Diff}^\kappa_{\boldsymbol{P}^\perp A} B\|_{L^1 L^2_f[I]} \lesssim \|\boldsymbol{P}^\perp A\|_{L^1 L^\infty_a[I]}\|B\|_{S^1_e[I]}, \tag{4-83}$$

where

$$f_k = \Big(\sum_{k'<k-\kappa} a_{k'}\Big) e_k.$$

The only remaining term is the paradifferential term $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A} B$. We first state the high-modulation bounds.

Proposition 4.22. *For any $\kappa > 0$, consider the splitting $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A} = \mathrm{Diff}^{\kappa}_{A_0} + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}_x A}$, where*

$$\mathrm{Diff}^{\kappa}_{A_0} B = -\sum_k 2[P_{<k-\kappa} A_0, \partial_t P_k B], \quad \mathrm{Diff}^{\kappa}_{\boldsymbol{P}_x A} B = \sum_k 2[P_{<k-\kappa} \boldsymbol{P}_\ell A, \partial^\ell P_k B].$$

For $\mathrm{Diff}_{A_0} B$, we have the bound

$$\|\mathrm{Diff}^{\kappa}_{A_0} B\|_{(X^{-1/2+b_1,-b_1} \cap \Box \underline{X}^1)_{cd}[I]} \lesssim \|A_0\|_{Y^1_c[I]} \|B\|_{S^1_d[I]}. \tag{4-84}$$

On the other hand, for $\mathrm{Diff}_{\boldsymbol{P}_x A} B$, we have the bounds

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}_x A} B\|_{(\Box \tilde{X}^1)_{cd}[I]} \lesssim \|A_x\|_{S^1_c[I]} \|B\|_{S^1_d[I]}, \tag{4-85}$$

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}_x A} B\|_{(\Box \underline{X}^1)_{cd}[I]} \lesssim \|A_x\|_{(S^1 \cap \tilde{X}^1)_c[I]} \|B\|_{S^1_d[I]}, \tag{4-86}$$

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}_x A} B\|_{(X^{-1/2+b_1,-b_1})_{cd}[I]} \lesssim \|A_x\|_{(S^1 \cap \underline{X}^1)_c[I]} \|B\|_{S^1_d[I]}. \tag{4-87}$$

Next, we consider the $N \cap L^2 \dot{H}^{1/2}$ norm of $\mathrm{Diff}_{\boldsymbol{P}A} B$. The contribution of each Littlewood–Paley projection $P_{k_0} \boldsymbol{P}A$ is perturbative, as the following proposition states:

Proposition 4.23. *Let $A_{t,x}$ be a caloric Yang–Mills wave on an interval I obeying*

$$\|A\|_{S^1[I]} \leq M. \tag{4-88}$$

Then for any $\kappa > 0$ and $k_0 \in \mathbb{Z}$, we have

$$\|\mathrm{Diff}^{\kappa}_{P_{k_0} \boldsymbol{P}A} B\|_{(N \cap L^2 \dot{H}^{-1/2})_d[I]} \lesssim_M \|B\|_{S^1_d[I]}. \tag{4-89}$$

However, we *cannot* sum up in k_0. The proper way to handle $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}$ is not to regard it as a perturbative nonlinearity, but rather as a part of the underlying linear operator. Indeed, for the operator $\Box + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}$, we have the following well-posedness result:

Theorem 4.24. *Let $A_{t,x}$ be a caloric Yang–Mills wave on an interval I obeying (4-88). Consider the following initial value problem on $I \times \mathbb{R}^4$:*

$$\begin{cases} \Box B + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A} B = G, \\ (B, \partial_t B)(t_0) = (B_0, B_1), \end{cases} \tag{4-90}$$

for some $\mathfrak{g}$-valued spatial 1-form $G \in N \cap L^2 \dot{H}^{-1/2}[I]$, $(B_0, B_1) \in \dot{H}^1 \times L^2$ and $t_0 \in I$.

Then for $\kappa \geq \kappa_1(M)$, where $\kappa_1(M) \gg 1$ is some function independent of $A_{t,x}$, there exists a unique solution $B \in S^1[I]$ to (4-90). Moreover, for any admissible frequency envelope c, the solution obeys the bound

$$\|B\|_{S^1_c[I]} \lesssim_M \|(B_0, B_1)\|_{(\dot{H}^1 \times L^2)_c} + \|G\|_{(N \cap L^2 \dot{H}^{-1/2})_c[I]}. \tag{4-91}$$

As a quick corollary of Propositions 4.19–4.20 and Theorem 4.24, we obtain well-posedness of the initial value problem associated to $\Box_A$; see Theorem 5.1(1) below.

Theorem 4.24 is proved in Sections 9, 10 and 11. The main ingredient for the proof is construction of a parametrix for $\Box + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}$ by renormalization with a pseudodifferential gauge transformation; for a more detailed discussion, see Section 9.

The paradifferential wave equation (4-90) leads to the following *weak divisibility* property of the S^1 norm, which will later play an important role in the energy induction argument.

Proposition 4.25. *Let $A_{t,x}$ be a caloric Yang–Mills wave on an interval I which obeys (4-88) for some $M > 0$. Let $B \in S^1[I]$ be a solution to the paradifferential wave equation (4-90) with the source $G \in N \cap L^2 \dot{H}^{-1/2}[I]$, which obeys the bound*

$$\sup_{t \in I} \|(B, \partial_t B)(t)\|_{L^2} \leq E \tag{4-92}$$

for some $E > 0$. Then there exists a partition $I = \bigcup_{i \in \mathcal{I}} I_i$ such that

$$\|B\|_{S^1[I_i]} \lesssim_E 1 \quad \text{for } i \in \mathcal{I}, \tag{4-93}$$

where

$$\#\mathcal{I} \lesssim_{E, M, \|B\|_{S^1[I]}, \|G\|_{N \cap L^2 \dot{H}^{-1/2}[I]}} 1.$$

The proof of this proposition also involves the parametrix construction (see Sections 9, 10 and 11), as well as Proposition 4.23.

We now state additional estimates satisfied by $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}$, which are needed to analyze the difference of two solutions (or even approximate solutions). For this purpose, it is necessary to exploit the so-called secondary null structure of the Yang–Mills equation, which becomes available after reiterating the equations for $\boldsymbol{P}A$.

We begin with simple bilinear estimates, which allow us to peel off the nonessential parts (in particular, the contribution of the cubic and higher-order nonlinearities) of A_0 and $\boldsymbol{P}A$.

Proposition 4.26. *We have*

$$\|\mathrm{Diff}^{\kappa}_{A_0} B\|_{(N \cap L^2 \dot{H}^{-1/2})_f[I]} \lesssim \|A_0\|_{(L^1 L^\infty \cap L^2 \dot{H}^{3/2})_a[I]} \|B\|_{S^1_e[I]}, \tag{4-94}$$

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}_x A} B\|_{(N \cap L^2 \dot{H}^{-1/2})_f[I]} \lesssim (\|\boldsymbol{P}A[t_0]\|_{(\dot{H}^1 \times L^2)_a} + \|\Box \boldsymbol{P}A\|_{L^1 L^2_a[I]}) \|B\|_{S^1_e[I]}, \tag{4-95}$$

where

$$f_k = \Big(\sum_{k' < k - \kappa} a_{k'} \Big) e_k.$$

The contribution of the quadratic nonlinearities $\mathcal{M}^2_0$ and $\mathcal{M}^2$ in the equations for A_0 and A_x, respectively, cannot be treated separately. This is precisely where we exploit the secondary null structure, which only manifests itself after combining the contribution of these nonlinearities in $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}$.

Proposition 4.27. *Let*

$$\Delta A_0 = [B^{(1)\ell}, \partial_t B^{(2)}_{\ell}], \tag{4-96}$$

$$\Box \boldsymbol{P}A = \boldsymbol{P}[B^{(1)\ell}, \partial_x B^{(2)}_{\ell}], \quad \boldsymbol{P}A[t_0] = 0, \tag{4-97}$$

where $B^{(1)}, B^{(2)} \in S^1[I]$. *Then we have*

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A} B\|_{(N \cap L^2 \dot{H}^{-1/2})_f[I]} \lesssim_{\tilde{M}} \|B^{(1)}\|_{S^1_c[I]} \|B^{(2)}\|_{S^1_d[I]} \|B\|_{S^1_e[I]}, \tag{4-98}$$

where

$$f_k = \Big(\sum_{k' < k - \kappa} c_{k'} d_{k'} \Big) e_k.$$

Next, we turn to the contribution of terms of the form $[A_\alpha, \partial^\alpha A]$ in the equation for $\boldsymbol{P}_x A$. The frequency envelope bound for this term is slightly involved, because it does not obey a good N-norm estimate.

Proposition 4.28. *Let* $A_0 = 0$ *and*

$$\Box \boldsymbol{P} A_j = \sum_{n=1}^{N} \boldsymbol{P}[B^{n(1)}_\alpha, \partial^\alpha B^{n(2)}_j], \quad \boldsymbol{P} A[t_0] = 0, \tag{4-99}$$

where

$$\|B^{n(1)}\|_{\underline{S}^1_{c^n}[I]} + \|(B^{n(1)}_0, \boldsymbol{P}^\perp B^{n(1)})\|_{Y^1_{c^n}[I]} \leq 1, \quad \|B^{n(2)}\|_{S^1_{d^n}[I]} \leq 1. \tag{4-100}$$

Assume furthermore that

$$\|\boldsymbol{P} A\|_{S^1_a[I]} \leq 1, \quad \|B\|_{S^1_e[I]} \leq 1. \tag{4-101}$$

Then we have

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}_x A} B\|_{(N \cap L^2 \dot{H}^{-1/2})_f[I]} \lesssim 1, \tag{4-102}$$

where

$$f_k = \Big(\sum_{k' < k - \kappa} (a_{k'} + \sum_{n=1}^{N} c^n_{k'} d^n_{k'}) \Big) e_k.$$

Next, we state a trilinear estimate for $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}$ in the presence of $\boldsymbol{w}^2_\mu$ which is analogous to Proposition 4.27. This is needed for analyzing the dynamic Yang–Mills heat flow of a caloric Yang–Mills wave.

Proposition 4.29. *Let*

$$\Delta A_0 = \boldsymbol{w}^2_0(B^{(1)}, B^{(2)}, s), \tag{4-103}$$

$$\Box \boldsymbol{P} A = \boldsymbol{P} \boldsymbol{w}^2_x(B^{(1)}, B^{(2)}, s), \quad \boldsymbol{P} A[t_0] = 0, \tag{4-104}$$

where $B^{(1)} \in S^1[I]$, $\boldsymbol{P}^\perp B^{(1)} \in Y^1[I]$ *and* $B^{(2)} \in S^1[I]$. *Then we have*

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A} B\|_{(N \cap L^2 \dot{H}^{-1/2})_f[I]} \lesssim_{\tilde{M}} (\|B^{(1)}\|_{S^1_c[I]} + \|\boldsymbol{P}^\perp B^{(1)}\|_{Y^1_c[I]}) \|B^{(2)}\|_{S^1_d[I]} \|B\|_{S^1_e[I]}, \tag{4-105}$$

where

$$f_k = \Big(\sum_{k' < k - \kappa} \langle s 2^{2k'} \rangle^{-10} \langle s^{-1} 2^{-2k'} \rangle^{-\delta_2} c_{k'} d_{k'} \Big) e_k.$$

Finally, we end this subsection with auxiliary estimates for $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}$, which are needed to justify approximate linear energy conservation for the paradifferential wave equation.

Proposition 4.30. *Let* $\kappa \geq 10$. *We have*

$$\||D|^{-1}[\nabla, \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}]B\|_{N_{cd}} \lesssim 2^{-\delta_2\kappa}(\|\boldsymbol{P}A_x\|_{S^1_c[I]} + \|DA_0\|_{L^2\dot{H}^{1/2}_c[I]})\|B\|_{S^1_d[I]}. \tag{4-106}$$

Moreover, consider the L^2*-adjoint of* $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}$, *which is given by*

$$(\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A})^* B = \sum_k P_k \partial^\alpha [\boldsymbol{P}_\alpha A_{<k-\kappa}, B].$$

Then we have

$$\|(\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A})^* B - \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A} B\|_{N_{cd}[I]} \lesssim 2^{-\delta_2\kappa}(\|\boldsymbol{P}A_x\|_{S^1_c[I]} + \|DA_0\|_{L^2\dot{H}^{1/2}_c[I]})\|B\|_{S^1_d[I]}. \tag{4-107}$$

5. Structure of caloric Yang–Mills waves

In this section, we use the results stated in Section 4 to study properties of subthreshold caloric Yang–Mills waves satisfying an a priori S^1-norm bound on an interval.

5A. *Structure of a caloric Yang–Mills wave with finite S^1-norm.* The following theorem provides detailed properties of a caloric Yang–Mills wave with finite S^1-norm. It will be useful for the proof of the key regularity result (Theorem 6.1), as well as the main results stated in Section 1C.

For a regular solution to the Yang–Mills equation in the caloric gauge, we have seen in Theorem 3.5 that (3-12), (3-13), (3-14) and (3-15) are satisfied. More generally, we say that a one-parameter family $A(t)$ ($t \in I$) of connections in $\mathcal{C}$ (which is quite rough in general) solves the Yang–Mills equation in the caloric gauge, or in short that A is a *caloric Yang–Mills wave* if $(A, \partial_t A) \in L^\infty(I; T^{L^2}\mathcal{C})$ and satisfies (3-12), (3-13), (3-14) and (3-15).

Theorem 5.1. *Let A be a caloric Yang–Mills wave on a time interval I with energy $\mathcal{E}$ obeying*

$$A(t) \in \mathcal{C}_{\mathcal{Q}} \quad \text{for all } t \in I, \tag{5-1}$$

$$\|A\|_{S^1[I]} \leq M \tag{5-2}$$

for some $0 < \mathcal{Q}, M < \infty$. Let c be a δ_5-frequency envelope for the initial data $(A, \partial_t A)(t_0)$ ($t_0 \in I$) in $\dot{H}^1 \times L^2$. Then the following properties hold:

(1) (*linear well-posedness for $\Box_A$*) *The initial value problem for the linear equation*

$$\Box_A u = f \tag{5-3}$$

is well-posed. Moreover,

$$\|u\|_{S^1_d[I]} \lesssim_{M,\mathcal{Q}} \|(u, \partial_t u)(t_0)\|_{(\dot{H}^1\times L^2)_d} + \|f\|_{(N\cap L^2\dot{H}^{-1/2})_d[I]} \tag{5-4}$$

for any δ_5-frequency envelope d.

(2) (*frequency envelope bound*)

$$\|A\|_{S^1_c[I]} + \|\Box_A A\|_{(N\cap L^2\dot{H}^{-1/2})_{c^2}[I]} \lesssim_{M,\mathcal{Q}} 1. \tag{5-5}$$

(3) (*elliptic component bounds*)

$$\|A_0\|_{Y^1_{c^2}[I]} + \|\boldsymbol{P}^\perp A\|_{Y^1_{c^2}[I]} \lesssim_{M,\mathcal{Q}} 1. \tag{5-6}$$

(4) (*high modulation bounds*)

$$\|\Box A\|_{\Box \underline{X}^1_{c^2}[I]} + \|\Box A\|_{X^{-1/2+b_1,-b_1}_{c^2}[I]} \lesssim_{M,\mathcal{Q}} 1. \tag{5-7}$$

(5) (*paradifferential formulation*) *For any* $\kappa \geq 10$,

$$\|\Box A + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A} A\|_{(N \cap L^2 \dot{H}^{-1/2})_{c^2}[I]} \lesssim_{M,\mathcal{Q}} 2^{C\kappa}. \tag{5-8}$$

(6) (*weak divisibility*) *There exists a partition* $I = \bigcup_{i \in \mathcal{I}} I_i$ *so that* $\#\mathcal{I} \lesssim_{M,\mathcal{Q}} 1$ *and*

$$\|A\|_{S^1[I_i]} \lesssim_{\varepsilon} 1. \tag{5-9}$$

(7) (*persistence of regularity*) *If* $(A, \partial_t A)(t_0) \in \dot{H}^N \times \dot{H}^{N-1}$ $(N \geq 1)$, *then* $A \in S^N \cap S^1[I]$ *and* $A_0 \in Y^N \cap Y^1[I]$. *Moreover,*

$$\|A\|_{S^N \cap S^1[I]} + \|A_0\|_{Y^N \cap Y^1[I]} \lesssim_{M,\mathcal{Q},N} \|(A, \partial_t A)(t_0)\|_{(\dot{H}^N \times \dot{H}^{N-1}) \cap (\dot{H}^1 \times L^2)}. \tag{5-10}$$

For the subsequent properties, let $\tilde{A}$ *be another caloric Yang–Mills wave on* I *obeying the same conditions* (5-1) *and* (5-2).

(8) (*weak Lipschitz dependence on data*) *For* $\sigma < 1$ *sufficiently close to* 1, *we have*

$$\|A - \tilde{A}\|_{S^\sigma[I]} \lesssim_{M,\mathcal{Q}} \|(A - \tilde{A}, \partial_t(A - \partial_t \tilde{A}))(t_0)\|_{\dot{H}^\sigma \times \dot{H}^{\sigma-1}}. \tag{5-11}$$

(9) (*elliptic component bound for the transport equation*)

$$\|A_0\|_{(|D|^{-2} L^2_x L^1_t)_{c^2}[I]} \lesssim_{M,\mathcal{Q}} 1. \tag{5-12}$$

Moreover, if d_k *is a* δ_5*-frequency envelope for* $A - \tilde{A}$ *in* $S^1[I]$, *then*

$$\|A_0 - \tilde{A}_0\|_{(|D|^{-2} L^2_x L^1_t)_{ce}[I]} \lesssim_{M,\mathcal{Q}} 1, \tag{5-13}$$

where $e_k = c_k + c_k (c \cdot d)_{\leq k}$.

Remark 5.2. The frequency envelope bound (5-5) implies a uniform-in-time positive lower bound on the energy concentration scale r_c; see Lemma 7.8 below. As a consequence, once Theorem 1.13 is proved, finiteness of the S^1-norm would imply that solution can be continued past finite endpoints of I (we note, however, that Theorem 5.1 will be used in the proof of Theorem 1.13).

Remark 5.3. The combination of (1), (2) and the divisibility of the norm $N \cap L^2 \dot{H}^{-1/2}[I]$ (see Proposition 4.6) show that a finite S^1-norm Yang–Mills wave on I exhibits some modified scattering behavior, i.e., that each A_j tends to a homogeneous solution to the equation $\Box_A u = 0$ towards infinite endpoints of I.

We start by establishing some weaker derived bounds.

Proposition 5.4. *Let* A *be a caloric Yang–Mills wave on a time interval* I, *which obeys* $A(t) \in \mathcal{C}_{\mathcal{Q}}$ *for all* $t \in I$ *and* $\|A\|_{S^1[I]} \leq M$. *Let* c *be a* $C\delta_5$*-frequency envelope for* A *in* $S^1[I]$, *i.e.*, $\|A\|_{S^1_c[I]} \leq 1$.

(1) *The following derived bounds for $A_{t,x}$ hold*:

$$\|A_0\|_{Y^1_{c^2}[I]} + \|\boldsymbol{P}^\perp A\|_{Y^1_{c^2}[I]} \lesssim_{M,\mathcal{Q}} 1, \tag{5-14}$$

$$\|\Box A\|_{\Box \underline{X}^1_{c^2}[I]} + \|\Box A\|_{X^{-1/2+b_1,-b_1}_{c^2}[I]} \lesssim_{M,\mathcal{Q}} 1. \tag{5-15}$$

(2) *Let $\tilde{A}$ be another caloric Yang–Mills wave on I that also obeys $\|\tilde{A}\|_{S^1[I]} \leq M$. Let d be a δ_5-frequency envelope for the difference $A - \tilde{A}$ in $S^1[I]$; i.e., $\|A - \tilde{A}\|_{S^1_d[I]} \leq 1$. Then we have*

$$\|A_0 - \tilde{A}_0\|_{Y^1_e[I]} + \|\boldsymbol{P}^\perp A - \boldsymbol{P}^\perp \tilde{A}\|_{Y^1_e[I]} \lesssim_{M,\mathcal{Q}} 1, \tag{5-16}$$

$$\|\Box(A - \tilde{A})\|_{\Box \underline{X}^1_e[I]} + \|\Box(A - \tilde{A})\|_{X_e^{-1/2+b_1,-b_1}[I]} \lesssim_{M,\mathcal{Q}} 1, \tag{5-17}$$

where $e_k = d_k + c_k (c \cdot d)_{\leq k}$.

As a quick consequence of Proposition 5.4, we see that any caloric Yang–Mills wave A with $A(t) \in \mathcal{C}_\mathcal{Q}$ for all $t \in I$ and $\|A\|_{S^1[I]} \leq M$ obeys

$$\|A\|_{\underline{S}^1[I]} \lesssim_{M,\mathcal{Q}} 1.$$

Remark 5.5. The reason why we state these weaker bounds as a separate proposition is for logical clarity. As will be evident, the proof of Proposition 5.4 depends *only* on Propositions 4.12–4.22. In fact, after these propositions are established in Section 8, Proposition 5.4 will be used in the proofs of Proposition 4.23, Theorem 4.24 and Proposition 4.25 in Sections 8 and 9.

Proof of Proposition 5.4. Since A is a caloric Yang–Mills wave, Theorem 3.5 determines A_0, $\partial_0 A_0$ and $\boldsymbol{P}^\perp_j A = \Delta^{-1} \partial_j \partial^\ell A_\ell$ in terms of A. To derive the equation for $\partial_t \boldsymbol{P}^\perp A$, we first compute

$$\begin{aligned}
\partial_t \boldsymbol{P}^\perp A = \partial_t \frac{\partial_x \partial^\ell}{\Delta} A_\ell &= \Delta^{-1} \partial_x \partial^\ell (F_{0\ell} + \partial_\ell A_0 + [A_\ell, A_0]) \\
&= \Delta^{-1} \partial_x (\boldsymbol{D}^\ell F_{0\ell} + \Delta A_0 + \partial^\ell [A_\ell, A_0] - [A^\ell, F_{0\ell}]).
\end{aligned}$$

By the constraint equation, we have $\boldsymbol{D}^\ell F_{0\ell} = 0$. Expanding $F_{0\ell}$ in terms of $A_{t,x}$, we arrive at

$$\partial_t \boldsymbol{P}^\perp_j A = \partial_j A_0 + \Delta^{-1} \partial_j (\partial^\ell [A_\ell, A_0] - [A^\ell, \partial_t A_\ell] + [A^\ell, \partial_\ell A_0] - [A^\ell, [A_0, A_\ell]]). \tag{5-18}$$

The rest of the proof consists of combining Theorem 3.5 with Propositions 4.12, 4.13 and 4.22 in the right order. We first sketch the proof of the nondifference bounds (5-14)–(5-15). We begin by verifying that

$$\||D| A_0\|_{Y_{c^2}[I]} + \||D| \boldsymbol{P}^\perp A\|_{Y_{c^2}[I]} \lesssim_{M,\mathcal{Q}} 1.$$

Indeed, by the mapping properties in Theorem 3.5 and the embeddings

$$L^1 \dot{H}^1 \cap L^2 \dot{H}^{\frac{1}{2}} \subseteq Y,$$

the contributions of $\boldsymbol{A}_0^3$ in A_0 and $\boldsymbol{D}\boldsymbol{A}^3$ in $\boldsymbol{P}^\perp A$ are handled easily. For the quadratic nonlinearities, we apply (4-29) for A_0, (4-37) with $\sigma = 0$ for $\boldsymbol{P}^\perp A$ and $\sigma = 1$ for A_0.

Next, we show that

$$\|\partial_t A_0\|_{L^2 \dot{H}^{1/2}_{c^2}[I]} + \|\partial_t \boldsymbol{P}^\perp A\|_{L^2 \dot{H}^{1/2}_{c^2}[I]} \lesssim_{M,\mathcal{Q}} 1.$$

For $\partial_t A_0$, we use Theorem 3.5 for $\boldsymbol{D}\boldsymbol{A}_0^3$ and (4-30) for the quadratic nonlinearity. For $\partial_t \boldsymbol{P}^\perp A$, we estimate the right-hand side of (5-18), where we use the $Y[I]$-norm bound for A_0 that was just established.

We now consider $\Box A$. We first prove the weaker bound

$$\|\Box A\|_{\Box \widetilde{X}^1_{c^2}[I]} \lesssim_{M,\mathcal{Q}} 1. \tag{5-19}$$

By the mapping properties in Theorem 3.5 and the embeddings

$$L^1 L^2 \cap L^2 \dot{H}^{-\frac{1}{2}} \subseteq \Box \underline{X}^1 \cap X^{-\frac{1}{2}+b_1,-b_1} \subseteq \Box \widetilde{X}^1$$

the contribution of R_j is acceptable in both cases. For the quadratic nonlinearities $\boldsymbol{P}\mathcal{M}^2 + \boldsymbol{P}^\perp \mathcal{M}^2$, and the contribution of $\Box A - \Box_A A$, we apply (4-42), (4-43), (4-74), (4-77), (4-84) and (4-85); note that we need to use (5-14) in both (4-77) and (4-84).

We are ready to prove (5-17). The desired estimate for the $\Box \underline{X}^1[I]$-norm follows by repeating the preceding argument with (4-85) replaced by (4-86), and using (5-19). On the other hand, for the $\Box X^{-1/2+b_1,-b_1}[I]$-norm, we replace (4-85) by (4-87) instead, and use the $\Box \underline{X}^1[I]$-norm bound that we have just proved.

Finally, the proof of the difference bounds (5-16)–(5-17) proceeds similarly, taking the difference of each of the equations (3-12)–(3-15). We leave the details to the reader. $\square$

We now prove Theorem 5.1, using the estimates stated in Section 4.

Proof of Theorem 5.1. Throughout this proof, we omit the dependence of constants on $\mathcal{Q}$.

<u>Proof of (1)</u>: We begin with a $\Box_A$ decomposition which will be repeatedly used in the sequel. Given $\kappa > 10$, we write

$$\Box_A = \Box + \mathrm{Diff}^\kappa_{\boldsymbol{P}A} - R^\kappa_A,$$

where, using the decomposition in (4-67), the remainder R^κ_A is given by

$$R^\kappa_A = \mathrm{Diff}^\kappa_{\boldsymbol{P}^\perp A} - \mathrm{Rem}^{\kappa,2}_A - \mathrm{Rem}^{\kappa,3}_A.$$

Lemma 5.6. *Let $J \subset I$. Let d be a δ_5-frequency envelope for u in $S^1[J]$. Then we have*

$$\|R^\kappa_A u\|_{(N \cap L^2 \dot{H}^{-1/2})_d[J]} \lesssim_M (2^{-\delta_2 \kappa} \|A\|_{S^1[J]} + 2^{C\kappa} C(A,J)) \|u\|_{S^1_d[J]}, \tag{5-20}$$

with

$$C(A,J) = \|\boldsymbol{P}^\perp A\|_{Y^1[J]} + \|\boldsymbol{P}^\perp A\|_{\ell^1 L^1 L^\infty[J]} + \|A\|_{\mathrm{Str}^1[J]} + \|(\nabla \boldsymbol{P}^\perp A, \nabla A_0)\|_{L^2 \dot{H}^{1/2}[J]}. \tag{5-21}$$

Proof. We successively bound the three terms in R^κ_A as follows. For the first of them we have

$$\|\mathrm{Diff}^\kappa_{\boldsymbol{P}^\perp A} u\|_{(N \cap L^2 \dot{H}^{-1/2})_d[J]} \lesssim_M (\|\boldsymbol{P}^\perp A\|_{Y^1[J]} + \|\boldsymbol{P}^\perp A\|_{\ell^1 L^1 L^\infty[J]}) \|u\|_{S^1_d[J]},$$

using the bounds (4-82) and (4-83), and noting that the second norm of A is estimated using (4-37) for the quadratic part and (3-22) by

$$\|\boldsymbol{P}^{\perp} A\|_{\ell^1 L^1 L^\infty[J]} \lesssim_M 1.$$

For the second term in R^κ_A in (5-22) we have

$$\|\mathrm{Rem}^{\kappa,2}_A u\|_{(N\cap L^2\dot{H}^{-1/2})_d[J]} \lesssim_M (2^{-\delta_2\kappa}\|A\|_{S^1[J]} + 2^{C\kappa}C(A,J))\|u\|_{S^1_d[J]},$$

as a consequence of (4-78), (4-79) and (4-80).

Finally, for the third term in R^κ_A we have

$$\|\mathrm{Rem}^{\kappa,3}_A u\|_{(N\cap L^2\dot{H}^{-1/2})_d[J]} \lesssim_M \|A\|_{\mathrm{Str}^1[J]}\|u\|_{S^1_d[J]}$$

due to (4-74). □

To prove (1) we rewrite (5-3) in the form

$$(\Box + \mathrm{Diff}^\kappa_{\boldsymbol{P}A})u = f - R^\kappa_A u. \tag{5-22}$$

The important fact is that all the A norms in $C(A,J)$ except for S^1 are divisible norms, and also controlled by M. On the other hand the S^1 norm of A has the redeeming $2^{-\delta_2\kappa}$ factor. To proceed we choose κ large enough,

$$\kappa \ll_{M,\mathcal{Q}} 1.$$

Then we can subdivide the interval $I = \bigcup_{j\in\mathcal{J}} J_k$ so that $\#\mathcal{J} \lesssim_M 1$, and so that in each interval J_j we have smallness,

$$\|R^\kappa_A u\|_{(N\cap L^2\dot{H}^{-1/2})_d[J_j]} \ll_M \|u\|_{S^1_d[J_j]}. \tag{5-23}$$

A second consequence of our choice for κ is that Theorem 4.24 applies. Then we can successively apply Theorem 4.24 in each interval J_k, treating R^κ_A perturbatively.

<u>Proof of (2)</u>: The argument here is similar to the previous one. For any interval $J \subset I$ and any $(-\delta_5, N)$-frequency envelope d for A in $S^1[J]$ we can use the bounds (4-44)–(4-49) and (3-21) to estimate

$$\|\Box_A A\|_{(N\cap L^2\dot{H}^{-1/2})_d[J]} \lesssim_M (2^{-\delta_2\kappa}\|A\|_{S^1[J]} + 2^{C\kappa}\|A\|_{DS^1[J]})\|A\|_{S^1_d[J]}. \tag{5-24}$$

As before we use the divisibility of the DS^1 norm to partition the interval I into finitely many subintervals J_k, whose number depends only on M, and so that in each subinterval we have

$$2^{-\delta_2\kappa}\|A\|_{S^1[J]} + 2^{C\kappa}\|A\|_{DS^1[J]} \le \varepsilon \ll_{M,\mathcal{Q}} 1.$$

We now specialize the choice of d, choosing it to be a minimal δ_5-frequency envelope for A in the first interval J_1. Applying the result in part (1) in J_1 we conclude that

$$d \lesssim_{M,\mathcal{Q}} c + \varepsilon d,$$

which by the smallness of ε implies that $d \lesssim_{M,\mathcal{Q}} c$. Then we reiterate.

<u>Proofs of (3) and (4)</u>: These follow from (5-5) and Proposition 5.4.

Proof of (5): This is obtained by combining the bound (5-20) for $J = I$ and $u = A$ with the bound (5-24).

Proof of (6): In view of (5), this is a direct consequence of Proposition 4.25.

Proof of (7): We use frequency envelopes. It suffices to show that if c_k is a $(-\delta_5, S)$-frequency envelope for the initial data in the energy space then $C(M)c_k$ is a frequency envelope for A in S^1 and A_0 in Y^1. We begin with a version of Lemma 5.6:

Lemma 5.7. *Let $J \subset I$. Let $d = d(J)$ be a $(-\delta_5, S)$-frequency envelope for A in $S^1[J]$. Then we have*

$$\|R^\kappa_A A\|_{(N\cap L^2\dot{H}^{-1/2})_d[J]} \lesssim_M (2^{-\delta_2\kappa}\|A\|_{S^1[J]} + 2^{C\kappa}C(A,J))\|A\|_{S^1_d[J]}. \tag{5-25}$$

Proof. The same argument as in the proof of (5-8) applies for the first term in R^κ_A, as there the output frequency and the u input frequency are the same. On the other hand for the two remaining terms, the frequency envelope d is inherited from the highest frequency input; see Propositions 4.19, 4.20. □

Combining the bound in the lemma with (5-24) we obtain the estimate

$$\|\Box A + \mathrm{Diff}^\kappa_{\boldsymbol{P}A} A\|_{(N\cap L^2\dot{H}^{-1/2})_d[J]} \lesssim_M (2^{-\delta_2\kappa}\|A\|_{S^1[J]} + 2^{C\kappa}C(A,J))\|A\|_{S^1_d[J]}. \tag{5-26}$$

Now we can conclude as in the proof of (2). We first choose κ large enough so that Theorem 4.24 applies, and also so that

$$2^{-\delta_2\kappa}\|A\|_{S^1[I]} \ll_M 1.$$

Then we divide the interval I into finitely many subintervals (again, depending only on M and $\mathcal{Q}$) so that for each subinterval J we have

$$2^{C\kappa}\|A\|_{DS^1[J]} \ll_M 1.$$

Thus, for each subinterval J we have ensured that

$$\|\Box A + \mathrm{Diff}^\kappa_{\boldsymbol{P}A} A\|_{(N\cap L^2\dot{H}^{-1/2})_d[J]} \ll_M \|A\|_{S^1_d[J]}.$$

Let c_k be a $(-\delta_5, S)$-frequency envelope for the initial data in the energy space, Then applying Theorem 4.24 in the first interval J_1 we conclude that

$$\|P_k A\|_{S^1[J_1]} \lesssim_{M,\mathcal{Q}} c_k + \varepsilon d_k, \quad \varepsilon \ll_M 1, \tag{5-27}$$

for any $(-\delta_5, S)$-frequency envelope d_k for A in $S^1[J_1]$. In particular if d_k is a minimal $(-\delta_5, S)$-frequency envelope for A in $S^1[J_1]$ then we obtain

$$d_k \lesssim_M c_k + \varepsilon d_k,$$

which leads to

$$d_k \lesssim_{M,\mathcal{Q}} c_k,$$

i.e., the desired bound in J_1. We now reiterate this bound in successive intervals J_j. Finally, the Y bound follows as in (3).

Proof of (8): Assume $0 < 1-\sigma \ll \delta_5$. We write the equation for $\delta A = A - \tilde{A}$ in the form

$$(\Box + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}\tilde{A}})\delta A = F^{\kappa},$$

where

$$F^{\kappa} = \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A-\boldsymbol{P}\tilde{A}} A + (R^{\kappa}_{A} A - R^{\kappa}_{\tilde{A}} \tilde{A}) + (\Box_A A - \Box_A \tilde{A}). \tag{5-28}$$

We claim that we can estimate the terms in F^{κ} as follows:

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A-\boldsymbol{P}\tilde{A}} A\|_{N^{\sigma-1}\cap L^2\dot{H}^{\sigma-1-1/2}[J]} \lesssim_M 2^{-c_\sigma \kappa}(\|A\|_{S^1} + \|\tilde{A}\|_{S^1})\|\delta A\|_{S^\sigma[J]}, \tag{5-29}$$

$$\|R^{\kappa}_{A} A - R^{\kappa}_{\tilde{A}} \tilde{A}\|_{N^{\sigma-1}\cap L^2\dot{H}^{\sigma-1-1/2}[J]} \lesssim_M 2^{C\kappa}(C(A,J) + C(\tilde{A},J))\|\delta A\|_{S^\sigma[J]}, \tag{5-30}$$

$$\|\Box_A A - \Box_A \tilde{A}\|_{N^{\sigma-1}\cap L^2\dot{H}^{\sigma-1-1/2}[J]} \lesssim_M (C(A,J) + C(\tilde{A},J))\|\delta A\|_{S^\sigma[J]}. \tag{5-31}$$

We first show how to conclude the proof of (8) using (5-29), (5-30) and (5-31). As in the proofs of (1), (2) and (7), we first choose κ large enough, $\kappa \gg_M 1$. Then we use divisibility for the expressions $C(A,J)$ and $C(\tilde{A},J)$ in order to divide the interval I into subintervals J_j so that on each subinterval F^{κ} is perturbative, i.e.,

$$\|F^{\kappa}\|_{N^{\sigma-1}\cap L^2\dot{H}^{\sigma-1-1/2}[J_j]} \ll_{M,\kappa} \|\delta A\|_{S^\sigma[J_j]}.$$

Finally, we apply Theorem 4.24 successively on the intervals J_j; then (8) follows.

It remains to prove the bounds (5-29), (5-30) and (5-31). The bounds (5-30) and (5-31) are the difference counterparts of (5-25) and (5-24), respectively, and are proved in a very similar fashion. Details are omitted. We only remark that the requirement $\sigma < 1$ is not needed here, and that these bounds hold for any δ_5-admissible frequency envelope c_k for δA in S^1.

We now turn our attention to the novel part of the argument, which is the bound for $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A-\boldsymbol{P}\tilde{A}} A$. It is here that the condition $\sigma < 1$ pays a critical role. This is done in the next lemma. For later use we state the result in a more general fashion. This will be needed again in the proof of Proposition 6.4. A variation of the same argument will also be needed in Proposition 6.3.

Lemma 5.8. *Let $J \subset I$. Let c_k, d_k, b_k be frequency envelopes for A, $\tilde{A}$, respectively δA and B in $S^1[J]$. Then the expression $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A-\boldsymbol{P}\tilde{A}} B$ can be estimated as*

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A-\boldsymbol{P}\tilde{A}} B\|_{(N\cap L^2\dot{H}^{-1/2})_f[J]} \lesssim_{M,\mathcal{Q}} 2^{-c_\sigma \kappa}\|\delta A\|_{S^\sigma_d[J]}\|B\|_{S^1_b[J]}, \tag{5-32}$$

where f_k is given by

$$f_k = \Big(\sum_{k' \le k-\kappa} d_{k'} + c_{k'}(c\cdot d)_{\le k'}\Big) b_k. \tag{5-33}$$

Before proving the lemma we show that it implies (5-29). To measure δA in S^σ we can choose the frequency envelope d_k with the property that $2^{(\sigma-1)k} d_k$ is a $(-\delta, 1-\sigma+\delta)$-admissible envelope with $\delta < \frac{1}{2}(1-\sigma)$, $\delta \ll \delta_5$, and so that

$$\|\delta A\|^2_{S^\sigma[J]} \approx \sum_k (2^{(\sigma-1)k} d_k)^2.$$

Then we have

$$f_k \lesssim_M d_{k-\kappa} c_k \lesssim_M 2^{-\frac{1}{2}(1-\sigma)\kappa} d_k,$$

and (5-29) follows. We return to the proof of the lemma:

Proof of Lemma 5.8. We first recall the equations for $\boldsymbol{P} A_x$ and A_0. Following Theorem 3.5, these have the form

$$\begin{aligned} \Box \boldsymbol{P} A_x &= \boldsymbol{P}[A^\ell, \partial_x A_\ell] - 2\boldsymbol{P}[A_\ell, \partial^\ell A_x] + \boldsymbol{P}(R(A) + [A_\ell, [A^\ell, A_x]]), \\ \Delta A_0 &= [A^\ell, \partial_x A_\ell] + \boldsymbol{Q}(A, \partial_0 A) + \Delta \boldsymbol{A}_0^3. \end{aligned} \tag{5-34}$$

Based on these equations we consider the following decomposition of $\boldsymbol{P} A = (\boldsymbol{P} A_x, A_0)$:

$$\boldsymbol{P} A = (A_x^{\text{main}}, A_0^{\text{main}}) + (A_x^2, 0) + (A_x^3, A_0^3),$$

where the three components are determined by the following three sets of equations:

$$\begin{aligned} \Box A_x^{\text{main}} &= \boldsymbol{P}[A^\ell, \partial_x A_\ell], \quad A_x^{\text{main}}[0] = 0, \\ \Delta A_0^{\text{main}} &= [A^\ell, \partial_x A_\ell], \end{aligned}$$

$A_0^2 = 0$, and

$$\Box A_x^2 = -2\boldsymbol{P}[A_\ell, \partial^\ell A_x], \quad A_x^2[0] = 0,$$

and finally

$$\begin{aligned} \Box A_x^3 &= \boldsymbol{P}(R(A) + \boldsymbol{P}[A_\ell, [A^\ell, A_x]]), \quad A_x^3[0] = \boldsymbol{P} A[0], \\ \Delta A_0^3 &= \boldsymbol{Q}(A, \partial_0 A) + \Delta \boldsymbol{A}_0^3. \end{aligned} \tag{5-35}$$

We also use the same set of equations and the same decomposition for $\boldsymbol{P}\tilde{A}$, and take the differences δA^{main}, δA^2 and δA^3. We are now ready to estimate the three contributions.

The contribution of δA^{main}. For this we use the estimates in Proposition 4.27, which yield

$$\|\text{Diff}^\kappa_{\boldsymbol{P} A^{\text{main}} - \boldsymbol{P}\tilde{A}^{\text{main}}} B\|_{(N \cap L^2 \dot{H}^{-1/2}[J])_f} \lesssim_M 2^{-\sigma\kappa} \|\delta A\|_{S^1_d[J]} \|B\|_{S^1_b[J]}, \tag{5-36}$$

where

$$f_k = \biggl(\sum_{k' \le k-\kappa} c_{k'} d_{k'}\biggr) b_k,$$

which suffices. For later use, we also record the following consequence of Proposition 4.15, which provides a bound for $\|\Box \delta A_x^{\text{main}}\|_{N \cap L^2 \dot{H}^{1/2}}$:

$$\|\delta A_x^{\text{main}}\|_{S^1_{cd}[J]} \lesssim \|\delta A\|_{S^1_d[J]} (\|A\|_{S^1_c[J]} + \|\tilde{A}\|_{S^1_c[J]}). \tag{5-37}$$

The contribution of δA^3. This is more easily dealt with using instead Proposition 4.26. We start with $A_0^3 - \tilde{A}_0^3$, which is estimated using the bounds (4-36) and (4-37) in Proposition 4.13 for the first term, and (3-23) for the second, by

$$\|A_0^3 - \tilde{A}_0^3\|_{(L^1 L^\infty \cap L^2 \dot{H}^{3/2})_{cd}[J]} \lesssim_M \|\delta A\|_{S^1_d[J]} (\|A\|_{S^1_c[J]} + \|\tilde{A}\|_{S^1_c[J]}). \tag{5-38}$$

Similarly, for $A_x^3 - \tilde{A}_x^3$ we can apply the difference bound associated to (3-21) for R_x and Strichartz estimates for the remaining cubic term to obtain

$$\|\Box(A_x^3 - \tilde{A}_x^3)\|_{(L^1L^2 \cap L^2\dot{H}^{-1/2})_{cd}[J]} \lesssim_M \|\delta A\|_{S_d^1[J]}(\|A\|_{S_c^1[J]} + \|\tilde{A}\|_{S_c^1[J]}). \tag{5-39}$$

As a consequence this also gives

$$\|A_x^3 - \tilde{A}_x^3\|_{S_{cd}^1[J]} \lesssim_M \|\delta A\|_{S_d^\sigma[J]}(\|A\|_{S_c^1[J]} + \|\tilde{A}\|_{S_c^1[J]}). \tag{5-40}$$

Using (5-38) and (5-40) in Proposition 4.26 yields the desired bound

$$\|\mathrm{Diff}^\kappa_{\delta A^3} B\|_{(N \cap L^2\dot{H}^{-1/2})_f[J]} \lesssim_{M,\mathcal{Q}} \|\delta A\|_{S_d^1[J]}\|B\|_{S_b^1[J]}(\|A\|_{S_c^1[J]} + \|\tilde{A}\|_{S_c^1[J]}), \tag{5-41}$$

with the same f_k as in the previous case.

The contribution of A^2. Here we will use Proposition 4.28. For this we need to verify its hypotheses. We begin with (4-101), for which we combine (5-37) and (5-40) to conclude that

$$\|\delta A_x^2\|_{S_d^1[J]} \lesssim_M \|\delta A\|_{S_d^1[J]}. \tag{5-42}$$

Next we consider (4-100). Using the second part of Proposition 5.4 we obtain

$$\|\delta A\|_{\underline{S}_e^1[J]} + \|(\delta A_0, \boldsymbol{P}^\perp \delta A)\|_{Y_e^1[J]} \lesssim_M \|\delta A\|_{S_d^1[J]}, \tag{5-43}$$

with

$$e_k = d_k + c_k(c \cdot d)_{<k}.$$

The last two bounds allow us to use Proposition 4.28. This yields

$$\|\mathrm{Diff}^\kappa_{\delta A^2} B\|_{(N \cap L^2\dot{H}^{-1/2})f[J])} \lesssim_{M,\mathcal{Q}} \|\delta A\|_{S_d^1[J]}\|B\|_{S_c^1[J]}(\|A\|_{S_c^1[J]} + \|\tilde{A}\|_{S_c^1[J]}), \tag{5-44}$$

where

$$f_k = \Big(\sum_{k' \leq k-\kappa} d_{k'} + e_{k'} d_{k'}\Big) b_k.$$

The proof of the lemma is now concluded. □

Proof of (9): This is a direct consequence of the bounds (4-39) and (3-23) for the quadratic part $\boldsymbol{A}_0^2$ of A_0, and its cubic and higher part $\boldsymbol{A}_0^3$. □

5B. *Caloric Yang–Mills waves with small energy dispersion on a short interval.* Next, we consider the effect of small inhomogeneous energy dispersion on a time interval with compatible scale.

Theorem 5.9. *Let A be a caloric Yang–Mills wave on a time interval I with energy $\mathcal{E}$, obeying (5-1), (5-2), as well as the smallness relations*

$$\|F\|_{ED_{\geq 0}[I]} \leq \varepsilon, \quad |I| \leq \varepsilon. \tag{5-45}$$

Let c be a δ_5-frequency envelope for A in $S^1[I]$. Then for sufficiently small $\varepsilon > 0$ depending on M and $\mathcal{Q}$, the following properties hold:

(1) (*small energy dispersion below scale* 1 *for* A)

$$\|A\|_{ED^1_{\geq 0}[I]} \lesssim_{\mathcal{E},\mathcal{Q}} \varepsilon^{\delta_2}. \tag{5-46}$$

(2) (*elliptic component bounds*)

$$\|A_0\|_{Y^1_c[I]} + \|\boldsymbol{P}^\perp A\|_{Y^1_c[I]} \lesssim_{M,\mathcal{Q}} \varepsilon^{\delta_2}. \tag{5-47}$$

(3) (*high modulation bounds*)

$$\|\Box A\|_{L^2 \dot{H}_c^{-1/2}[I]} \lesssim_{M,\mathcal{Q}} \varepsilon^{\delta_2}. \tag{5-48}$$

(4) (*paradifferential formulation*)

$$\|\Box A + \mathrm{Diff}^\kappa_{\boldsymbol{P}A} A\|_{(N\cap L^2\dot{H}^{-1/2})_c[I]} \lesssim_{M,\mathcal{Q}} \varepsilon^{\delta_4} 2^{C\kappa}. \tag{5-49}$$

(5) (*approximate linear energy conservation*) *For any* $t_1, t_2 \in I$,

$$\big| \|\nabla A(t_1)\|^2_{L^2} - \|\nabla A(t_2)\|^2_{L^2} \big| \lesssim_{M,\mathcal{Q}} \varepsilon^{\delta_4}. \tag{5-50}$$

(6) (*approximate conservation of* $\mathcal{Q}$) *For any* $t_1, t_2 \in I$,

$$|\mathcal{Q}(A(t_1) - \mathcal{Q}(A(t_2))| \lesssim_{\mathcal{E},\mathcal{Q}} \varepsilon^{\delta_4}. \tag{5-51}$$

Proof. Again, we omit the dependence of constants on $\mathcal{Q}$. The property that will be used here repeatedly is (4-21), which asserts that all nonsharp Strichartz norms are small. We recall it here for convenience:

$$\sup_k \|P_k F\|_{\mathrm{Str}} \lesssim_M \varepsilon^{\delta_1} \lesssim \varepsilon^{\delta_2}. \tag{5-52}$$

Proof of (1): This is a consequence of the caloric bound (3-7) applied with $d_k = \varepsilon$.

Proof of (2): We repeat the arguments in the proof of Proposition 5.4(1). The bounds for the cubic and higher terms in Theorem 3.5 use only the Strichartz Str^1 norms, so the contributions of $\boldsymbol{A}_0^3$ in A_0, $\boldsymbol{DA}^3$ in $\boldsymbol{P}^\perp A$ and $\boldsymbol{DA}_0^3$ in $\partial_t A_0$ are easily estimated. For the quadratic terms we replace (4-29) with (4-33) in the case of A_0, and then (4-37) with (4-38) in the case of $\boldsymbol{P}^\perp A$ and $\partial_t A_0$; again the smallness comes from Str^1.

Proof of (3): We consider the terms in the A_x equation in Theorem 3.5. The cubic terms R_x and $[A_\ell, [A^\ell, A]]$ are estimated only in terms of $\|A\|_{\mathrm{Str}^1}$. For the quadratic terms we use instead the bounds (4-30), (4-36), (4-63) and (4-65); all smallness comes from Str^1.

Proof of (4): We first establish the similar bound for $\Box_A A$, which is given by (3-12). For the quadratic terms we use (4-50) and (4-51). For the cubic term we use (3-21). Hence it remains to estimate the difference

$$R^\kappa_A A = \mathrm{Diff}^\kappa_{\boldsymbol{P}^\perp A} A - \mathrm{Rem}^{\kappa,2}_A A - \mathrm{Rem}^{\kappa,3}_A A.$$

For the first term we use (4-83), where the ε smallness comes from the L^1L^∞ norm of $\boldsymbol{P}^\perp A$ due to the bounds (4-38) and (3-22) for the quadratic and cubic parts of $A^\perp$ respectively.

For the second term we use the bound (4-81). The second term on the right is small due to (5-47), so we obtain

$$\|\mathrm{Rem}^{\kappa,2}_A A\|_{(N\cap L^2\dot{H}^{-1/2})_c} \lesssim_M (2^{-\delta_2\kappa} + 2^{C\kappa}\varepsilon^{\delta_2})\|A\|_{\underline{S}^1_c}.$$

Now we observe that on the right we can replace κ with any $\kappa' > \kappa$ without any change in the proof. Then it suffices to optimize with respect to κ'.

For the third term we use directly (4-74).

Proof of (5): This statement is a corollary of (5-49). For the proof, we introduce the *linear energy*

$$E_{\mathrm{lin}}(A)(t) = \frac{1}{2}\int_{\mathbb{R}^4} \sum_{\mu=0}^{4} |\partial_\mu A(t)|^2 \, dx.$$

Given any interval $I' = (t_1, t_2) \subseteq I$, we consider

$$\mathcal{I} = \int_{\mathbb{R}\times\mathbb{R}^4} \chi_{I'} \langle (\Box + \mathrm{Diff}^\kappa_{\boldsymbol{P}A}) A, \partial_t A\rangle \, dt\, dx.$$

Integrating by parts, we may rewrite

$$\mathcal{I} = E_{\mathrm{lin}}(A)(t_1) - E_{\mathrm{lin}}(A)(t_2) + \frac{1}{2}\int \langle \mathrm{Diff}^\kappa_{\boldsymbol{P}A} A, A\rangle(t_2)\, dx - \frac{1}{2}\int \langle \mathrm{Diff}^\kappa_{\boldsymbol{P}A} A, A\rangle(t_1)\, dx$$
$$- \frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}^4} \chi_{I'} \langle [\partial_t, \mathrm{Diff}^\kappa_{\boldsymbol{P}A}] A, A\rangle \, dt\, dx + \frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}^4} \chi_{I'} \langle (\mathrm{Diff}^\kappa_{\boldsymbol{P}A} - (\mathrm{Diff}^\kappa_{\boldsymbol{P}A})^*) A, \partial_t A\rangle \, dt\, dx.$$

By Proposition 4.30 and the straightforward bound

$$\int \langle \mathrm{Diff}^\kappa_{\boldsymbol{P}A} A, A\rangle(t) \lesssim 2^{-\kappa} \|(A, A_0)(t)\|_{\dot{H}^1} \|\nabla A(t)\|^2_{L^2} \lesssim_M 2^{-\kappa},$$

we see that

$$|\mathcal{I} - (E_{\mathrm{lin}}(A)(t_1) - E_{\mathrm{lin}}(A)(t_2))| \lesssim_M 2^{-c\kappa}. \tag{5-53}$$

On the other hand, by duality, we may put $\chi_{I'}(\Box + \mathrm{Diff}^\kappa_{\boldsymbol{P}A})A$ and $\chi_{I'}\partial_t A$ in N and N^*, respectively. Then by Proposition 4.6, (5-2) and (5-49), we have

$$|\mathcal{I}| \lesssim_M \varepsilon^{\delta_4} 2^{C\kappa}. \tag{5-54}$$

Optimizing the choice of κ, (5-50) follows.

Proof of (6): We will use the caloric flow in order to compare $\mathcal{Q}(A(t_1))$ and $\mathcal{Q}(A(t_2))$. Denote by $A(t,s)$ the caloric flow of A. We will split the difference in three as

$$\mathcal{Q}(A(t_1)) - \mathcal{Q}(A(t_2)) = \mathcal{Q}(A(t_1,1)) - \mathcal{Q}(A(t_2,1)) + \mathcal{Q}(A(t_1)) - \mathcal{Q}(A(t_1,1)) - \mathcal{Q}(A(t_2)) + \mathcal{Q}(A(t_2,1)).$$

For the first difference we estimate at parabolic time $s = 1$ as follows:

$$\begin{aligned}|\mathcal{Q}(A(t_1,1)) - \mathcal{Q}(A(t_2,1))| &\lesssim \int_{t_1}^{t_2}\int_{\mathbb{R}^4} \frac{d}{dt}|F(s,t,x)|^3 \, dx\, dt \\ &\lesssim \int_{t_1}^{t_2}\int_{\mathbb{R}^4} |F(1,t,x)|^2 |\partial_t F(1,x,t)| \, dx\, dt \\ &\lesssim \int_{t_1}^{t_2}\int_{\mathbb{R}^4} |F(s,t,x)|^2 |\partial_t F| \, dx\, dt \lesssim_{\mathcal{E},\mathcal{Q}} |t_1 - t_2| c_1^3,\end{aligned}$$

where at the last step we have simply used the fixed-time L^2 bounds given by Proposition 3.1(1) and Bernstein's inequality. Now we gain smallness from the time interval.

For the remaining two differences we only need fixed-time estimates, which for reference we state in the following.

Lemma 5.10. *Let $a \in \mathcal{C}$ be a caloric connection with energy $\mathcal{E}$ and $\mathcal{Q}(A) = \mathcal{Q}$, and A its caloric Yang–Mills flow.*

(a) *Assume that a is energy-dispersed at high frequencies,*

$$\|f\|_{\mathrm{ED}_{\geq m}} \leq \varepsilon. \tag{5-55}$$

Then for its caloric Yang–Mills heat flow $A(s)$ we have

$$\mathcal{Q}(a) - \mathcal{Q}(A(2^{-2m})) \lesssim_{\mathcal{E},\mathcal{Q}} \varepsilon^c. \tag{5-56}$$

(b) *If a is fully energy-dispersed,*

$$\|f\|_{\mathrm{ED}} \leq \varepsilon, \tag{5-57}$$

then we have

$$\mathcal{Q}(a) \lesssim_{\mathcal{E},\mathcal{Q}} \varepsilon^c. \tag{5-58}$$

Proof. (a) By scaling we can set $m = 0$. Denote by c_k a frequency envelope for f in L^2, and by d_k a frequency envelope for f in $\dot{W}^{-2,\infty}$. By the energy dispersion bound we have $d_k \leq \varepsilon$ for $k \geq 0$. By Proposition 3.2 we have the L^2 bound

$$\|P_k F\|_{L^2} \lesssim_{\mathcal{E},\mathcal{Q}} c_k \langle 2^{2k} s\rangle^{-N},$$

and the L^∞ bound

$$\|P_k F\|_{L^\infty} \lesssim_{\mathcal{E},\mathcal{Q}} 2^{2k} d_k^{\frac{1}{2}} \langle 2^{2k} s\rangle^{-N}.$$

We use these bounds to estimate the difference

$$\begin{aligned}
\mathcal{Q}(a) - \mathcal{Q}(A(1)) &= \int_0^1 \int_{\mathbb{R}^4} |F(s,t,x)|^3 \, dx \, ds \\
&\lesssim \sum_{k_1 \leq k_2 \leq k_3} \int_0^1 \int_{\mathbb{R}^4} |P_{k_1} F(s,t,x)| \, |P_{k_2} F(s,t,x)| \, |P_{k_3} F(s,t,x)| \, dx \, ds \\
&\lesssim_{\mathcal{E},\mathcal{Q}} \sum_{k_1 \leq k_2 \leq k_3} \frac{1}{1 + 2^{2k_3}} 2^{2k_1} d_{k_1}^{\frac{1}{2}} c_{k_2} c_{k_3} \\
&\lesssim \sum_{1 \leq k_3} d_{k_3}^{\frac{1}{2}} c_{k_3}^2 \lesssim \varepsilon^{\frac{1}{2}},
\end{aligned}$$

where at the next to last step we have used both the low-frequency decay and the off-diagonal decay for the summation in k_1 and k_2.

(b) This follows by letting $m \to -\infty$ in part (a). The proof of the lemma is concluded. □

The proof of (5-51) is also concluded. □

5C. ***The dynamic Yang–Mills heat flow of a caloric Yang–Mills wave.*** Here we investigate the structure of the dynamic Yang–Mills heat flow of a caloric Yang–Mills wave A with finite S^1-norm. As before, we consider two cases: (1) when A only obeys a finite S^1-norm bound; and (2) when A has small inhomogeneous energy dispersion on a short time interval of compatible scale.

In the general case, we have the following structure theorem.

Theorem 5.11. *Let A be a caloric Yang–Mills wave with energy $\mathcal{E}$ on a time interval I, obeying (5-1) and (5-2). Let $A_{t,x}(s)$ be the dynamic Yang–Mills heat flow of $A_{t,x}$ at heat-time $s > 0$ in the caloric gauge. Then the following properties hold*:

(1) (*fixed-time bounds*) *For any $t \in I$, let $c^{(0)}(t)$ be a δ_5-frequency envelope for $\nabla A(t)$ in L^2. Then*

$$\|P_k(\nabla A(s) - \nabla e^{s\Delta} A)(t)\|_{L^2} \lesssim_{\mathcal{E},\mathcal{Q}} \langle 2^{-2k} s^{-1}\rangle^{-\delta_4} \langle 2^{2k} s\rangle^{-10} c_k^{(0)}(t)^2, \tag{5-59}$$

$$\|P_k \partial^\ell A_\ell(t,s)\|_{L^2} \lesssim_{\mathcal{E},\mathcal{Q}} \langle 2^{2k} s\rangle^{-10} c_k^{(0)}(t)^2, \tag{5-60}$$

$$\|P_k \nabla A_0(t,s)\|_{L^2} \lesssim_{\mathcal{E},\mathcal{Q}} \langle 2^{2k} s\rangle^{-10} c_k^{(0)}(t)^2, \tag{5-61}$$

$$\|P_k \Box A(t,s)\|_{\dot{H}^{-1}} \lesssim_{\mathcal{E},\mathcal{Q}} \langle 2^{2k} s\rangle^{-10} c_k^{(0)}(t)^2. \tag{5-62}$$

(2) (*frequency envelope bounds*) *Let c be a δ_5-frequency envelope for A in $S^1[I]$. Then*

$$\|P_k(A(s) - e^{s\Delta} A)\|_{\underline{S}^1[I]} \lesssim_{M,\mathcal{Q}} \langle 2^{-2k} s^{-1}\rangle^{-\delta_4} \langle 2^{2k} s\rangle^{-10} c_k^2, \tag{5-63}$$

$$\|P_k A_0(s)\|_{Y^1[I]} \lesssim_{M,\mathcal{Q}} \langle 2^{2k} s\rangle^{-10} c_k^2, \tag{5-64}$$

$$\|P_k \boldsymbol{P}^\perp A(s)\|_{Y^1[I]} \lesssim_{M,\mathcal{Q}} \langle 2^{2k} s\rangle^{-10} c_k^2. \tag{5-65}$$

(3) (*derived difference bounds*) *Let $\tilde{A}$ be a caloric Yang–Mills wave on I obeying $\|\tilde{A}\|_{S^1[I]} \le \widetilde{M}$, and let d be a δ_5-frequency envelope for the difference $A(s) - \tilde{A}$ in $S^1[I]$. Then*

$$\|P_k(A_0(s) - \tilde{A}_0)\|_{Y^1[I]} + \|P_k(\boldsymbol{P}^\perp A(s) - \boldsymbol{P}^\perp \tilde{A})\|_{Y^1_d[I]} \lesssim_{M,\widetilde{M},\mathcal{Q}} e_k + \min\{1, (s^{-\frac{1}{2}}|I|)^{\delta_4}\}\langle 2^{-2k} s^{-1}\rangle^{-\delta_4} \langle 2^{2k} s\rangle^{-10} c_k^2, \tag{5-66}$$

$$\|P_k \Box(A(s) - \tilde{A})\|_{\Box \underline{X}^1[I]} + \|P_k \Box(A(s) - \tilde{A})\|_{X^{-1/2+b_1,-b_1}[I]} \lesssim_{M,\widetilde{M},\mathcal{Q}} e_k + \min\{1, (s^{-\frac{1}{2}}|I|)^{\delta_4}\}\langle 2^{-2k} s^{-1}\rangle^{-\delta_4} \langle 2^{2k} s\rangle^{-10} c_k^2, \tag{5-67}$$

where $e_k = d_k + c_k (c \cdot d)_{\le k}$.

Remark 5.12. Combining (5-63) with the obvious bound for $e^{s\Delta} A$, we get the simple bound

$$\|P_k A(s)\|_{\underline{S}^1[I]} \lesssim_{M,\mathcal{Q}} \langle 2^{2k} s\rangle^{-10} c_k. \tag{5-68}$$

Next, we consider the effect of small inhomogeneous energy dispersion on a time interval of compatible scale.

Theorem 5.13. *Let A be a caloric Yang–Mills wave with energy $\mathcal{E}$ on a time interval I, obeying (5-1), (5-2) and (5-45), and $A_{t,x}(s)$ be the dynamic Yang–Mills heat flow of $A_{t,x}$ at heat-time $s > 0$ in the caloric gauge. Let c be a δ_5-frequency envelope for A in $S^1[I]$. Then the following properties hold*:

(1) (*fixed-time smallness bound*)

$$\|\nabla P_k(A(s)-e^{s\Delta}A)(t)\|_{L^2} \lesssim_{\mathcal{E},\mathcal{Q}} 2^{\delta_4(m-k)_+}\varepsilon^{\delta_4}\langle 2^{-2k}s^{-1}\rangle^{-\delta_4}\langle 2^{2k}s\rangle^{-10}c_k^{(0)}(t), \tag{5-69}$$

$$\|P_k\partial^\ell A_\ell(t,s)\|_{L^2} \lesssim_{\mathcal{E},\mathcal{Q}} 2^{\delta_4(m-k)_+}\varepsilon^{\delta_4}\langle 2^{2k}s\rangle^{-10}c_k^{(0)}(t). \tag{5-70}$$

(2) (*small energy dispersion below scale* 1 *for* $A(s)$)

$$\|A(s)\|_{ED^{-1}_{\geq 0}[I]} \lesssim_{\mathcal{E},\mathcal{Q}} \varepsilon^{\delta_4}. \tag{5-71}$$

(3) (*frequency envelope bounds*)

$$\|P_k(A(s)-e^{s\Delta}A)\|_{\underline{S}^1[I]} \lesssim_{M,\mathcal{Q}} \varepsilon^{\delta_4}\langle 2^{-2k}s^{-1}\rangle^{-\delta_4}\langle 2^{2k}s\rangle^{-10}c_k, \tag{5-72}$$

$$\|P_kA_0(s)\|_{Y^1[I]} \lesssim_{M,\mathcal{Q}} \varepsilon^{\delta_4}\langle 2^{2k}s\rangle^{-10}c_k, \tag{5-73}$$

$$\|P_k\boldsymbol{P}^\perp A(s)\|_{Y^1[I]} \lesssim_{M,\mathcal{Q}} \varepsilon^{\delta_4}\langle 2^{2k}s\rangle^{-10}c_k. \tag{5-74}$$

(4) (*derived difference bounds*) *Let $\tilde{A}$ be a caloric Yang–Mills wave on I with $\|\tilde{A}\|_{S^1[I]} \leq \widetilde{M}$, and let d be a δ_5-frequency envelope for the difference $A(s)-\tilde{A}$ in $S^1[I]$. Then*

$$\|P_k(A_0(s)-\tilde{A}_0)\|_{Y^1[I]} + \|P_k(\boldsymbol{P}^\perp A(s)-\boldsymbol{P}^\perp\tilde{A})\|_{Y^1_d[I]} \lesssim_{M,\widetilde{M},\mathcal{Q}} e_k + \varepsilon^{\delta_4}\langle 2^{-2k}s^{-1}\rangle^{-\delta_4}\langle 2^{2k}s\rangle^{-10}c_k, \tag{5-75}$$

$$\|P_k\Box(A(s)-\tilde{A})\|_{\Box\underline{X}^1[I]} + \|P_k\Box(A(s)-\tilde{A})\|_{X^{-1/2+b_1,-b_1}[I]} \lesssim_{M,\widetilde{M},\mathcal{Q}} e_k + \varepsilon^{\delta_4}\langle 2^{-2k}s^{-1}\rangle^{-\delta_4}\langle 2^{2k}s\rangle^{-10}c_k, \tag{5-76}$$

where $e_k = d_k + c_k(c\cdot d)_{\leq k}$.

We now turn to the proof of each theorem.

Proof of Theorem 5.11. In the proof, we omit the dependence of constants on M and $\mathcal{Q}$. We introduce the notation

$$\boldsymbol{A}(t,s) = A(t,s) - e^{s\Delta}A(t).$$

Proof of (1): By (3-2) in Proposition 3.1 (note that $\partial_t A$ here corresponds to B in the proposition) we get

$$\|\nabla P_k\boldsymbol{A}(t,s)\|_{L^2[I]} \lesssim \langle 2^{-2k}s^{-1}\rangle^{-\delta_1}\langle 2^{2k}s\rangle^{-10}(c_k^{(0)})^2. \tag{5-77}$$

Now the second bound follows from (3-18) for $\boldsymbol{DA}^3$ and Proposition 4.13 for $Q(A,A)$.

Proof of (2): We proceed in several substeps.

Step 2.1: Our first (and main) goal is to prove

$$\|P_k\boldsymbol{A}(s)\|_{S^1[I]} \lesssim \langle 2^{-2k}s^{-1}\rangle^{-c\delta_3}\langle 2^{2k}s\rangle^{-10}c_k^2. \tag{5-78}$$

We begin by invoking (3-4) with $(\sigma,p) = \left(\frac{1}{4},4\right)$ and $(\sigma_1,p_1) = \left(\frac{1}{2},2\right)$. Since $S^1[I] \subseteq \mathrm{Str}^1[I] \subseteq L^4\dot{W}^{1/4,4}[I]$, we also obtain (after taking $L^2_t[I]$)

$$\|\nabla P_k\boldsymbol{A}(s)\|_{L^2\dot{H}^{1/2}[I]} \lesssim \langle 2^{-2k}s^{-1}\rangle^{-\delta_1}\langle 2^{2k}s\rangle^{-10}c_k^2. \tag{5-79}$$

In view of the embedding $P_k L^2 \dot{H}^{1/2}[I] \subseteq P_k X_1^{0,1/2}[I] \subseteq 2^{-k} S_k[I]$, we have

$$\|\nabla P_k \boldsymbol{A}(s)\|_{S_k[I]} \lesssim \langle 2^{-2k} s\rangle^{-\delta_1} \langle 2^{2k}\rangle^{-10} c_k^2. \tag{5-80}$$

To complete the proof of (5-78), it only remains to establish (recall (4-2))

$$\|\Box P_k \boldsymbol{A}(s)\|_{L^2 \dot{H}^{-1/2}[I]} \lesssim \langle 2^{-2k} s\rangle^{-\delta_1} \langle 2^{2k}\rangle^{-10} c_k^2. \tag{5-81}$$

We argue differently depending on whether $s2^{2k} \gtrsim 1$ or $s2^{2k} \ll 1$. In the former case, we consider $e^{s\Delta} A$ and $A(s)$ separately. In view of (5-7), note that

$$\|\Box P_k e^{s\Delta} A\|_{L^2 \dot{H}^{-1/2}[I]} \lesssim \langle 2^{2k}\rangle^{-10} c_k^2,$$

so it suffices to prove

$$\|\Box P_k A(s)\|_{L^2 \dot{H}^{-1/2}[I]} \lesssim \langle 2^{2k}\rangle^{-10} c_k^2.$$

For this, we need to use the wave equation for $A(s)$ (see Theorem 3.6):

$$\Box A(s) = (\Box - \Box_{A(s)}) A(s) + \mathcal{M}^2(A(s), A(s)) + R_j(A(s)) + \boldsymbol{P}\boldsymbol{w}_x^2(A, A, s) + R_{j;s}(A). \tag{5-82}$$

As in the proof of Proposition 5.4, we note that $\Box - \Box_{A(s)}$ contains the terms $A_0(s)$, $\partial^\ell A(s)$ and $\partial_0 A_0(s)$ that are in turn determined by A, $A(s)$ (see Theorem 3.6). By (5-80) and an obvious bound for $e^{s\Delta} A$, we see that $\langle 2^{2k} s\rangle^{-10} c_k$ is a frequency envelope for $A(s)$ in $\mathrm{Str}^1[I]$. The desired estimate is proved by applying the $L^2 L^2$-type estimates in Section 4 (observe that they only involve the Str^1-norm of A!) and Theorem 3.6.

In the case $s2^{2k} \ll 1$, we begin by writing $\boldsymbol{A}(s) = (A(s) - A) + (1 - e^{s\Delta}) A$. For the second term, again by (5-7), we have

$$\|\Box P_k (1 - e^{s\Delta}) A\|_{L^2 \dot{H}^{-1/2}[I]} \lesssim \langle 2^{-2k} s^{-1}\rangle^{-\delta_1} c_k^2.$$

Thus, for $s2^{2k} \ll 1$, it suffices to establish

$$\|\Box P_k (A(s) - A)\|_{L^2 \dot{H}^{-1/2}[I]} \lesssim \langle 2^{-2k} s^{-1}\rangle^{-c\delta_3} c_k^2. \tag{5-83}$$

Here, we use the equation $\Box(A(s) - A)$ obtained by taking the difference of the equations in Theorems 3.5 and 3.6:

$$\begin{aligned}\Box(A(s) - A) = (\Box - \Box_{A(s)}) A(s) - (\Box - \Box_A) A + \mathcal{M}^2(A(s), A(s)) - \mathcal{M}^2(A, A)& \\ + R_j(A(s)) - R_j(A) + \boldsymbol{P}_j \boldsymbol{w}_x^2(A, A, s) + R_{j;s}(A).&\end{aligned} \tag{5-84}$$

We note that $(\Box - \Box_{A(s)}) A(s) - (\Box - \Box_A) A$ contains the differences $A_0(s) - A_0$, $\partial_\ell^\ell A(s) - \partial^\ell A_\ell$ and $\partial_0 A_0(s) - \partial_0 A_0$, for which similar difference equations may be derived from Theorems 3.5 and 3.6.

As before, c_k is a δ_5-frequency envelope for A and $A(s)$ in $\mathrm{Str}^1[I]$, whereas $d_k = \langle 2^{-2k} s^{-1}\rangle^{-c\delta_3} c_k$ is a δ_3-frequency envelope for $A(s) - A$ in $\mathrm{Str}^1[I]$ by (5-80) and an obvious bound for $(1 - e^{s\Delta}) A$. Hence the difference envelope e_k in Theorem 3.5 obeys the bound

$$e_k = d_k + c_k (c \cdot d)_{\leq k} \lesssim \langle 2^{-2k} s^{-1}\rangle^{-c\delta_3} c_k.$$

The desired estimate (5-83) is proved by applying the L^2L^2-type estimates in Section 4 (again, they only involve the Str^1-norms of ∇A, $\nabla A(s)$ and $\nabla(A(s)-A)$) and Theorem 3.6.

Step 2.2: To complete the proof, it remains to show that (5-78) implies (5-63)–(5-65). This is proved in a completely analogous way to Proposition 5.4(1), replacing Theorem 3.5 by Theorem 3.6 (where we use Propositions 4.16 and 4.17 for $\boldsymbol{w}_0$ and $\boldsymbol{w}_x$, respectively).

Proof of (3): This is analogous to the proof of Proposition 5.4(1). The only difference in the analysis arises from the extra terms

(i) $\boldsymbol{P}_j w_x^2(\partial_t A, \partial_t A, s) + R_{j;s}(A)$ in $\Box_{A(s)} A(s)$,

(ii) $\boldsymbol{A}_{0;s} = \Delta^{-1} \boldsymbol{w}_0^2(A, A, s) + \boldsymbol{A}^3_{0;s}(A)$ in $A_0(s)$,

(iii) $\boldsymbol{D}\boldsymbol{A}_{0;s}(A)$ in $\partial_t A_0(s)$.

For the first term in (5-75) we need to estimate

$$\||D|^{-1} \boldsymbol{w}_0^2(A, A, s)\|_Y + \||D| \boldsymbol{A}^3_{0;s}(A)\|_Y + \|\boldsymbol{D}\boldsymbol{A}_{0;s}(A)\|_Y.$$

The last two terms are estimated directly using (3-36) and (3-37) and Bernstein's inequality. The first term is estimated via (4-54).

For the extra gain when $s^{1/2} > |I|$ we rebalance by using Hölder in time t and Bernstein in x. Because of this, in that range it suffices to use $L^\infty L^2$ bounds instead of Y, and thus rely instead on (3-33) and (3-34), and (4-52).

For the second term in (5-75) we follow the computation for $\partial_t \boldsymbol{P}^\perp A(s)$ in the proof of Proposition 5.4. The extra contributions there are

$$\Delta^{-1}\partial_j(\partial^\ell [A_\ell(s), \boldsymbol{A}_{0;s}] + [A^\ell(s), \partial_\ell \boldsymbol{A}_{0;s}] + [A^\ell, [A_\ell, \boldsymbol{A}_{0;s}]]).$$

For these it suffices to use (4-53) and (3-36) for long intervals I, and (4-52) and (4-52) and (3-33) for short intervals.

Finally, for the two terms in (5-76) we need to bound

$$\|\boldsymbol{P}_j w_x^2(\partial_t A, \partial_t A, s)\|_{\Box \underline{X}^1 \cap X^{-1/2+b+1,-b_1}} + \|R_{j;s}(A)\|_{\Box \underline{X}^1 \cap X^{-1/2+b+1,-b_1}}.$$

For this it suffices to use the bounds (4-58) and (3-35) in the range $|I| > s^{1/2}$, and (4-56) and (3-32) in the range $|I| \le s^{1/2}$. □

Proof of Theorem 5.13. As before, we omit the dependence of constants on M and Q.

Proof of (1) and (2): The three bounds follow directly from Proposition 3.2, precisely in order from the estimates (3-8), (3-9) and (3-7).

Proof of (3): We repeat the arguments in the proof of Theorem 5.11(2). The bound (5-79) for $P_k \boldsymbol{A}(s)$ goes through the Str^1 norm, so by the same proof we also obtain for $k \ge 0$

$$\|\nabla P_k \boldsymbol{A}(s)\|_{L^2 \dot{H}^{1/2}[I]} \lesssim \langle 2^{-2k} s^{-1}\rangle^{-c\delta_3} \langle 2^{2k} s\rangle^{-10} \varepsilon^{\delta_2} c_k. \tag{5-85}$$

On the other hand for $k \le 0$ we can use (5-69) and Hölder's inequality in time to gain smallness.

Similarly, the bound (5-81) also uses only Str^1 norms so it can be replaced by

$$\|\Box P_k \boldsymbol{A}(s)\|_{L^2 \dot{H}^{-1/2}[I]} \lesssim \langle 2^{-2k} s^{-1}\rangle^{-c\delta_3} \langle 2^{2k}\rangle^{-10} \varepsilon^{\delta_2} c_k \tag{5-86}$$

for $k \geq 0$. Again for $k \leq 0$ we can use a simpler $L^\infty \dot{H}^{-1}$ bound and then Hölder's inequality in time. Together, the bounds (5-85) and (5-86) imply (5-72).

Finally, it remains to establish (5-73) and (5-74). Here the same considerations as in the proof of (5-47) apply, but using Theorem 3.6 instead of Theorem 3.5, as well as Proposition 4.16.

Proof of (4): This repeats the proof of Theorem 5.11(3), but taking advantage of the Str^1 norm in estimating $\boldsymbol{A}^3_{0;s}$ and $\boldsymbol{DA}_{0;s}$ and using (4-55) instead of (4-54). As before, the ε gain is due to energy dispersion if $k \geq 0$ and to the interval size otherwise. □

6. Energy-dispersed caloric Yang–Mills waves

The goal of this section is to prove the following key theorem for energy-dispersed subthreshold caloric Yang–Mills waves, which is essentially a restatement of Theorem 1.20 in terms of the linear energy:

Theorem 6.1. *There exist nondecreasing positive functions $M(E, \mathcal{Q})$ and nonincreasing positive functions $\varepsilon(E, \mathcal{Q})$ and $T(E, \mathcal{Q})$ so that the following holds. Let A be a regular caloric Yang–Mills wave on a time interval I satisfying*

$$\inf_{t\in I} \|\nabla A(t)\|^2_{L^2} \leq E, \quad A(t) \in \mathcal{C}_\mathcal{Q} \text{ for all } t \in I. \tag{6-1}$$

If A moreover obeys the smallness bounds

$$\|F\|_{\mathrm{ED}_{\geq m}[I]} \leq \varepsilon(E, \mathcal{Q}), \quad |I| \leq 2^{-m} T(E, \mathcal{Q}), \tag{6-2}$$

then we have

$$\|A\|_{S^1[I]} \leq M(E, \mathcal{Q}). \tag{6-3}$$

We next show that Theorem 1.16 immediately follows. Indeed, for caloric waves we have (see Theorem 1.6)

$$\|\nabla A\|_{L^2} \lesssim_{\mathcal{E},\mathcal{Q}} 1,$$

as well as

$$\mathcal{E} \lesssim_{\|\nabla A\|_{L^2}} 1.$$

Thus the linear and nonlinear energy are interchangeable in the statement of the theorem. The (minor) difference is that the nonlinear energy is exactly conserved, whereas the linear energy is only approximately conserved for energy-dispersed Yang–Mills waves; see Theorem 5.9(5).

For the remainder of this section, we fix $\mathcal{Q}$. We omit any dependence of constants on $\mathcal{Q}$ and write $\varepsilon(E) = \varepsilon(E, \mathcal{Q})$, $T(E) = T(E, \mathcal{Q})$, $M = M(E, \mathcal{Q})$ etc.

Theorem 6.1 is proved by an induction-on-energy argument of similar structure to [Sterbenz and Tataru 2010a; Oh and Tataru 2018]. For the initial step, we show that it holds for small E (Proposition 6.2). For the induction step, we assume that the result holds for all solutions with $\inf_I E_{\mathrm{lin}}(A) \leq E$, and we seek to show that it holds up to $\inf_I E_{\mathrm{lin}}(A) \leq E + c(E)$ for some small $c(E) > 0$. Notably, in order to continue the induction argument, we do not want $c(E)$ to depend on $F(E)$ or $\varepsilon(E)$.

6A. ***Induction on energy argument.*** As remarked earlier, the initial step of the proof of Theorem 6.1 is essentially small-energy global regularity for the Yang–Mills equation in the caloric gauge, which is a quick consequence of Theorem 5.1.

Proposition 6.2. *There exists a small universal constant $E_* > 0$ (in particular, independent of I) such that if a classical caloric Yang–Mills connection satisfies*

$$\inf_{t \in I} \|\nabla A(t)\|_{L^2}^2 \leq E_*, \tag{6-4}$$

then we have

$$\|A\|_{S^1[I]} \lesssim \sqrt{E_*}. \tag{6-5}$$

Proof. We will follow a standard continuity argument, similar to the one used in the Coulomb gauge in [Krieger and Tataru 2017]. Start from a near minimum t_0 for $\|\nabla A(t)\|_{L^2}^2$. Denote by c a frequency envelope for $A[t_0]$ in $\dot{H}^1 \times L^2$. For a short time, there exists a classical solution, which satisfies

$$\|A\|_{S^1[I]} \lesssim E_*.$$

We now consider the maximal interval I containing t_0 and where the solution A exists as a classical solution and satisfies

$$\|A\|_{S^1[I]} \leq 1. \tag{6-6}$$

This in particular implies

$$Q(A) \lesssim 1.$$

Hence by Theorem 5.1(2) it follows that

$$\|A\|_{S^1_c[I]} \lesssim 1,$$

and in particular

$$\|A\|_{S^1[I]} \lesssim E_*. \tag{6-7}$$

Assume now by contradiction that I has a finite end T. The S^1 (6-6) bound implies that A is uniformly bounded near $t = T$ and has a limit as a classical solution. Hence it can be extended further as a classical solution (for a precise statement, see in particular Theorem 7.6). However, in view of (6-7), if E_* is sufficiently small then by continuity we can find a larger interval $I \subsetneq J$ where (6-6) holds. This is a contradiction. It follows that the solution A is global and satisfies (6-7). □

For the induction step, consider a regular caloric Yang–Mills wave A on I such that

$$E < \inf_{t \in I} \|\nabla A(t)\|_{L^2}^2 \leq E + c(E), \quad \|F\|_{\mathrm{ED}_{\geq 0}(I)} \leq \varepsilon, \quad |I| \leq T. \tag{6-8}$$

Our goal is to establish a uniform bound

$$\|A\|_{S^1[I]} \leq M \tag{6-9}$$

for appropriately chosen $c(E) > 0$ (depending *only* on E), ε, T and M (which may depend on E, $\varepsilon(E)$, $T(E)$, $M(E)$ and $c(E)$).

Once this goal is achieved, we may extend $M(E)$, $\varepsilon(E)$ and $T(E)$ to $[0, E + c(E)]$ so that $M(E + c(E)) = M$, $\varepsilon(E + c(E)) = \varepsilon$ and $T(E + c(E)) = T$, while keeping validity of Theorem 6.1

in this range of energy. Since $c(E)$ is a positive number depending only on E, this procedure can be continued until Theorem 6.1 holds for all regular subthreshold caloric Yang–Mills waves.

We now turn to the proof of (6-9). By translating and reversing t, we may assume without any loss of generality that $I = [0, T_+)$ for some $T_+ > 0$ and

$$E < \|\nabla A(0)\|_{L^2}^2 \le E + 2c(E).$$

Since A is regular, it can be easily seen that $\|A\|_{S^1[0,T)}$ is a continuous function of T satisfying

$$\limsup_{T\to 0+} \|A\|_{S^1[0,T)} \lesssim \|\nabla A(t)\|_{L^2} \lesssim E^{\frac{1}{2}}.$$

Therefore, on a subinterval $J = [0, T) \subseteq I$, we may make the bootstrap assumption

$$\|A\|_{S^1[J]} \le 2M. \tag{6-10}$$

In order to improve (6-10) to (6-9), we compare A with a caloric Yang–Mills wave $\tilde{A}$ with $S^1[I]$-norm $\le M(E)$ (eventually), which we construct as follows.

To begin with, we view the space-time connection $A_{t,x}$ on $I \times \mathbb{R}^4$ as a caloric initial data and solve the dynamic Yang–Mills heat flow in the local caloric gauge, i.e.,

$$\partial_s A_\mu(t,x,s) = \boldsymbol{D}^k F_{k\mu}(t,x,s),$$
$$A_\mu(t,x,0) = A_\mu(t,x).$$

From the results in Section 3, we obtain a global-in-heat-time solution $A_{t,x}(t,x,s)$ on $I \times \mathbb{R}^4 \times [0,\infty)$. Note that $\partial_t A$ solves the linearized Yang–Mills heat flow in local caloric gauge, and we have $(A, \partial_t A)(t,s) \in T^{L^2}\mathcal{C}$ for every $(t,s) \in I \times [0,\infty)$.

By the caloric gauge condition, the linear energy

$$\|(A, \partial_t A)(t,s)\|_{\dot{H}^1 \times L^2}^2 = \|\nabla A(t,s)\|_{L^2}^2$$

eventually tends to zero as $s \to \infty$. Thus there exists a heat-time $s'_* > 0$ such that

$$\|(A, \partial_t A)(0,s)\|_{\dot{H}^1 \times L^2}^2 = E.$$

To eliminate ambiguity, we take s'_* to be the minimum such heat-time. In order to choose the cut-off heat-time s_*, we distinguish two scenarios:

(1) If $s'_* \ge 1$, then we define $s_* = 1$.

(2) If $s'_* < 1$, then we define $s_* = s'_*$.

With s_* chosen as above, we define $\tilde{A}$ to be the caloric Yang–Mills wave with initial data

$$(\tilde{A}, \partial_t \tilde{A})(0) = (A, \partial_t A)(0, s_*).$$

In both scenarios, we aim to prove that $\tilde{A}$ exists on J and is well-approximated by $A(s_*)$. Moreover, by the induction hypothesis, $\tilde{A}$ should obey a nice S^1-norm bound.

Proposition 6.3. *Let $\tilde{A}$ be defined as above. For sufficiently small $\varepsilon, T > 0$ depending on M, $M(E)$, $T(E)$, $\varepsilon(E)$ and $c(E)$, the regular caloric Yang–Mills wave $\tilde{A}$ exists on the interval J and obeys*

$$\|\tilde{A}\|_{S^1[J]} \leq M(E) + C_0\sqrt{E}, \tag{6-11}$$

$$\|A(s_*) - \tilde{A}\|_{\underline{S}^1_{c^*}[J]} \lesssim_M \varepsilon^{\delta_6}, \tag{6-12}$$

$$\|A_0(s_*) - \tilde{A}_0\|_{Y^1_{c^*}[J]} \lesssim_M \varepsilon^{\delta_6}, \tag{6-13}$$

$$\|\boldsymbol{P}^\perp A(s_*) - \boldsymbol{P}^\perp \tilde{A}\|_{Y^1_{c^*}[J]} \lesssim_M \varepsilon^{\delta_6}, \tag{6-14}$$

where C_0 is a universal constant and c^* is a frequency envelope defined as

$$c^*_k = 2^{-\delta_*|k-k(s_*)|}. \tag{6-15}$$

On the other hand, viewing A as a "high frequency perturbation" of $\tilde{A}$, we show below that A stays close to $\tilde{A}$ in the space S^1.

Proposition 6.4. *Let $\tilde{A}$ be defined as above on the interval J. Provided that $c = c(E) > 0$ is chosen small enough compared to E (but independent of $M(E)$, $T(E)$ or $\varepsilon(E)$) and $T, \varepsilon > 0$ are also sufficiently small depending on M, $M(E)$, $T(E)$, $\varepsilon(E)$ and $c(E)$, we have*

$$\|A - \tilde{A}\|_{S^1[J]} \lesssim_{M(E),E} 1. \tag{6-16}$$

Assuming the preceding two propositions, we may choose M sufficiently large compared to $M(E)$ and E, then choose ε and T accordingly, so that the desired estimate (6-9) follows from (6-11) and (6-16).

It remains to prove Propositions 6.3 and 6.4, which are the subjects of Sections 6B and 6C, respectively.

6B. *Control of $\tilde{A} - A(s_*)$: proof of Proposition 6.3.* We introduce the notation

$$\delta A^{\text{low}} = \tilde{A} - A(s_*). \tag{6-17}$$

We proceed differently depending on how s_* was chosen.

<u>Scenario 1</u>: $s_* = 1 (\leq s'_*)$. This scenario is simpler to handle, and we do not need to invoke the induction hypothesis.

Step 1.1: S^1-norm bound for $\tilde{A}$. We first prove the S^1-norm bound (6-11). The idea is to exploit the smoothing property of the Yang–Mills heat flow, which implies control of higher Sobolev norms of $(\tilde{A}, \partial_t \tilde{A})(0) = (A, \partial_t A)(0, 1)$ in terms of $\sqrt{E}$, and use subcritical local regularity of Yang–Mills in the caloric gauge, which works in a time interval of length $O_E(1)$.

Fix a large integer N (say $N = 10$). We claim that $\tilde{A}$ exists on J and

$$\|\tilde{A}\|_{S^N \cap S^1[J]} \lesssim \sqrt{E}, \tag{6-18}$$

provided that T is sufficiently small depending only on E (so that $|J| \ll_E 1$).

By the smoothing property for the Yang–Mills heat flow and its linearization in the caloric gauge (see Section 3), we have

$$\|(\tilde{A}, \partial_t \tilde{A})(0)\|_{(\dot{H}^N \times \dot{H}^{N-1}) \cap (\dot{H}^1 \times L^2)} \lesssim \sqrt{E}.$$

For T sufficiently small (depending only on E), the following local-in-time a priori estimates at subcritical regularity hold:

$$\sup_{t\in J}\|(\tilde{A}, \partial_t \tilde{A})(t)\|_{(\dot{H}^N\times\dot{H}^{N-1})\cap(\dot{H}^1\times L^2)} + |J|\|\Box\tilde{A}\|_{L^\infty(\dot{H}^{N-1}\cap L^2)[J]} \lesssim \sqrt{E},$$
$$\sup_{t\in J}\|(\tilde{A}_0, \partial_t \tilde{A}_0)(t)\|_{(\dot{H}^N\times\dot{H}^{N-1})\cap(\dot{H}^1\times L^2)} \lesssim \sqrt{E}.$$

The proof is via Theorem 3.5 and, as usual, the Sobolev embedding into L^∞; we omit the details.

As a consequence of the preceding a priori bounds, we obtain (6-18) as desired. Moreover, by Theorem 3.5 and the fixed-time bounds in Section 4, we have

$$\|\Box\tilde{A}\|_{L^\infty\dot{H}^{-1}[J]} \lesssim_E 1. \tag{6-19}$$

Step 1.2: S^1-norm bound for $A(s_*) - \tilde{A}$. As a preparation for the proof of (6-12), we claim that

$$\|A(s_*) - \tilde{A}\|_{S^1_{c^*}[J]} \lesssim_M \varepsilon^c. \tag{6-20}$$

In the present case, $2^{k(s_*)} = 1$. For frequencies higher than 1, we simply use (6-18) with smoothing estimates for $A(s_*)$ in S^1. For frequencies lower than 1, we control $\Box(\tilde{A} - A(s_*))$ in $L^\infty\dot{H}^{-1}$ and integrate in time.

By Theorem 5.11, we have

$$\|P_k A(s_*)\|_{S^1[J]} \lesssim_M 2^{-20k_+}, \tag{6-21}$$
$$\|P_k \Box A(t,s)\|_{\dot{H}^{-1}} \lesssim_M 2^{-20k_+}. \tag{6-22}$$

Let $\kappa_0 \geq k(s_*)$ be a parameter to be fixed below. By (6-20) and (6-21), we have

$$\|P_k \delta A^{\mathrm{low}}\|_{S^1[J]} \leq \|P_k\tilde{A}\|_{S^1[J]} + \|P_k A(s_*)\|_{S^1[J]} \lesssim_M 2^{-c\kappa_0}c^*_k \quad \text{for } k \geq \kappa_0, \tag{6-23}$$

where $0 < c \ll 1$ is a universal constant. Since

$$P_k(L^\infty\dot{H}^{-1}[J]) \hookrightarrow |J|2^k N \cap (|J|2^k)^{\frac{1}{2}} L^2\dot{H}^{-\frac{1}{2}},$$

for $k \leq \kappa_0$ it follows from (6-19) and (6-22) that

$$\begin{aligned}\|P_k\Box\delta A^{\mathrm{low}}\|_{(N\cap L^2\dot{H}^{-1/2})[J]} &\leq \|P_k\Box\tilde{A}\|_{(N\cap L^2\dot{H}^{-1/2})[J]} + \|P_k\Box A(s_*)\|_{(N\cap L^2\dot{H}^{-1/2})[J]}\\ &\lesssim_M ((|J|2^{\kappa_0})^{1/2} + (|J|2^{\kappa_0}) + \varepsilon^c)c^*_k.\end{aligned}$$

Since $\delta A^{\mathrm{low}}[0] = 0$, we arrive at

$$\|P_k \delta A^{\mathrm{low}}\|_{S^1[J]} \lesssim_M ((|J|2^{\kappa_0})^{\frac{1}{2}} + (|J|2^{\kappa_0}) + \varepsilon^c)c^*_k \quad \text{for } k \leq \kappa_0. \tag{6-24}$$

Step 1.3: completion of proof. Finally, the bounds (6-12)–(6-14) follow from (6-20) and Theorem 5.11(3) with $d_k = c^*_k$ provided that $|J| \leq T$ is sufficiently small. Here, note that

$$e_k = c^*_k + c_k(c\cdot c^*)_{\leq k} \lesssim_M c^*_k.$$

Scenario 2: $s_* = s'_* > 1$. In the second scenario, we analyze the equation satisfied by the difference $\delta A^{\text{low}} = A(s_*) - \tilde{A}$ to prove (6-12), then make use of the induction hypothesis to derive (6-11). By another continuous induction in time, we may make the following extra bootstrap assumptions:

$$\|\tilde{A}\|_{S^1[J]} \leq 2(M(E) + C_0\sqrt{E}), \tag{6-25}$$

as well as

$$\|\delta A^{\text{low}}\|_{S^1_{c*}[J]} \leq \varepsilon^{c\delta_6}. \tag{6-26}$$

Here we use a smaller power of ε, so this last bound will only serve to ensure some a priori smallness of δA^{low} in S^1_{c*}.

By Theorem 5.13, we have

$$\|P_k A(s_*)\|_{S^1[J]} \lesssim_M c_k \langle 2^{2k} s_* \rangle^{-10}, \tag{6-27}$$

$$\|A(s_*)\|_{\text{ED}^1_{\geq 0}[J]} \lesssim_E \varepsilon^{\delta_4}, \tag{6-28}$$

$$\|\Box A(s_*)\|_{L^2 \dot{H}^{-1/2}[J]} \lesssim_M \varepsilon^{\delta_4}. \tag{6-29}$$

Therefore, $(A(s_*), J)$ is (ε, M_*)-energy-dispersed for $M_* \lesssim_M 1$ and $\varepsilon \leq \varepsilon^{\delta_4}$.

Step 2.1: bounds for δA^{low}. Here we establish (6-12). We write an equation for δA^{low} of the form

$$\Box_{\tilde{A}} \delta A^{\text{low}} = F, \quad \delta A^{\text{low}}[0] = 0.$$

We claim that in each subinterval J_1 of J and for each $\kappa > 10$ we have the bound

$$\|F\|_{(N \cap L^2 \dot{H}^{-1/2})_{c*}[J_1]} \lesssim_M (2^{-c\delta_* \kappa} \|\tilde{A}\|_{S^1[J_1]} + 2^{C\kappa} C(\tilde{A}, J_1)) \|\delta A^{\text{low}}\|_{S^1_{c*}[J_1]} + \varepsilon^{\delta_6}, \tag{6-30}$$

where $C(\tilde{A}, J_1)$ contains only divisible norms of $\tilde{A}$; see (5-21).

We first verify that the bound (6-30) implies (6-12). Using the well-posedness for the $\Box_{\tilde{A}}$ equation, given by Theorem 5.1, in the time interval $J_1 = [t_1, t_2]$, we obtain the bound

$$\|\delta A^{\text{low}}\|_{S^1_{c*}[J_1]} \leq C(M)\big(\|\delta A^{\text{low}}[t_1]\|_{\mathcal{H}_{c*}} + (2^{-c\delta_* \kappa} \|\tilde{A}\|_{S^1[J_1]} + 2^{C\kappa} C(\tilde{A}, J_1)) \|\delta A^{\text{low}}\|_{S^1_{c*}[J_1]} + \varepsilon^{\delta_6}\big).$$

For this to be useful we need to ensure that the coefficient of $\|\delta A^{\text{low}}\|_{S^1_{c*}[J_1]}$ on the right is small. To achieve that we first choose κ large enough, $k \gg_M 1$, depending only on M, so that

$$C(M) 2^{-c\delta_* \kappa} \|\tilde{A}\|_{S^1[J]} \ll 1.$$

Then we divide the interval J into subintervals J_j so that

$$C(M) 2^{C\kappa} C(\tilde{A}, J_j) \ll 1.$$

The number of such intervals depends only on M. On each subinterval $J_j = [t_{j-1}, t_j]$ we have the bound

$$\|\delta A^{\text{low}}\|_{S^1_{c*}[J_1]} + \|\delta A^{\text{low}}[t_j]\|_{\mathcal{H}_{c*}} \leq C(M)(\|\delta A^{\text{low}}[t_{j-1}]\|_{\mathcal{H}_{c*}} + \varepsilon^{\delta_6}).$$

Reiterating this we obtain (6-12).

If remains to prove the bound (6-30). We relabel J_1 by J for simplicity. As a preliminary step, we observe that, by Theorem 5.13 and the bootstrap assumption (6-26), we have

$$\|\delta A^{\mathrm{low}}\|_{\underline{S}^1_{c^*}[J]} + \|\delta A_0^{\mathrm{low}}\|_{Y^1_{c^*}[J]} + \|\boldsymbol{P}^{\perp}\delta A^{\mathrm{low}}\|_{Y^1_{c^*}[J]} \lesssim_M \|\delta A^{\mathrm{low}}\|_{S^1_{c^*}[J]}. \tag{6-31}$$

In particular, this proves the bounds (6-13) and (6-14) once (6-12) is known.

The expression for F is obtained from Theorems 3.5 and 3.6,

$$F := \Box_{\tilde{A}}\delta A^{\mathrm{low}} = \Box_{\tilde{A}}\tilde{A} - \Box_{A(s_*)}A(s_*) + (\Box_{A(s_*)} - \Box_{\tilde{A}})A(s_*),$$

where we further expand the two terms as

$$\Box_{\tilde{A}}\tilde{A} - \Box_{A(s_*)}A(s_*) = \mathcal{M}^2(\tilde{A},\tilde{A}) - \mathcal{M}^2(A(s_*),A(s_*)) + R(\tilde{A}) - R(A(s_*)) \\ + \boldsymbol{P}w_x^2(\partial_t A,\partial_t A,s) + R_{j;s}(A),$$

and

$$(\Box_{A(s_*)} - \Box_{\tilde{A}})A(s_*) = -\mathrm{Diff}^{\kappa}_{\boldsymbol{P}\delta A^{\mathrm{low}}}A(s_*) - \mathrm{Diff}^{\kappa}_{\boldsymbol{P}^{\perp}\delta A^{\mathrm{low}}}A(s_*) - \mathrm{Rem}^{\kappa,2}_{\delta A^{\mathrm{low}}}A(s_*) \\ + (\mathrm{Rem}^3(A(s_*)) - \mathrm{Rem}^3(\tilde{A}))A(s_*) + \mathrm{Rem}^3_{s_*}(A)A(s_*).$$

We successively estimate the terms above as in (6-30):

(a) For $\mathcal{M}^2(\tilde{A},\tilde{A}) - \mathcal{M}^2(A(s_*),A(s_*))$ we use the estimate (4-50). We inherit the envelope c_* from δA^{low} but we also gain an additional power of ε from the energy dispersion of $A(s_*)$.

(b) For $R(\tilde{A}) - R(A(s_*))$ we use the difference version of the bound (3-21), with a similar gain.

(c) For $\boldsymbol{P}w_x^2(\partial_t A,\partial_t A,s)$ we use (4-59), taking advantage of the energy dispersion for A.

(d) For $R_{j;s}(A)$ we use (3-35), gaining a power of ε from the Str^1 norm.

(e) For $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}^{\perp}\delta A^{\mathrm{low}}}A(s_*)$ we use (4-82) combined with (6-31) for the high modulations, and (4-83) combined with (4-37) and (3-22) for low modulations.

(f) For $\mathrm{Rem}^{\kappa,2}_{\delta A^{\mathrm{low}}}A(s_*)$ we use (4-81).

(g) For $(\mathrm{Rem}^3(A(s_*)) - \mathrm{Rem}^3(\tilde{A}))A(s_*)$ we use (4-74).

(h) For $\mathrm{Rem}^3_{s_*}(A)A(s_*)$ we use (4-76).

This leaves us with the most difficult term $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}\delta A^{\mathrm{low}}}A(s_*)$, for which we claim that

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}\delta A^{\mathrm{low}}}A(s_*)\|_{(N\cap L^2\dot{H}^{-1/2})_{c^*}[J]} \lesssim_M 2^{-c\delta_*\kappa}\|\delta A^{\mathrm{low}}\|_{S^1[J]}. \tag{6-32}$$

For $\boldsymbol{P}\delta A^{\mathrm{low}}$ we consider the same type of decomposition as in the proof of Lemma 5.8,

$$\boldsymbol{P}\delta A^{\mathrm{low}} = \boldsymbol{P}\delta A^{\mathrm{low,main}} + \boldsymbol{P}\delta A^{\mathrm{low,main,2}} + \boldsymbol{P}\delta A^{\mathrm{low,rem,2}} + \boldsymbol{P}\delta A^{\mathrm{low,rem,3}},$$

where

$$
\begin{aligned}
\delta A_0^{\mathrm{low,main}} &= \Delta^{-1}([\tilde{A}, \partial_t \tilde{A}] - [A(s_*), \partial_t A(s_*)]),\\
\delta A_0^{\mathrm{low,main,2}} &= \Delta^{-1} \boldsymbol{w}_0(A, A, s),\\
\delta A_0^{\mathrm{low,rem,2}} &= 2\Delta^{-1}(\boldsymbol{Q}(\tilde{A}, \partial_t \tilde{A}) - \boldsymbol{Q}(A(s_*), \partial_t A(s_*))),\\
\delta A_0^{\mathrm{low,rem,3}} &= A_0^3(\tilde{A}, \partial_t \tilde{A}) - A_0^3(A(s_*), \partial_t A(s_*)) + A_{0;s}^3(A, \partial_t A),
\end{aligned}
$$

and

$$
\begin{aligned}
\delta A_x^{\mathrm{low,main}} &= \Box^{-1}(\boldsymbol{P}\mathcal{M}^2(\tilde{A}, \tilde{A}) - \boldsymbol{P}\mathcal{M}^2(A(s_*), A(s_*))),\\
\delta A_x^{\mathrm{low,main,2}} &= \Box^{-1} \boldsymbol{P}\boldsymbol{w}_x(A, A, s),\\
\delta A_x^{\mathrm{low,rem,2}} &= \Box^{-1} \boldsymbol{P}([\tilde{A}_\alpha, \partial^\alpha \tilde{A}] - [A_\alpha(s_*), \partial^\alpha A(s_*)]),\\
\delta A_x^{\mathrm{low,rem,3}} &= \Box^{-1} \boldsymbol{P}(R(\tilde{A}) - R(A(s_*)) - \mathrm{Rem}^3(\tilde{A})\tilde{A} + \mathrm{Rem}^3(A(s_*))A(s_*))\\
&\qquad + \Box^{-1} \boldsymbol{P}(R_{j;s}(A) - \mathrm{Rem}_s^3(A)A(s_*)),
\end{aligned}
$$

where $\Box^{-1}$ is the wave parametrix with zero Cauchy data at $t = 0$.

As a preliminary observation we note that

$$
\|\delta A_x^{\mathrm{low,main}}\|_{S^1_{c^*}} + \|\delta A_x^{\mathrm{low,main,2}}\|_{S^1_{c^*}} + \|\delta A_x^{\mathrm{low,rem,2}}\|_{S^1_{c^*}} + \|\delta A_x^{\mathrm{low,rem,3}}\|_{S^1_{c^*}} \lesssim_M \|\delta A^{\mathrm{low}}\|_{S^1_{c^*}} + \varepsilon^{\delta_2}. \tag{6-33}
$$

This is a consequence of (4-42) for the first term, (4-59) and (5-47) for the second, and (3-21), (3-35), (4-74) and (4-76) for the last term. The bound for the third term follows indirectly since they all add up to δA^{low}.

Now we consider the contributions of each of these terms to $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}\delta A^{\mathrm{low}}} A(s_*)$.

The contributions of $\delta A_x^{\mathrm{low,main}}$ and $\delta A_0^{\mathrm{low,main}}$. These are considered together, and estimated using Proposition 4.27. This yields the frequency envelope

$$
f_k = \Big(\sum_{k' < k - \kappa} c^*_{k'} c_{k'} \langle 2^{2k'} s_* \rangle^{-N} \Big) c_k \langle 2^{2k'} s_* \rangle^{-N} \|\delta A^{\mathrm{low}}\|_{S^1_{c^*}[J]} \lesssim_M 2^{-c\delta_* \kappa} c^*_k \|\delta A^{\mathrm{low}}\|_{S^1_{c^*}[J]},
$$

as needed.

The contributions of $\delta A_x^{\mathrm{low,main,2}}$ and $\delta A_0^{\mathrm{low,main,2}}$. These are also considered together, but now we want to use Proposition 4.29. As they involve no δA^{low} differences, we need to estimate these contributions by ε^{δ_6}. Unfortunately Proposition 4.29 provides no source for an energy dispersion gain, so we use a trick, decomposing

$$
\mathrm{Diff}^{\kappa}_{\delta A^{\mathrm{low,main,2}}} A(s_*) = \mathrm{Diff}^{\kappa'}_{\delta A^{\mathrm{low,main,2}}} A(s_*) + \mathrm{Diff}^{[\kappa', \kappa]}_{\delta A^{\mathrm{low,main,2}}} A(s_*),
$$

where $\kappa' > \kappa$ is a secondary parameter to be chosen shortly. For the first term we apply Proposition 4.29, which yields

$$
\|\mathrm{Diff}^{\kappa'}_{\delta A^{\mathrm{low,main,2}}} A(s_*)\|_{(N \cap L^2 \dot{H}^{-1/2})_{c^*}[J]} \lesssim_M 2^{-c\delta_* \kappa'}.
$$

For the second term, on the other hand, we use instead the bounds (4-55) and (4-59), which capture both the c^* decay and the energy dispersion. The price to pay is that this way we only have access to the

S^1 norm of $\delta A^{\text{low,main},2}$, so we are only allowed to use (4-77). This yields

$$\|\mathrm{Diff}^{[\kappa',\kappa]}_{\delta A^{\text{low,main},2}} A(s_*)\|_{(N\cap L^2\dot{H}^{-1/2})_{c^*}[J]} \lesssim_M \varepsilon^{c\delta_c\delta_g} 2^{C\kappa'}.$$

We now add the last two bounds and then optimize in κ' to obtain the desired estimate

$$\|\mathrm{Diff}^{\kappa}_{\delta A^{\text{low,main},2}} A(s_*)\|_{(N\cap L^2\dot{H}^{-1/2})_{c^*}[J]} \lesssim_M \varepsilon^{\delta_h}.$$

The contribution of $\delta A^{\text{low,rem},2}$. The $\delta A_x^{\text{low,rem},2}$ part is estimated using Proposition 4.28, with (6-33) serving to verify the hypothesis. For the output this yields the frequency envelope

$$f_k = \Big(\sum_{k'<k-\kappa} c^*_{k'}\Big) c_k \langle 2^{2k'} s_*\rangle^{-N} \lesssim_M 2^{-c\delta_*\kappa} c^*_k.$$

A simpler analysis applies for the contribution of $\delta A_0^{\text{low,rem},2}$ where we can use Proposition 4.13.

The contribution of $\delta A^{\text{low,rem},3}$. For the contribution of $\delta A_0^{\text{low,rem},3}$ we use (3-23) and (3-36), while for the contribution of $\delta A_x^{\text{low,rem},3}$ we use (3-21), (3-35), (4-74) and (4-76), all combined with Proposition 4.26.

Step 2.2: S^1-norm bound for $\tilde{A}$ via induction hypothesis. Taking ε sufficiently small and using the bootstrap assumption (6-26), we may ensure that

$$\|\widetilde{F}\|_{ED_{\geq 0}[J]} \leq \varepsilon(E). \tag{6-34}$$

By the induction hypothesis, we may thus assume that

$$\|\tilde{A}\|_{S^1[J]} \leq M(E). \tag{6-35}$$

6C. ***Control of $A-\tilde{A}$: proof of Proposition 6.4.*** Here, we seek to bound

$$\delta A^{\text{high}} = A - \tilde{A}.$$

We begin by observing that

$$\|\tilde{A}\|_{ED^{-1}_{\geq 0}[J]} + \|\Box\tilde{A}\|_{L^2\dot{H}^{-1/2}[J]} \lesssim_M \varepsilon^{\delta_6}.$$

Therefore, both (A, J) and $(\tilde{A}, J)$ are (ε, M)-dispersed, where $\varepsilon \lesssim_M \varepsilon^{\delta_6}$.

Step 1: consequence of approximate linear energy conservation. We claim that

$$\sup_{t\in J}\|(\delta A^{\text{high}}, \partial_t \delta A^{\text{high}})(t)\|^2_{\dot{H}^1\times L^2} \lesssim c(E) + C_M \varepsilon^{\delta_6}. \tag{6-36}$$

Note that

$$\delta A^{\text{high}} = (1-e^{s_*\Delta})A + e^{s_*\Delta}A - A(s_*) + A(s_*) - \tilde{A}.$$

We begin with the inequality

$$\|\nabla A(t)\|^2_{L^2} \geq \|\nabla(1-e^{s_*\Delta})A(t)\|^2_{L^2} + \|e^{s_*\Delta}A(t)\|^2_{L^2},$$

which follows from Plancherel and nonnegativity of the symbol of $(1-e^{s_*\Delta})e^{s_*\Delta}$. By Theorem 5.13(1) and (6-12), we have

$$\|\nabla e^{s_*\Delta} A(t)\|_{L^2}^2 = \|\nabla \tilde{A}(t)\|_{L^2}^2 + C_M \varepsilon^{\delta_6}, \tag{6-37}$$

$$\|\nabla(1-e^{s_*\Delta})A(t)\|_{L^2}^2 = \|\nabla(A-\tilde{A})(t)\|_{L^2}^2 + C_M \varepsilon^{\delta_6}. \tag{6-38}$$

Hence, by Theorem 5.9(5), we have

$$\begin{aligned}\|\nabla(A-\tilde{A})(t)\|_{L^2}^2 &\le \|\nabla A(t)\|_{L^2}^2 - \|\nabla \tilde{A}(t)\|_{L^2}^2 + C_M \varepsilon^{\delta_6}\\ &\le \|\nabla A(0)\|_{L^2}^2 - \|\nabla \tilde{A}(0)\|_{L^2}^2 + C_M \varepsilon^{\delta_6}\\ &\le c(E) + C_M \varepsilon^{\delta_6}.\end{aligned}$$

Step 2: weak divisibility and reinitialization. By Theorem 5.1(7) there exists a partition $J = \bigcup_{k=1}^{K} J_k$ such that $K \lesssim_{M(E)} 1$ and

$$\|\tilde{A}\|_{S^1[J_k]} \lesssim_E 1, \tag{6-39}$$

so that the number of such intervals is also controlled $K \lesssim_{M(E)} 1$. Using the uniform control of the energy of δA^{high} in Step 1, it suffices to estimate δA^{high} in S^1 separately in each of these intervals.

We will make a bootstrap assumption

$$\|\delta A^{\mathrm{high}}\|_{S^1[J_k]} \le 2. \tag{6-40}$$

Then our goal is to improve (6-40) to

$$\|\delta A^{\mathrm{high}}\|_{S^1[J_k]} \le 1 \tag{6-41}$$

by taking $c \ll_E 1$, $\varepsilon \ll_M 1$ and $T \ll_{M,\varepsilon} 1$.

In view of (6-39) and (6-40), in all the estimates below within a single interval J_k, all implicit constants will depend on E rather than $M(E)$. To simplify the notation we drop the subscript and replace J_k by J in what follows.

Step 3: frequency envelope bounds. Let c_k be a frequency envelope for A in $S^1[J]$. Then by Proposition 3.1, the initial data in J_k for $A(s)$ has the frequency envelope $2^{-(k-k^*)_+}c_k$. By Theorem 5.1, we have a similar envelope in S^1,

$$\|P_k \tilde{A}(s)\|_{S^1[J[} \lesssim_E 2^{-(k-k^*)_+} c_k. \tag{6-42}$$

On the other hand, by the estimate (6-12) we have, under the assumption $\varepsilon \ll_E 1$, the bound

$$\|P_k(\tilde{A} - A(s))\|_{S^1[J]} \lesssim_E 2^{-\delta_*|k-k^*|} c_k. \tag{6-43}$$

Hence for the high-frequency difference A^h we have the bound

$$\|P_k \delta A^{\mathrm{high}}\|_{S^1[J]} \lesssim_E 2^{-\delta_*(k-k^*)_-} c_k. \tag{6-44}$$

Step 4: control of nonlinearity. By Theorem 5.9(4) applied separately to A and $\tilde{A}$ we have

$$\|(\Box + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A})\delta A^{\mathrm{high}} + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}\delta A^{\mathrm{high}}} \tilde{A}\|_{N \cap L^2 \dot{H}^{-1/2}[J]} \lesssim_E 2^{C\kappa} \varepsilon^{\delta_4 \delta_6}, \tag{6-45}$$

where the parameter $\kappa \geq 10$ is arbitrary for now, to be chosen later. We claim that the second term can be estimated separately as

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}\delta A^{\mathrm{high}}} \tilde{A}\|_{N \cap L^2 \dot{H}^{-1/2}[J]} \lesssim_E 2^{-c\delta_* \kappa}. \tag{6-46}$$

This is a consequence of Lemma 5.8. To see that we use the bounds (6-42) and (6-44) to compute the frequency envelope f_k in Lemma 5.8. We have

$$f_k \lesssim_E \Bigg(\sum_{k' < k - \kappa} 2^{-c\delta_*(k'-k^*)_-} c_{k'} + 2^{-(k'-k^*)_+} c_{k'} (c^2 c^*)_{<k'} \Bigg) 2^{-(k-k^*)_+} c_k \lesssim_E 2^{-c\delta_* |k-k^*|} c_k,$$

and thus (6-46) follows. Combining (6-45) with (6-46) yields

$$\|(\Box + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}) \delta A^{\mathrm{high}}\|_{N \cap L^2 \dot{H}^{-1/2}[J]} \lesssim_E 2^{-c\delta_* \kappa} + 2^{C\kappa} \varepsilon^{\delta_4 \delta_6}. \tag{6-47}$$

Hence by Theorem 5.1(1) we conclude that

$$\|\delta A^{\mathrm{high}}\|_{S^1[J_k]} \lesssim_E c + 2^{-c\delta_* \kappa} + 2^{C\kappa} \varepsilon^{\delta_4 \delta_6}.$$

Hence by taking $\kappa \gg_E 1$, $c \ll_E 1$, $\varepsilon \ll_{E,\kappa} 1$ and $T \ll_{E,\varepsilon,\kappa} 1$, the desired conclusion (6-41) follows.

7. Proof of the main results

The purpose of this short section is to deduce Theorems 1.13, 1.20 and 1.18 from Theorem 6.1.

7A. *Higher-regularity local well-posedness.* In this subsection, we sketch the proof of higher-regularity local well-posedness of the hyperbolic Yang–Mills equation. We first use the temporal gauge, which works for general connections, and then turn to the caloric gauge, which works for data satisfying (1-12).

7A1. *Temporal gauge.* Here we write the Yang–Mills equations in the temporal gauge,

$$A_0 = 0. \tag{7-1}$$

They take the form

$$\Box_A A_j = \boldsymbol{D}^k \partial_j A_k, \tag{7-2}$$

with the additional constraint equation

$$\boldsymbol{D}^j \partial_0 A_j = 0. \tag{7-3}$$

This can be viewed as a semilinear system of wave equations for the curl of A, coupled with a second-order transport equation for the divergence of A.

We consider the Cauchy problem with initial data

$$A[0] = (A_j(0), \partial_t A_j(0)).$$

The initial data is uniquely determined by the Yang–Mills initial data and the gauge condition (7-1).

The system (7-2) together with the constraint equation (7-3) is well-posed in regular Sobolev spaces. Precisely, we have:

Theorem 7.1. *The system (7-2) is locally well-posed in $H^N \times H^{N-1}$ for $N \geq 2$, with Lipschitz dependence on the initial data.*

We further remark that the temporal gauge fully describes all classical solutions to the Yang–Mills system:

Theorem 7.2. *Let A be a solution to the Yang–Mills system which has local-in-time regularity $(A, \partial_t A) \in C([0,T]; H^N \times H^{N-1})$ for $N \geq 3$. Then A has a temporal gauge equivalent $\tilde{A}$ with the same regularity $(\tilde{A}, \partial_t \tilde{A}) \in C([0,T]; H^N \times H^{N-1})$.*

To see this, it suffices to solve an equation for the gauge transformation O, namely

$$O^{-1}\partial_0 O = A_0, \quad O(0,x) = I,$$

which is an ODE on the Lie group $\boldsymbol{G}$. If $A \in C(H^N)$ then this yields a unique solution $O \in C(H^N)$. This in turn yields a temporal gauge equivalent solution

$$(\tilde{A}, \partial_t \tilde{A}) \in C([0,T]; H^{N-1} \times H^{N-2}).$$

This argument loses one derivative. However, the initial data is in $H^N \times H^{N-1}$, which by the well-posedness result yields a $C([0,T]; H^N \times H^{N-1})$ solution. But by the $H^{N-1} \times H^{N-2}$ well-posedness the two must agree, so we obtain a unique representation in the temporal gauge with the same data and without loss of derivatives.

Remark 7.3. Analogues of Theorems 7.1 and 7.2 hold for the space $H^N_{\mathrm{loc}} \times H^{N-1}_{\mathrm{loc}}$ instead of $H^N \times H^{N-1}$, where H^N_{loc} is equipped with the norm $\sup_{x\in\mathbb{R}^4} \|\cdot\|_{H^N(B_1(x))}$.

7A2. *Caloric gauge.* In view of Theorem 1.11 we can fully describe caloric Yang–Mills waves as continuous functions

$$I \ni t \to (A_x(t), \partial_0 A_x(t)) \in T^{L^2} C.$$

For higher-regularity Yang–Mills waves we have the following:

Theorem 7.4. *Let A be a solution to the Yang–Mills system which has local-in-time regularity $(A, \partial_t A) \in C([0,T]; H^N \times H^{N-1})$ for $N \geq 2$. Assume in addition that the bound (1-12) is uniformly satisfied by its caloric extension, globally in parabolic time. Then A has a caloric gauge equivalent $\tilde{A}$ with the same regularity $(\tilde{A}, \partial_t \tilde{A}) \in C([0,T]; H^N \times H^{N-1})$.*

This result is a direct consequence of Theorem 1.11, with one minor exception. Precisely, Theorem 1.11 does not directly yield the $C_t L^2_x$ regularity for $\partial_0 A_0$. For that we instead need to refer to the expression (3-15) and the bounds (3-18) and (4-28) for the two terms in (3-15).

Remark 7.5. The same result will easily hold for $(A, \partial_t A) \in C([0,T]; \boldsymbol{H} \times L^2)$. However, if we only assume that $(A, \partial_t A) \in C([0,T]; \dot{H}^1 \times L^2)$ then one would also need to resolve the remaining gauge freedom. For that it suffices to observe that if two A's have a small difference in L^2, then the two O's can be chosen in tandem so that they agree at infinity.

In particular this says that a caloric gauge solution exists for as long as a regular solution exists and the L^3 bound in (1-12) remains finite. This will allow us to bootstrap the existence time for as long as we have good bounds in the caloric gauge. Precisely, for[8] $N \geq 3$ suppose that an H^N solution exists in the caloric gauge up to time T. If this solution has uniform H^N bounds up to time T, then its temporal gauge representation has uniform H^N bounds up to time T. Thus it can be extended further in the temporal gauge, and hence also in the caloric gauge. This shows that a maximal caloric gauge solution must either explode in H^N at the (finite) end of its lifespan, or the L^3 norm in (1-12) must explode. The latter cannot happen for subthreshold solutions. Thus we have:

Theorem 7.6. *The Yang–Mills system in the caloric gauge is locally well-posed in $H^N \times H^{N-1}$ for $N \geq 2$. Further, the solution extends for as long as the $H^N \times H^{N-1}$ norm remains bounded and the L^3 norm in (1-12) remains bounded.*

For regular data, this result reduces the problem of global well-posedness to that of obtaining uniform bounds for caloric solutions.

7B. *Local well-posedness in the caloric manifold $\mathcal{C}$: proof of Theorem 1.13.* For $\varepsilon_* > 0$, recall that the energy concentration scale $r_c^{\varepsilon_*}$ was defined as

$$r_c^{\varepsilon_*}[a,e] = \sup\{r : \mathcal{E}_{B_r}(a,e) \leq \varepsilon_*^2\} = \sup\left\{ r > 0 : \sup_{x \in \mathbb{R}^4} \frac{1}{2} \sum_{\alpha<\beta} \| f_{\alpha\beta} \|_{L^2(B_r(x))}^2 \leq \varepsilon_*^2 \right\},$$

where f_{jk} is the curvature form corresponding to a_j, $f_{0j} = -f_{j0} = e_j$ and $f_{00} = 0$. Since the definition only involves $f_{\alpha\beta}$, we will slightly abuse the notation and simply write $r_c^{\varepsilon_*}[f]$ for $r_c^{\varepsilon_*}[a,e]$.

Lemma 7.7. *Let A be a regular caloric Yang–Mills wave on $I = (-T_0, T_0)$. For any $\varepsilon > 0$, if ε_* is sufficiently small compared to ε and*

$$T_0 \leq r_c^{\varepsilon_*}[a,e],$$

then we have

$$\|F\|_{\mathrm{ED}_{\geq m}[I]} \leq \varepsilon, \quad \textit{with } 2^m = \varepsilon (r_c^{\varepsilon_*}[a,e])^{-1}.$$

Proof. By our notation, $f_{\alpha\beta} = F_{\alpha\beta}(0)$. After rescaling, we may set $r_c^{\varepsilon_*}(F(0)) = 1$. We begin with the observation that

$$\|P_k F(t)\|_{L^\infty} \lesssim 2^{ck_-} 2^{-2k} \sup_{x \in \mathbb{R}^4} \|F(t)\|_{L^2(B_1(x))}, \tag{7-4}$$

which follows from the properties of the convolution kernel of P_k; in particular, it is rapidly decaying on the scale 2^{-k} and its L^2-norm is bounded by 2^{-2k}. Then, by the localized energy estimate for the hyperbolic Yang–Mills equation, i.e.,

$$\mathcal{E}_{\{t\} \times B_{R-|t|}}(F) \leq \mathcal{E}_{\{0\} \times B_R}(F) \quad (0 < |t| < R), \tag{7-5}$$

the lemma follows. □

Proof of Theorem 1.13. We prove the theorem in several steps:

[8] The requirement $N \geq 3$ is so that there is no loss of regularity in the transition to the temporal gauge. Precisely, we want to ensure that $A_0 \in C(\dot{H}^1 \cap \dot{H}^{N+1})$.

Step 1: regular solutions. Let A be a regular caloric Yang–Mills wave with energy $\mathcal{E}$ and initial caloric size $\mathcal{Q}$. For ε_* small enough, to be chosen later, let $r_c := r_c^{\varepsilon_*}$ be the corresponding energy concentration scale for the initial data.

Our goal is to prove that if ε_* is small enough, depending only on $\mathcal{E}$ and $\mathcal{Q}$, then the solution A persists as a regular caloric solution up to time r_c. Precisely, we will apply Theorem 6.1 to the solution A in order to show that the solution A exists in $[-r_c, r_c]$ and satisfies the bound

$$\|A\|_{S^1[-r_c,r_c]} \leq M(\mathcal{E}, 3\mathcal{Q}). \tag{7-6}$$

We use a continuity argument. Let $T_0 \leq r_c$ be a maximal time with the property that the solution A given by Theorem 7.4 exists as a classical caloric solution in $(-T_0, T_0)$, and further satisfies the bound

$$\sup_{t\in[-T_0,T_0]} \mathcal{Q}(A(t)) \leq 3\mathcal{Q}. \tag{7-7}$$

For $0 < T < T_0$ we seek to apply Theorem 6.1 to A in $I = [-T, T]$. To verify the hypothesis of Theorem 6.1 we need to ensure that for a suitable choice of m we have

$$\|F\|_{\mathrm{ED}_{\geq m}} \leq \varepsilon(\mathcal{E}, 3\mathcal{Q}), \quad |I| \leq 2^{-m} T(\mathcal{E}, 3\mathcal{Q}).$$

For this it suffices to apply Lemma 7.7 with

$$\varepsilon = \min\{\varepsilon(\mathcal{E}, 3\mathcal{Q}), T(\mathcal{E}, 3\mathcal{Q})\},$$

which yields the appropriate choice of ε_*.

Now by Theorem 6.1 we obtain the uniform bound

$$\|A\|_{S^1[-T,T]} \leq M(\mathcal{E}, 3\mathcal{Q}), \quad 0 < T < T_0.$$

By the structure theorem, Theorem 5.1, it follows that higher-regularity bounds are also uniformly propagated,

$$\sup_{t\in(-T_0,T_0)} \|(A, \partial_t A)(t)\|_{H^N} < \infty.$$

Thus by the local result for regular solutions in Theorem 7.6 we can continue the regular caloric Yang–Mills connection A beyond the time interval $[-T_0, T_0]$.

Finally, we consider the bounds for $\mathcal{Q}(A)$. These we can propagate using Theorem 5.9, which implies that

$$\sup_{t\in[-T_0,T_0]} \mathcal{Q}(A(t)) - \mathcal{Q} \lesssim_{\mathcal{Q},\mathcal{E}} \varepsilon^{\delta_4}.$$

Readjusting ε if needed, it follows that

$$\sup_{t\in[-T_0,T_0]} \mathcal{Q}(A(t)) \leq 2\mathcal{Q}. \tag{7-8}$$

This implies that the bound (7-7) also can be propagated beyond $\pm T_0$. This contradicts the maximality of T_0 unless $T_0 = r_c$. Hence the classical caloric Yang–Mills wave exists in $[-r_c, r_c]$ and (7-6) holds.

Step 2: rough solutions. Given any caloric initial data (a, b) with finite energy $\mathcal{E}$ and caloric size $\mathcal{Q}$, we consider the corresponding regularized data $(a(s), b(s))$ obtained using the Yang–Mills heat flow. We have the uniform bounds

$$\mathcal{E}(a(s), b(s)) \leq \mathcal{E}(a, b), \quad \mathcal{Q}(a(s), b(s)) \leq \mathcal{Q}(a, b).$$

In particular, we have $(f(s), e(s)) \to (f, e)$ in $\dot{H}^1 \times L^2$. This implies that the energy concentration scales for $(a(s), e(s))$ converge to those for (a, e). Thus, by the analysis in the smooth case above, for small enough s the corresponding solutions $A(s)$ exist as smooth caloric Yang–Mills waves in $I = [-r_c, r_c]$ and satisfy the uniform S^1 bound (7-6).

Now we use the structure theorem, Theorem 5.1, to consider the limit as $s \to 0$. If c_k is a frequency envelope for (a, e), then by Proposition 3.1 it follows that:

(i) For $(a(s), b(s))$ we have the frequency envelope in $\dot{H}^1 \times L^2$

$$c_k(s) = c_k \langle 2^{2k} s \rangle^{-c\delta_5}.$$

(ii) For the difference $(a, b) - (a(s), b(s))$ we have the envelope in $\dot{H}^1 \times L^2$

$$\delta c_k(s) = c_k \langle 2^{-2k} s^{-1} \rangle^{-c\delta_5}.$$

(iii) For the difference $(a(s), b(s)) - (a(2s), b(2s))$ we have the envelope in $\dot{H}^1 \times L^2$

$$c_k^*(s) = c_{k(s)} 2^{-c\delta_5 |k - k(s)|}.$$

By Theorem 5.1(2), it follows that $c_k(s)$ is a frequency envelope for $A(s)$ in S_1. Combining this with Theorem 5.1(8), it follows that $c_k^*(s)$ is a frequency envelope for $A(s) - A(2s)$. Summing up such differences, we obtain the general difference bound

$$\|A(s_1) - A(s_2)\|_{S^1} \lesssim_{\mathcal{E}, \mathcal{Q}} c_{[k(s_1), k(s_2)]}. \tag{7-9}$$

This implies that the limit

$$A = \lim_{s \to 0} A(s)$$

exists in s. We define A to be the caloric Yang–Mills wave associated to the (a, b) data. We remark that by (7-9) we have the difference bound

$$\|A - A(s)\|_{S^1} \lesssim_{\mathcal{E}, \mathcal{Q}} c_{\geq k(s)}. \tag{7-10}$$

Step 3: difference bound. The difference bound in part (4) of the theorem is a direct consequence of the difference bound in Theorem 5.1(8).

Step 4: continuous dependence. We consider a convergent sequence of caloric initial data

$$(a^{(n)}, b^{(n)}) \to (a, b) \quad \text{in } \dot{H}^1 \times L^2. \tag{7-11}$$

Let $A^{(n)}(s)$ and $A(s)$ be the corresponding solutions with regularized data.

Denote by c_k^n a corresponding sequence of frequency envelopes for the initial data $(a^{(n)}, b^{(n)})$ in $\dot{H}^1 \times L^2$. By Theorem 5.1(2), these are also frequency envelopes for the solutions $A^{(n)}(s)$.

By Theorem 7.4 we know that for each s we have

$$A^{(n)}(s) \to A(s) \quad \text{in } S^1$$

and in effect in stronger topologies. Then we estimate

$$\begin{aligned}\limsup_{n\to\infty} \|A^{(n)} - A\|_{S^1} &\lesssim \lim_{s\to\infty} \limsup_{n\to\infty} \|A^{(n)}(s) - A(s)\|_{S^1} + c^n_{\geq k(s)} + c_{\geq k(s)} \\ &\lesssim \lim_{s\to\infty} \limsup_{n\to\infty} c^n_{\geq k(s)}.\end{aligned}$$

But the last limit is zero in view of the convergence in (7-11). The continuous dependence follows. ☐

We end this subsection with a lemma that bounds the energy concentration scale from below by an L^2-frequency envelope for F, which proves Remark 5.2.

Lemma 7.8. *Let c be a frequency envelope for $F_{\alpha\beta}$ in L^2 for all $\alpha, \beta \in \{0, 1, \dots, 4\}$. Suppose that $\|c\|_{\ell^2_{\geq m}} < C^{-1}\varepsilon_*$ for some $m \in \mathbb{Z}$ and a sufficiently large universal constant $C > 0$. Then $r_c^{\varepsilon_*}(F) \geq 2^{-m}$.*

Proof. It suffices to establish the bound

$$\|F\|_{L^2(B(x,2^{-k})} \lesssim c_{\geq k}.$$

To see this we use Bernstein's inequality to estimate

$$\|F\|_{L^2(B(x,2^{-k})} \lesssim \|F_{\geq k}\|_{L^2} + \sum_{j<k} 2^{-2k} \|F_j\|_{L^\infty} \lesssim c_{\geq k} + \sum_{j<k} 2^{2j-2k} c_j \approx c_{\geq k}.$$ ☐

7C. *Regularity of energy-dispersed solutions*: *proof of Theorem 1.20*. Consider a time t_0 where $\mathcal{Q}(A(t))$ is nearly minimal. From Lemma 5.10 we have the estimate

$$\mathcal{Q}(A(t_0)) \lesssim_\varepsilon \varepsilon^c.$$

If ε is small enough this allows us to conclude first that $\mathcal{Q} \leq 1$, and then that

$$\mathcal{Q} \lesssim_E \varepsilon^c.$$

Now a straightforward continuity argument shows that

$$Q(A(t)) \leq 1, \quad t \in I,$$

which again by Lemma 5.10 yields

$$\mathcal{Q}(A(t)) \lesssim_\varepsilon \varepsilon^c, \quad t \in I.$$

Then we can apply directly the result in Theorem 6.1 for any $m \in \mathbb{Z}$. This eliminates any restriction on the size of the interval I.

7D. *Gauge transformation into temporal gauge*: *proof of Theorem 1.18*. To produce a temporal gauge solution to (1-1) from the caloric gauge solution we use a gauge transformation O defined as the solution to the ODE

$$O^{-1}\partial_t O = A_0, \quad O(0) = I. \tag{7-12}$$

Here for A_0 we have the regularity given by Theorem 5.1(9), namely

$$A_0 \in \ell^1 |D|^{-2} L^2_x L^1_t. \tag{7-13}$$

We use this to compute the regularity of O:

Lemma 7.9. (a) *Assume that* A_0 *is as in* (7-13). *Then the solution* O *to the ODE has the following properties*:

(i) $O_{;x} \in C_t(\dot{H}^1)$.

(ii) *O is continuous in both x and t.*

(b) *Consider two solutions O and $\widetilde{O}$ arising from A_0 and $\tilde{A}_0$. Then we have*:

(i) ($\dot{H}^1$ *bound*)

$$\|O^{-1}\partial_x O - \widetilde{O}^{-1}\partial_x \widetilde{O}\|_{\dot{H}^1} \lesssim \|A_0 - \tilde{A}_0\|_{\ell^1|D|^{-2}L^2_x L^1_t}.$$

(ii) (*uniform bound*)

$$\|d(O, \widetilde{O})\|_{L^\infty} \lesssim \|A_0 - \tilde{A}_0\|_{\ell^1|D|^{-2}L^2_x L^1_t}.$$

Proof. (a) We first consider the ODE

$$O^{-1}\partial_t O = F, \quad O(0) = I, \tag{7-14}$$

and observe that for smooth F this is easily solvable.

Next we consider a smooth one-parameter family of solutions $O(h)$. For this we compute

$$\frac{d}{dt}(O^{-1}\partial_h O) = \partial_h F - [F, O^{-1}\partial_h O],$$

which immediately leads to

$$|O^{-1}\partial_h O(t)| \le \int_0^t |\partial_h F(s)|\, ds.$$

Comparing two solutions O and $\widetilde{O}$ generated by F and $\widetilde{F}$ using the straight line between them, it follows that

$$d(O(t), \widetilde{O}(t)) \le \int_0^t |F(s) - \widetilde{F}(s)|\, ds. \tag{7-15}$$

This yields a Lipschitz property for the map

$$L^1_t \ni F \to O \in C_t,$$

which is thus by density extended to all $F \in L^1_t$.

Next we turn our attention to A_0, which by Bernstein's inequality satisfies

$$A_0 \in C_x L^1_t.$$

This implies the desired continuity of O.

Finally we consider the evolution of $O^{-1}\partial_x O$,

$$\frac{d}{dt}(O^{-1}\partial_x O) = \partial_x A_0 - [A_0, O^{-1}\partial_x O].$$

Since $\partial_x A_0 \in L^4_x L^1_t$, this immediately gives

$$O^{-1}\partial_x O \in L^4_x C_t \subset CL^4.$$

A second differentiation yields as well

$$\partial_x(O^{-1}\partial_x O) \in L^2_x C_t \subset C L^2.$$

(b) The uniform bound for the difference follows directly from (7-15). For the difference of the derivatives we compute

$$\partial_t(O^{-1}\partial_j O - \tilde{O}^{-1}\partial_j \tilde{O}) + [A_0, O^{-1}\partial_j O - \tilde{O}^{-1}\partial_j \tilde{O}] = \partial_j A_0 - \partial_j \tilde{A}_0 - [A_0 - \tilde{A}_0, \tilde{O}\partial_j \tilde{O}].$$

As above, we can estimate this first in L^4 and then in $\dot{H}^1$. □

To conclude the proof of Theorem 1.18 it remains to verify (i) that gauge transformations O having the properties in the above lemma yield temporal connections $A^{[t]} \in C(\dot{H}^1)$, and (ii) these connections depend continuously on the initial data.

For the continuity in time we write

$$A^{[t]} = O(A - O^{-1}\partial_x O)O^{-1}.$$

The second term above is in $C_t \dot{H}^1$ due to the previous lemma. For the first term we differentiate, then use again the lemma combined with the continuity of O and dominated convergence.

For the continuous dependence of the temporal solutions with caloric data the same argument as above applies. However, we also need to consider general finite-energy initial data sets. Here the construction of the temporal gauge solutions starting from a general initial data (a, e) goes as follows:

(1) Given the initial position $a \in \dot{H}^1$, we consider the gauge transformation $O = O(a)$ which turns a into $(\tilde{a}, \tilde{e})$, its caloric gauge counterpart.

(2) Given the caloric data $(\tilde{a}, \tilde{e})$ we have as above a unique temporal solution $\tilde{A}$.

(3) To return to the data (a, e) we apply to $\boldsymbol{A}$ the inverse gauge transformation O^{-1} to obtain the temporal solution A.

The regularity of the gauge transformation O is $O^{-1}\partial_x O \in \dot{H}^1$, which suffices in order for it to map $C(\dot{H}^1)$ connections into $C(\dot{H}^1)$ connections. It remains to prove the continuous dependence. Consider a convergent sequence of data $(a^{(n)}, e^{(n)}) \to (a, e)$ in $\dot{H}^1 \times L^2$. Without any restriction in generality we can assume that (a, e) is caloric. Denote by $O^{(n)}$ the corresponding gauge transformations, which, we recall, are only unique up to constant gauge transformations. Then we need to show that for a well chosen (sub-)sequence of representatives $O^{(n)}$ we have the following properties:

(1) $(O^{(n)})^{-1}\partial_x O^{(n)} \to 0$ in $\dot{H}^1$.

(2) $O^{(n)}(x) \to I$ a.e. in x.

But this is a consequence of Theorem 1.2; see also Remark 1.3 (recall also that $O_{;x} = \mathrm{Ad}(O)(O^{-1}\partial_x O)$).

8. Multilinear estimates

The purpose of this section is to prove most of the results stated without proof in Section 4. The exceptions are Theorem 4.24 and Proposition 4.25, which involve construction of a parametrix for $\Box + \mathrm{Diff}^{\kappa}_{\boldsymbol{PA}}$; their proofs are given in the next section.

8A. ***Disposable operators and null forms.*** In this subsection we collect preliminary materials that are needed for analysis of the multilinear operators in the nonlinearity of the Yang–Mills equation in the caloric gauge.

8A1. *Disposable operators.* Boundedness properties of the multilinear operators arising in caloric gauge (see Section 3) can be conveniently phrased in terms of *disposability* (after multiplication with appropriate weights) of these operators.

We begin by considering the multilinear operator $\boldsymbol{Q}$ with the symbol

$$\boldsymbol{Q}(\xi,\eta) = \frac{|\xi|^2 - |\eta|^2}{2(|\xi|^2+|\eta|^2)} = \frac{(\xi+\eta)\cdot(\xi-\eta)}{2(|\xi|^2+|\eta|^2)},$$

which arose in the wave equation for A_x (most notably through the expression for $\partial^\ell A_\ell$) in the caloric gauge.

Lemma 8.1. *For any $k, k_1, k_2 \in \mathbb{Z}$, the bilinear operator*

$$2^{k_{\max}-k} P_k \boldsymbol{Q}(P_{k_1}(\cdot), P_{k_2}(\cdot))$$

is disposable.

Proof. To begin with, note the symbol bound

$$|\boldsymbol{Q}(\xi,\eta)| \lesssim \frac{|\xi+\eta|}{(|\xi|^2+|\eta|^2)^{\frac{1}{2}}},$$

which implies that the symbol of $2^{k_{\max}-k} P_k \boldsymbol{Q}(P_{k_1}(\cdot), P_{k_2}(\cdot))$ is uniformly bounded. In the case $k_2 < k_1 - 5$ so that $|k_{\max} - k| \leq 3$, it can also be checked that

$$2^{n_1 k_1} 2^{n_2 k_2} |\partial_\xi^{(n_1)} \partial_\eta^{(n_2)} (P_k(\xi+\eta)\boldsymbol{Q}(\xi,\eta) P_{k_1}(\xi) P_{k_2}(\eta))| \lesssim_{n_1,n_2} 1,$$

which proves the desired disposability property. By symmetry, the case $k_1 < k_2 - 5$ follows as well. In the case $|k_1 - k_2| < 5$ (so that $|k_{\max} - k_1| < 10$), making the change of variables $(\xi,\zeta) = (\xi, \xi+\eta)$, it can be seen that

$$2^{k_1-k} 2^{n_1 k_1} 2^{n_2 k} |\partial_\xi^{(n_1)} \partial_\zeta^{(n_2)} (P_k(\zeta)\boldsymbol{Q}(\xi,\zeta-\xi) P_{k_1}(\xi) P_{k_2}(\zeta-\xi))| \lesssim_{n_1,n_2} 1,$$

which implies disposability of $2^{k_{\max}-k} P_k \boldsymbol{Q}(P_{k_1}(\cdot), P_{k_2}(\cdot))$. ☐

Next, we consider the multilinear operator $\boldsymbol{W}(s)$ with the symbol

$$\boldsymbol{W}(\xi,\eta,s) = -\frac{1}{2\xi\cdot\eta} e^{-s|\xi+\eta|^2}(1-e^{2s\xi\cdot\eta}),$$

which arose in the wave equation for the Yang–Mills heat flow development $A_x(s)$ of a caloric Yang–Mills wave.

Lemma 8.2. *For any $k, k_1, k_2 \in \mathbb{Z}$ and $s > 0$, the bilinear operator*

$$\langle s 2^{2k}\rangle^{10} \langle s^{-1} 2^{-2k_{\max}}\rangle 2^{2k_{\max}} P_k \boldsymbol{W}(P_{k_1}(\cdot), P_{k_2}(\cdot), s) \tag{8-1}$$

is disposable.

Proof. Without loss of generality, we may assume that $s = 1$ by scaling. We distinguish two scenarios:

Case 1: high-low or low-high, $k = \max\{k_1, k_2\} + O(1)$. To prove disposability of (8-1), it suffices to show that

$$\langle 2^{2k_{\max}} \rangle^{11} 2^{n_1 k_1} 2^{n_2 k_2} \left| \partial_\xi^{(n_1)} \partial_\eta^{(n_2)} \left(P_k(\xi+\eta) e^{-|\xi+\eta|^2} \frac{1 - e^{2\xi\cdot\eta}}{\xi \cdot \eta} P_{k_1}(\xi) P_{k_2}(\eta) \right) \right| \lesssim_{n_1,n_2} 1$$

for any $n_1, n_2 \in \mathbb{N}$. Since the derivatives of $P_k(\xi+\eta)P_{k_1}(\xi)P_{k_2}(\eta)$ already obey desirable bounds, it only remains to prove

$$\langle 2^{2k_{\max}} \rangle^{11} 2^{n_1 k_1} 2^{n_2 k_2} \left| \partial_\xi^{(n_1)} \partial_\eta^{(n_2)} \left(e^{-|\xi+\eta|^2} \frac{1 - e^{2\xi\cdot\eta}}{\xi \cdot \eta} \right) \right| \lesssim_{n_1,n_2} 1 \tag{8-2}$$

for ξ, η in the support of the symbol (8-1).

Since $k = \max\{k_1, k_2\} + O(1)$, we have $2^{2k_{\max}} \simeq |\xi|^2 + |\eta|^2 \simeq |\xi+\eta|^2$. On the one hand, it is straightforward to verify

$$\begin{aligned} 2^{n_1 k_1} 2^{n_2 k_2} |\partial_\xi^{(n_1)} \partial_\eta^{(n_2)} e^{-|\xi+\eta|^2}| &\lesssim_{n_1,n_2} 2^{n_1 k_1} 2^{n_2 k_2} (1 + |\xi+\eta|^2)^{\frac{n_1+n_2}{2}} e^{-|\xi+\eta|^2} \\ &\lesssim_{n_1,n_2} 2^{(n_1+n_2)k_{\max}} \langle 2^{2k_{\max}} \rangle^{\frac{n_1+n_2}{2}} e^{-|\xi+\eta|^2}. \end{aligned}$$

On the other hand, we also have

$$\begin{aligned} 2^{n_1 k_1} 2^{n_2 k_2} \left| \partial_\xi^{(n_1)} \partial_\eta^{(n_2)} \left(\frac{1 - e^{2\xi\cdot\eta}}{\xi\cdot\eta} \right) \right| &\lesssim_{n_1,n_2} 2^{n_1 k_1} 2^{n_2 k_2} (1 + |\xi|^2 + |\eta|^2)^{\frac{n_1+n_2}{2}} (1 + e^{2\xi\cdot\eta}) \\ &\lesssim_{n_1,n_2} 2^{(n_1+n_2)k_{\max}} \langle 2^{2k_{\max}} \rangle^{\frac{n_1+n_2}{2}} (1 + e^{2\xi\cdot\eta}). \end{aligned}$$

The key point here is that when $|\xi \cdot \eta| \ll 1$, the denominator $\xi \cdot \eta$ cancels with the first term in the Taylor expansion of the numerator $1 - \xi \cdot \eta$; we omit the details. Combining (8-3) and (8-3), it follows that

$$2^{n_1 k_1} 2^{n_2 k_2} \left| \partial_\xi^{(n_1)} \partial_\eta^{(n_2)} \left(e^{-|\xi+\eta|^2} \frac{1 - e^{2\xi\cdot\eta}}{\xi\cdot\eta} \right) \right| \lesssim_{n_1,n_2} \langle 2^{2k_{\max}} \rangle^{n_1+n_2} e^{-|\xi+\eta|^2} (1 + e^{2\xi\cdot\eta}).$$

Since $e^{-|\xi+\eta|^2}(1 + e^{2\xi\cdot\eta}) = e^{-|\xi+\eta|^2} + e^{-(|\xi|^2+|\eta|^2)} \lesssim e^{-C^{-1}2^{2k_{\max}}}$, (8-2) follows.

Case 2: high-high, $k < \max\{k_1, k_2\} - C$. As usual, we make the change of variables $(\xi, \zeta) = (\xi, \xi + \eta)$. It suffices to prove

$$\langle 2^{2k} \rangle^{10} \langle 2^{2k_{\max}} \rangle 2^{n_1 k_1} 2^{n_2 k} \left| \partial_\xi^{(n_1)} \partial_\zeta^{(n_2)} \left(P_k(\zeta) e^{-|\zeta|^2} \frac{1 - e^{2\xi\cdot(\zeta-\xi)}}{\xi \cdot (\zeta - \xi)} P_{k_1}(\xi) P_{k_2}(\xi - \zeta) \right) \right| \lesssim_{n_1,n_2} 1.$$

Note that the derivatives of $\langle 2^{2k} \rangle^{10} P_k(\zeta) e^{-|\zeta|^2} P_{k_1}(\xi) P_{k_2}(\xi - \zeta)$ already obey desirable bounds. Hence we are only left to show

$$\langle 2^{2k_{\max}} \rangle 2^{n_1 k_1} 2^{n_2 k} \left| \partial_\xi^{(n_1)} \partial_\zeta^{(n_2)} \left(\frac{1 - e^{2\xi\cdot(\zeta-\xi)}}{\xi\cdot(\zeta-\xi)} \right) \right| \lesssim_{n_1,n_2} 1 \tag{8-3}$$

for ξ, ζ in the support of (8-1).

Note that $k_1 = k_{\max} + O(1)$. In the case $2^{2k_{\max}} \lesssim 1$, (8-3) follows from

$$|\partial_\xi^{(n_1)} \partial_\zeta^{(n_2)}((2\xi \cdot (\zeta - \xi))^{-1}(1 - e^{2\xi\cdot(\zeta-\xi)}))| \lesssim_{n_1,n_2} 1,$$

which follows by Taylor expansion at $\xi \cdot (\zeta - \xi) = 0$. In the case $2^{2k_{\max}} \gtrsim 1$, we use

$$2^{n_1 k_1} 2^{n_2 k} |\partial_\xi^{(n_1)} \partial_\zeta^{(n_2)} (\xi \cdot (\zeta - \xi))^{-1}| \lesssim 2^{-2k_{\max}},$$
$$2^{n_1 k_1} 2^{n_2 k} |\partial_\xi^{(n_1)} \partial_\zeta^{(n_2)} (1 - e^{2\xi\cdot(\zeta-\xi)})| \lesssim 1,$$

both of which follow from simple computation, whose details we omit. □

8A2. *Null forms.* We now discuss the null forms that arise in caloric gauge, which occur in conjunction with various (disposable) translation-invariant operators. To treat these in a systematic fashion, it is useful to define null forms in terms of an appropriate decomposition property of the symbol.

Definition 8.3 (null forms). Let $\mathcal{T}$ be a translation-invariant bilinear operator on $\mathbb{R}^{1+4}$ and let $\pm \in \{+, -\}$ be a sign. Given $k_1, k_2 \in \mathbb{Z}$, $\ell, \ell' \in -\mathbb{N}$, $\omega, \omega' \in \mathbb{S}^3$, define

$$\theta_\pm = \max\{|\angle(\omega, \pm\omega')|, 2^{\ell}, 2^{\ell'}\}.$$

(1) We say that $\mathcal{T}$ is a *null form of type* $\mathcal{N}_\pm$, and write

$$\mathcal{T}(\cdot,\cdot) = \mathcal{N}_\pm(\cdot,\cdot),$$

if for every $k_1, k_2 \in \mathbb{Z}$, $\ell, \ell' \in -\mathbb{N}$ and $\omega, \omega' \in \mathbb{S}^3$, $\mathcal{T}$ admits a decomposition of the form

$$\mathcal{T}((\tau,\xi),(\sigma,\eta))(P_{k_1} P_\ell^\omega)(\xi)(P_{k_2} P_{\ell'}^{\omega'})(\eta) = \theta_\pm 2^{k_1+k_2} \mathcal{O}((\tau,\xi),(\sigma,\eta)) \sum_{i_1,i_2\in\mathbb{N}} a_{i_1}(\xi) b_{i_2}(\eta),$$

where the Fourier multipliers

$$(1+|i_1|)^{100} a_{i_1}, \quad (1+|i_2|)^{100} b_{i_2} \tag{8-4}$$

are disposable, and the translation-invariant bilinear operator with symbol

$$\mathcal{O}((\tau,\xi),(\sigma,\eta))$$

is disposable as well.

(2) We say that $\mathcal{T}$ is a *null form of type* $\mathcal{N}$ if $\mathcal{T}(\cdot,\cdot) = \mathcal{N}_+(\cdot,\cdot)$ and $\mathcal{T}(\cdot,\cdot) = \mathcal{N}_-(\cdot,\cdot)$.

(3) We say that $\mathcal{T}$ is a *null form of type* $\mathcal{N}_{0,\pm}$, and write

$$\mathcal{T}(\cdot,\cdot) = \mathcal{N}_{0,\pm}(\cdot,\cdot),$$

if for every $k_1, k_2 \in \mathbb{Z}$, $\ell, \ell' \in -\mathbb{N}$ and $\omega, \omega' \in \mathbb{S}^3$, $\mathcal{T}$ admits a decomposition of the form

$$\mathcal{T}(\xi,\eta)(P_{k_1} P_\ell^\omega)(\xi)(P_{k_2} P_{\ell'}^{\omega'})(\eta) = \theta_\pm^2 2^{k_1+k_2} \mathcal{O}((\tau,\xi),(\eta,\sigma)) \sum_{i_1,i_2\in\mathbb{N}} a_{i_1}(\xi) b_{i_2}(\eta),$$

where the Fourier multipliers

$$(1+|i_1|)^{100} a_{i_1}, \quad (1+|i_2|)^{100} b_{i_2} \tag{8-5}$$

are disposable, and also the translation-invariant bilinear operator which has symbol $\mathcal{O}((\tau,\xi),(\sigma,\eta))$ is disposable as well.

In particular, $\mathcal{O}$, a_{i_1} and b_{i_2} may depend on $k_1, k_2, \ell, \ell', \omega, \omega'$, but the disposability bounds stated above do not.

Remark 8.4 (null form gain). To exploit the null form, it is convenient to make the following observation: as an immediate consequence of the definition, we may write

$$\mathcal{N}_\pm(P_{k_1} P_\ell^\omega u, P_{k_2} P_{\ell'}^{\omega'} v) = C\theta_\pm 2^{k_1+k_2} \widetilde{\mathcal{O}}(P_{k_1} P_\ell^\omega u, P_{k_2} P_{\ell'}^{\omega'} v)$$

for a universal constant $C > 0$ and some disposable $\widetilde{\mathcal{O}}$. Analogous statements hold for $\mathcal{N}$ and $\mathcal{N}_{0,\pm}$.

Remark 8.5 (behavior under symbol multiplication). The properties of $\mathcal{T}$ in Definition 8.3 seem complicated at first, but its usefulness comes from the fact that it is well-behaved under symbol-multiplication with a disposable multilinear operator. More precisely, if $\mathcal{O}(\cdot,\cdot)$ is a disposable translation-invariant bilinear operator and $\mathcal{T}(\cdot,\cdot)$ is a null form in the sense of Definition 8.3, then the translation-invariant bilinear operator with symbol $\mathcal{O}(\xi,\eta)\mathcal{T}(\xi,\eta)$ is clearly also a null form of the same type.

We now verify that the standard null forms are indeed null forms according to Definition 8.3. We have the following separation-of-variables result for the symbols of the standard null forms.

Lemma 8.6 (standard null forms). *Consider the symbols*

$$\boldsymbol{N}_{ij}(\xi,\eta) = \xi_i\eta_j - \xi_j\eta_i, \quad \boldsymbol{N}_{0,\pm}(\xi,\eta) = \pm|\xi||\eta| - \xi\cdot\eta.$$

These symbols admit the decompositions

$$|\xi|^{-1}|\eta|^{-1}\boldsymbol{N}_{ij}(\xi,\eta)(P_{k_1}P_\ell^\omega)(\xi)(P_{k_2}P_{\ell'}^{\omega'})(\eta) = \min\{\theta_+,\theta_-\} \sum_{i_1,i_2\in\mathbb{N}} a_{i_1}(\xi) b_{i_2}(\eta), \tag{8-6}$$

$$|\xi|^{-1}|\eta|^{-1}\boldsymbol{N}_{0,\pm}(\xi,\eta)(P_{k_1}P_\ell^\omega)(\xi)(P_{k_2}P_{\ell'}^{\omega'})(\eta) = \theta_\pm^2 \sum_{i_1,i_2\in\mathbb{N}} a'_{i_1}(\xi) b'_{i_2}(\eta), \tag{8-7}$$

where

$$(1+|i_1|)^{100} a_{i_1}, \quad (1+|i_1|)^{100} a'_{i_1}, \quad (1+|i_2|)^{100} b_{i_2}, \quad (1+|i_2|)^{100} b'_{i_2} \tag{8-8}$$

are disposable.

As a corollary, it follows that $\boldsymbol{N}_{ij}$ is a null form of type $\mathcal{N}$, whereas $\boldsymbol{N}_{0,\pm}$ are null forms of type $\mathcal{N}_\pm$.

As before, a_{i_1}, a'_{i_1}, b_{i_2} and b'_{i_2} depend on $k_1, k_2, \ell, \ell', \omega, \omega'$, but the disposability bounds stated in (8-8) do not.

This lemma can be proved by performing separation of variables using Fourier series on an appropriate rectangular box containing the support of $P_{k_1}P_\ell^\omega(\xi)P_{k_2}P_\omega^{\ell'}(\xi')$. For the details in the case of $|\xi|^{-1}|\eta|^{-1}\boldsymbol{N}_{ij}(\xi,\eta)$, we refer to [Gavrus and Oh 2016, Proof of Proposition 7.8]. For $\boldsymbol{N}_{0,\pm}$, observe that $\widetilde{\boldsymbol{N}}_{0,\pm}(\xi,\eta) := |\xi|^{-1}|\eta|^{-1}\boldsymbol{N}_{0,\pm}(\xi,\eta)$ obeys

$$|\widetilde{\boldsymbol{N}}_{0,\pm}(\xi,\eta)| \lesssim \theta_\pm^2, \quad |\partial_\xi \widetilde{\boldsymbol{N}}_{0,\pm}(\xi,\eta)| \lesssim 2^{-k_1}\theta_\pm, \quad |\partial_\eta \widetilde{\boldsymbol{N}}_{0,\pm}(\xi,\eta)| \lesssim 2^{-k_2}\theta_\pm,$$
$$|\partial_\xi^{(n_1)}\partial_\eta^{(n_2)} \widetilde{\boldsymbol{N}}_{0,\pm}(\xi,\eta)| \lesssim 2^{-n_1k_1}2^{-n_2k_2} \quad (n_1+n_2 \geq 2)$$

for ξ, η in the support of $P_{k_1} P^{\omega}_{\ell}(\xi) P_{k_2} P^{\ell'}_{\omega}(\eta)$. Using these symbol bounds, the case of $N_{0,\pm}$ can be handled by essentially the same proof as in [Gavrus and Oh 2016, Proof of Proposition 7.8]. See also [Gavrus 2019, Section 8].

We now present algebraic lemmas, which are used to identify null forms in the Yang–Mills equation in the caloric gauge. The following lemma identifies all bilinear null forms.

Lemma 8.7. *Let $\mathcal{O}$ be a disposable bilinear operator on $\mathbb{R}^{1+4}$. Let A be a spatial 1-form and let u, v be functions in the Schwartz class on $\mathbb{R}^{1+4}$. Then we have*

$$\mathcal{O}(\boldsymbol{P}^{\ell} A, \partial_{\ell} u) = \sum_j \mathcal{N}(|D|^{-1} A_j, u), \tag{8-9}$$

$$\boldsymbol{P}_x \mathcal{O}(u, \partial_x v) = |D|^{-1} \mathcal{N}(u, v). \tag{8-10}$$

Moreover, we also have

$$\mathcal{O}(\partial^{\alpha} u, \partial_{\alpha} v) = \mathcal{N}_{0,+}(Q^{+}u, Q^{+}v) + \mathcal{N}_{0,+}(Q^{-}u, Q^{-}v) + \mathcal{N}_{0,-}(Q^{+}u, Q^{-}v) + \mathcal{N}_{0,-}(Q^{-}u, Q^{+}v) + \mathcal{R}_0(u, v), \tag{8-11}$$

where

$$\mathcal{R}_0(u', v') = \mathcal{O}((D_t - |D|)Q^{+}u' + (D_t + |D|)Q^{-}u, D_t v') + \mathcal{O}(|D|(Q^{+}u' - Q^{-}u'), (D_t - |D|)Q^{+}v' + (D_t + |D|)Q^{-}v'). \tag{8-12}$$

Remark 8.8. As is evident from the proof below, Lemma 8.7 readily generalizes to a disposable multilinear operator $\mathcal{O}$ that has one of the above structures with respect to two inputs. We omit the precise statement, as the notation gets unnecessarily involved. However, we point out that this is all we need in order to handle the trilinear secondary null structure.

Remark 8.9. An alternative way to make use of the null form $\mathcal{O}(\partial^{\alpha} u, \partial_{\alpha} v)$ is to rely on the simple algebraic identity

$$2\mathcal{O}(\partial^{\alpha} u, \partial_{\alpha} v) = \Box \mathcal{O}(u, v) - \mathcal{O}(\Box u, v) - \mathcal{O}(u, \Box v). \tag{8-11$'$}$$

We have elected to use the decomposition (8-11) to unify the treatment of null forms.

Proof. We begin with (8-9) and (8-10). By Remark 8.5, it suffices to consider the case when $\mathcal{O}(u, v)$ is the product uv. Then it is a well-known fact (going back to [Klainerman and Machedon 1994; 1995]) that $\boldsymbol{P}^{\ell} A \partial_{\ell} u$ and $\boldsymbol{P}_j(u \partial_x v)$ are standard null forms, i.e.,

$$\boldsymbol{P}^{\ell} A \partial_{\ell} u = \boldsymbol{N}_{\ell j}((-\Delta)^{-1} \partial^{\ell} A^j, u), \tag{8-13}$$

$$\boldsymbol{P}_j(u \partial_x v) = (-\Delta)^{-1} \partial^{\ell} \boldsymbol{N}_{\ell j}(u, v). \tag{8-14}$$

We omit the simple symbol computation. Hence (8-9) and (8-10) follow.

Next, we prove (8-11), which is essentially the well-known fact that $\partial^{\alpha} u \partial_{\alpha} v = -D^{\alpha} u D_{\alpha} v$ is a null form. To verify (8-11), we first decompose $u = Q^{+}u + Q^{-}u$ and $v = Q^{+}v + Q^{-}$, then we substitute

$$D_t Q^{\pm} u = \pm |D| Q^{\pm} u + (D_t \mp |D|) Q^{\pm} u, \quad D_t Q^{\pm'} v = \pm' |D| Q^{\pm'} v + (D_t \mp' |D|) Q^{\pm'} v.$$

When $\mathcal{O}(u,v)=uv$, the contribution of the first terms gives

$$\sum_{\pm,\pm'}(\pm\pm'|D|Q^{\pm}u|D|Q^{\pm'}v-D^{\ell}Q^{\pm}uD_{\ell}Q^{\pm'}v)=\sum_{\pm,\pm'}\boldsymbol{N}_{0,\pm\pm'}(Q^{\pm}u,Q^{\pm'}v).$$

By Remark 8.5, the same contribution constitutes the first four terms in (8-11) in general. Note moreover that the remainder makes up $\mathcal{R}_0(u,v)$, which proves (8-11). □

Next, we present an algebraic computation, which will be used to reveal the trilinear secondary null form of the caloric Yang–Mills wave equation.

Lemma 8.10. *Let $\mathcal{O},\mathcal{O}'$ be disposable bilinear operators on $\mathbb{R}^{1+4}$. Then we have*

$$\begin{aligned}
&\mathcal{O}'(\Delta^{-1}\mathcal{O}(u^{(1)},\partial_0u^{(2)}),\partial^0u^{(3)})+\mathcal{O}'(\Box^{-1}\boldsymbol{P}_i\mathcal{O}(u^{(1)},\partial_xu^{(2)}),\partial^iu^{(3)})\\
&\quad=\mathcal{O}'(\Box^{-1}\mathcal{O}(u^{(1)},\partial_\alpha u^{(2)}),\partial^\alpha u^{(3)})-\mathcal{O}'(\Box^{-1}\Delta^{-1}\partial_t\partial_\alpha\mathcal{O}(u^{(1)},\partial^\alpha u^{(2)}),\partial_tu^{(3)})\\
&\qquad-\mathcal{O}'(\Box^{-1}\Delta^{-1}\partial_\ell\partial_\alpha\mathcal{O}(u^{(1)},\partial^\ell u^{(2)}),\partial^\alpha u^{(3)}),
\end{aligned}$$

provided that $\Delta^{-1}\mathcal{O}$, $\Box^{-1}\mathcal{O}$ and $\Box^{-1}\Delta^{-1}\mathcal{O}$ are well-defined in the sense that their kernels have finite masses.

Of course, the requirement that the kernels of $\Delta^{-1}\mathcal{O}$, $\Box^{-1}\mathcal{O}$ and $\Box^{-1}\Delta^{-1}\mathcal{O}$ have finite masses is excessively strong for the validity of the lemma, but it will be verified in the applications below.

Proof. The proof of this lemma is the same as in [Krieger et al. 2015, Appendix]. Using the identities

$$\Delta^{-1}-\Box^{-1}=\Box^{-1}\Delta^{-1}(-\partial_t^2),\qquad \boldsymbol{P}_iB=B_i-\Delta^{-1}\partial_i\partial^\ell B_\ell,\qquad \partial^0=-\partial_0=-\partial_t$$

and adding and subtracting $\mathcal{O}'(\Box^{-1}\Delta^{-1}\partial_t\partial^\ell\mathcal{O}(u^{(1)},\partial_\ell u^{(2)}),\partial_tu^{(3)})$, we may write

$$\begin{aligned}
&\mathcal{O}'(\Delta^{-1}\mathcal{O}(u^{(1)},\partial_0u^{(2)}),\partial^0u^{(3)})+\mathcal{O}'(\Box^{-1}\boldsymbol{P}_i\mathcal{O}(u^{(1)},\partial_xu^{(2)}),\partial^iu^{(3)})\\
&\quad=\mathcal{O}'(\Box^{-1}\mathcal{O}(u^{(1)},\partial_0u^{(2)}),\partial^0u^{(3)})+\mathcal{O}'(\Box^{-1}\mathcal{O}(u^{(1)},\partial_iu^{(2)}),\partial^iu^{(3)})\\
&\qquad-\mathcal{O}'(\Box^{-1}\Delta^{-1}\partial_t\partial^0\mathcal{O}(u^{(1)},\partial_0u^{(2)}),\partial_tu^{(3)})-\mathcal{O}'(\Box^{-1}\Delta^{-1}\partial_i\partial^\ell\mathcal{O}(u^{(1)},\partial_\ell u^{(2)}),\partial^iu^{(3)})\\
&\qquad\quad-\mathcal{O}'(\Box^{-1}\Delta^{-1}\partial_t\partial^\ell\mathcal{O}(u^{(1)},\partial_\ell u^{(2)}),\partial_tu^{(3)})-\mathcal{O}'(\Box^{-1}\Delta^{-1}\partial_0\partial^\ell\mathcal{O}(u^{(1)},\partial_\ell u^{(2)}),\partial^0u^{(3)})\\
&\quad=\mathcal{O}'(\Box^{-1}\mathcal{O}(u^{(1)},\partial_\alpha u^{(2)}),\partial^\alpha u^{(3)})-\mathcal{O}'(\Box^{-1}\Delta^{-1}\partial_t\partial_\alpha\mathcal{O}(u^{(1)},\partial^\alpha u^{(2)}),\partial_tu^{(3)})\\
&\qquad-\mathcal{O}'(\Box^{-1}\Delta^{-1}\partial_\ell\partial_\alpha\mathcal{O}(u^{(1)},\partial^\ell u^{(2)}),\partial^\alpha u^{(3)}).
\end{aligned}$$

In the last equality, we paired the first and the second, the third and the fifth, and the fourth and the sixth terms, respectively, from the preceding lines. □

8B. *Summary of global-in-time dyadic estimates.* In what follows, we denote by $\mathcal{O}$ a disposable translation-invariant bilinear operator on $\mathbb{R}^{1+4}$, and by $\mathcal{N}$ a bilinear null form as in Definition 8.3(2). Let u and v be test functions on $\mathbb{R}^{1+4}$. For convenience, we also introduce test functions u' and v', which stand for inputs of the form ∇u and ∇v, respectively, in the applications.

Given $k,k_1,k_2\in\mathbb{Z}$, we define $k_{\max}=\max\{k,k_1,k_2\}$ and $k_{\min}=\min\{k,k_1,k_2\}$. We use the shorthand $u_{k_1}=P_{k_1}u$, $v_{k_2}=P_{k_2}v$ and $v'_{k_2}=P_{k_2}v'$.

8B1. *Bilinear estimates for elliptic components.* We start with simple bilinear bounds which do not involve any null forms.

Proposition 8.11. *We have*

$$\|P_k\mathcal{O}(u_{k_1}, v'_{k_2})\|_{L^2\dot{H}^{-1/2}} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})}\|Du_{k_1}\|_{\mathrm{Str}^0}\|v'_{k_2}\|_{\mathrm{Str}^0}, \tag{8-15}$$

$$\|P_k\mathcal{O}(u_{k_1}, v'_{k_2})\|_{L^{9/5}\dot{H}^{-4/9}} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})}\|Du_{k_1}\|_{\mathrm{Str}^0}\|v'_{k_2}\|_{\mathrm{Str}^0}, \tag{8-16}$$

$$\|P_k\mathcal{O}(u_{k_1}, v'_{k_2})\|_{L^1\dot{W}^{-2,\infty}} \lesssim 2^{-\delta_1|k_1-k_2|}\|Du_{k_1}\|_{S}\|v'_{k_2}\|_{S}. \tag{8-17}$$

Furthermore, we have the following simpler variants of (8-15), (8-16) *and* (8-17):

$$\|P_k\mathcal{O}(u_{k_1}, v'_{k_2})\|_{L^2\dot{H}^{-1/2}} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})}\|u_{k_1}\|_{L^2\dot{H}^{3/2}}\|v'_{k_2}\|_{S}, \tag{8-18}$$

$$\|P_k\mathcal{O}(u_{k_1}, v'_{k_2})\|_{L^{9/5}\dot{H}^{-4/9}} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})}\|u_{k_1}\|_{L^2\dot{H}^{3/2}}\|v'_{k_2}\|_{S}, \tag{8-19}$$

$$\|P_k\mathcal{O}(u_{k_1}, v'_{k_2})\|_{L^1\dot{W}^{-2,\infty}} \lesssim 2^{\frac{2}{3}k_{\min}}2^{-\frac{4}{3}k}2^{-\frac{1}{6}k_1}2^{\frac{5}{6}k_2}(2^{\frac{1}{6}k_1}\|u_{k_1}\|_{L^2L^6})(2^{-\frac{5}{6}k_2}\|v'_{k_2}\|_{L^2L^6}). \tag{8-20}$$

8B2. *Bilinear estimates concerning the N-norm.* Next, we state the N-norm estimates which will be used for the bilinear expressions arising from $\boldsymbol{P}\mathcal{M}$, $\boldsymbol{P}^{\perp}\mathcal{M}$ and $\mathrm{Rem}^{\kappa,2}$.

Proposition 8.12. *We have*

$$\|P_k\mathcal{N}(u_{k_1}, v_{k_2})\|_{N} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})}2^{k}\|Du_{k_1}\|_{S}\|Dv_{k_2}\|_{S}, \tag{8-21}$$

$$\|P_k\mathcal{O}(\partial^\alpha u_{k_1}, \partial_\alpha v_{k_2})\|_{N} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})}2^{k_{\max}}\|Du_{k_1}\|_{S}\|Dv_{k_2}\|_{S}, \tag{8-22}$$

$$\|P_k\mathcal{O}(u'_{k_1}, v_{k_2})\|_{L^1L^2} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})}\|u'_{k_1}\|_{L^2\dot{H}^{1/2}}(2^{\frac{1}{6}k_2}\|v_{k_2}\|_{L^2L^6}). \tag{8-23}$$

Furthermore, for any $\kappa \in \mathbb{N}$, we have the low-modulation gain

$$\|P_k Q_{<k_{\min}-\kappa}\mathcal{N}(Q_{<k_{\min}-\kappa}u_{k_1}, Q_{<k_{\min}-\kappa}v_{k_2})\|_{N} \lesssim 2^{-\delta_1\kappa}2^{k}\|Du_{k_1}\|_{S}\|Dv_{k_2}\|_{S}, \tag{8-24}$$

$$\|P_k Q_{<k_{\min}-\kappa}\mathcal{O}(\partial^\alpha Q_{<k_{\min}-\kappa}u_{k_1}, \partial_\alpha Q_{<k_{\min}-\kappa}v_{k_2})\|_{N} \lesssim 2^{-\delta_1\kappa}2^{k_{\max}}\|Du_{k_1}\|_{S}\|Dv_{k_2}\|_{S}. \tag{8-25}$$

For the term $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}B$, we need to distinguish the case when the low-frequency input A has a dominant modulation. For this purpose, we borrow the bilinear operator $\mathcal{H}^*_k$ (and its "dual" $\mathcal{H}_k$) from [Krieger et al. 2015].

Given a bilinear translation-invariant operator $\mathcal{O}$, we introduce the expression $\mathcal{H}_k\mathcal{O}$ (resp. $\mathcal{H}^*_k\mathcal{O}$), which essentially separates out the case when the modulation of the output (resp. the first input) is dominant. More precisely, we define

$$\mathcal{H}_k\mathcal{O}(u,v) = \sum_{j:j<k+C} Q_j\mathcal{O}(Q_{<j-C}u, Q_{<j-C}v),$$

$$\mathcal{H}^*_k\mathcal{O}(u,v) = \sum_{j:j<k+C} Q_{<j-C}\mathcal{O}(Q_j u, Q_{<j-C}v)$$

for some universal constant C such that $C < C_0$, where C_0 is the constant in Lemma 8.21. We also define

$$\mathcal{H}\mathcal{O}(u,v) = \sum_{k,k_1,k_2:k<k_2-C} P_k \mathcal{H}_k \mathcal{O}(P_{k_1}u, P_{k_2}v),$$

$$\mathcal{H}^*\mathcal{O}(u,v) = \sum_{k,k_1,k_2:k_1<k_2-C} \mathcal{H}^*_{k_1} P_k \mathcal{O}(P_{k_1}u, P_{k_2}v).$$

We are now ready to state our estimates for the N-norm of the term $\mathrm{Diff}_{\boldsymbol{P}A} B$.

Proposition 8.13. *For $k_1 < k - 10$, we have*

$$\|P_k(1-\mathcal{H}^*_{k_1})\mathcal{N}(|D|^{-1}u_{k_1}, v_{k_2})\|_N \lesssim \|Du_{k_1}\|_S \|Dv_{k_2}\|_S, \tag{8-26}$$

$$\|P_k(1-\mathcal{H}^*_{k_1})\mathcal{O}(u_{k_1}, v'_{k_2})\|_N \lesssim \|u_{k_1}\|_{L^2\dot{H}^{3/2}} \|v'_{k_2}\|_S, \tag{8-27}$$

$$\|P_k\mathcal{H}^*_{k_1}\mathcal{N}(|D|^{-1}u_{k_1}, v_{k_2})\|_N \lesssim \|u_{k_1}\|_{Z^1} \|Dv_{k_2}\|_S, \tag{8-28}$$

$$\|P_k\mathcal{H}^*_{k_1}\mathcal{O}(u_{k_1}, v'_{k_2})\|_N \lesssim \|u_{k_1}\|_{\Delta^{-1/2}\Box^{1/2}Z^1} \|v'_{k_2}\|_S. \tag{8-29}$$

Furthermore, for $k_1 < k - 10$ and any $\kappa \in \mathbb{N}$, we have

$$\|P_k\mathcal{H}^*_{k_1}\mathcal{N}(|D|^{-1}Q_{<k_1-\kappa}u_{k_1}, v_{k_2})\|_N \lesssim 2^{-\delta_1\kappa}\|u_{k_1}\|_{Z^1} \|Dv_{k_2}\|_S, \tag{8-30}$$

$$\|P_k\mathcal{H}^*_{k_1}\mathcal{O}(Q_{<k_1-\kappa}u_{k_1}, v'_{k_2})\|_N \lesssim 2^{-\delta_1\kappa}\|u_{k_1}\|_{\Delta^{-1/2}\Box^{1/2}Z^1} \|v'_{k_2}\|_S. \tag{8-31}$$

8B3. *Bilinear estimates concerning $X^{s,b,p}_r$-type norms.* We now state the Z^1-, $Z^1_{p_0}$- and $\widetilde{Z}^1_{p_0}$-norm bounds. We begin with the ones for the bilinear expressions arising from $\boldsymbol{P}\mathcal{M}^2$, $\mathrm{Rem}^{\kappa,2}_A$ and $\mathcal{M}^2_0$.

Proposition 8.14. *We have*

$$\|P_k\mathcal{N}(u_{k_1}, v_{k_2})\|_{\Box Z^1_{p_0}} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})}2^k \|Du_{k_1}\|_S \|Dv_{k_2}\|_S, \tag{8-32}$$

$$\|P_k\mathcal{N}(u_{k_1}, v_{k_2})\|_{\Box Z^1} \lesssim 2^{-\delta_1|k_1-k_2|}2^k \|Du_{k_1}\|_S \|Dv_{k_2}\|_S. \tag{8-33}$$

Furthermore, for $k \le k_1 - C$, we have

$$\|P_k(1-\mathcal{H}_k)\mathcal{N}(u_{k_1}, v_{k_2})\|_{\Box Z^1} \lesssim 2^{-\delta_1(k_1-k)}2^k \|Du_{k_1}\|_S \|Dv_{k_2}\|_S, \tag{8-34}$$

$$\|P_k(1-\mathcal{H}_k)\mathcal{O}(u_{k_1}, v'_{k_2})\|_{\Delta^{1/2}\Box^{1/2}Z^1} \lesssim 2^{-\delta_1(k_1-k)} \|Du_{k_1}\|_S \|v'_{k_2}\|_S. \tag{8-35}$$

The following bounds are for the null form arising from $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}_x A} B$; we remark that this is the only place where we need to use the intermediate $\widetilde{Z}^1_{p_0}$-norm.

Proposition 8.15. *We have*

$$\|P_k\mathcal{N}(|D|^{-1}u_{k_1}, v_{k_2})\|_{\Box \widetilde{Z}^1_{p_0}} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})} \|u_{k_1}\|_{S^1} \|Dv_{k_2}\|_S, \tag{8-36}$$

$$\|P_k\mathcal{N}(|D|^{-1}u_{k_1}, v_{k_2})\|_{\Box Z^1_{p_0}} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})} \|u_{k_1}\|_{S^1\cap \widetilde{Z}^1_{p_0}} \|Dv_{k_2}\|_S, \tag{8-37}$$

$$\|P_k\mathcal{N}(|D|^{-1}u_{k_1}, v_{k_2})\|_{\Box Z^1} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})} \|u_{k_1}\|_{S^1\cap Z^1_{p_0}} \|Dv_{k_2}\|_S, \tag{8-38}$$

$$\|P_k\mathcal{N}(|D|^{-1}u_{k_1}, v_{k_2})\|_{X^{-1/2+b_1,-b_1}} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})} \|u_{k_1}\|_{S^1\cap Z^1_{p_0}} \|Dv_{k_2}\|_S. \tag{8-39}$$

Finally, the following bounds are used to handle $\mathrm{Diff}^{\kappa}_{A_0} B$ and $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}^{\perp}A} B$.

Proposition 8.16. *We have*

$$\|P_k \mathcal{O}(u_{k_1}, v'_{k_2})\|_{\square Z^1_{p_0}} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})}\|Du_{k_1}\|_Y \|v'_{k_2}\|_S, \tag{8-40}$$

$$\|P_k \mathcal{O}(u_{k_1}, v'_{k_2})\|_{\square Z^1} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})}\|Du_{k_1}\|_Y \|v'_{k_2}\|_S, \tag{8-41}$$

$$\|P_k \mathcal{O}(u_{k_1}, v'_{k_2})\|_{X^{-1/2+b_1,-b_1}} \lesssim 2^{-\delta_1(k_{\max}-k_{\min})}\|Du_{k_1}\|_Y \|v'_{k_2}\|_S. \tag{8-42}$$

8B4. *Trilinear null form estimate.* Let $u^{(1)}$, $u^{(2)}$, $u^{(3)}$ be test function on $\mathbb{R}^{1+4}$. Given $k_i \in \mathbb{Z}$, we introduce the shorthand $u^{(i)}_{k_i} = P_{k_i} u^{(i)}$ $(i = 1, 2, 3)$.

Proposition 8.17. *Let $\mathcal{O}$ and $\mathcal{O}'$ be disposable bilinear operators on $\mathbb{R}^{1+4}$. Let $j < k - C$ and $k < \min\{k_0, k_1, \dots, k_3\} - C$. Consider the expression*

$$\mathcal{N}^{\text{cubic}}_{k,j}(u^{(1)}_{k_1}, u^{(2)}_{k_2}, u^{(3)}_{k_3}) = Q_{<j-C}\mathcal{O}'(\Delta^{-1} P_k Q_j \mathcal{O}(Q_{<j-C} u^{(1)}_{k_1}, \partial_0 Q_{<j-C} u^{(2)}_{k_2}), \partial^0 Q_{<j-C} u^{(3)}_{k_3})$$
$$+ Q_{<j-C}\mathcal{O}'(\square^{-1} P_k Q_j \boldsymbol{P}_\ell \mathcal{O}(Q_{<j-C} u^{(1)}_{k_1}, \partial_x Q_{<j-C} u^{(2)}_{k_2}), \partial^\ell Q_{<j-C} u^{(3)}_{k_3}).$$

Then we have

$$\|\mathcal{N}^{\text{cubic}}_{k,j}(u^{(1)}_{k_1}, u^{(2)}_{k_2}, u^{(3)}_{k_3})\|_{L^1L^2} \lesssim 2^{-\delta_1(k_1-k)} 2^{-\delta_1(k-j)} \|Du^{(1)}_{k_1}\|_S \|Du^{(2)}_{k_2}\|_S \|Du^{(3)}_{k_3}\|_S. \tag{8-43}$$

In fact, for later use (in Section 11), it is convenient to also state a more atomic form of (8-43). Given $k_i \in \mathbb{Z}$ and a rectangular box $\mathcal{C}^{(i)}$, we use the shorthand $u^{(i)}_{k_i, \mathcal{C}^{(i)}} = P_{k_i} P_{\mathcal{C}^{(i)}} u^{(i)}$ $(i = 1, 2)$.

Proposition 8.18. *Suppose $\mathcal{O}$ and $\mathcal{O}'$ are translation-invariant bilinear operators on $\mathbb{R}^{1+4}$ such that $\mathcal{O}(P^\omega_\ell \cdot, P^{\omega'}_{\ell'} \cdot)$ and $\mathcal{O}'(P^\omega_\ell \cdot, P^{\omega'}_{\ell'} \cdot)$ are disposable for every $\ell, \ell' \in -\mathbb{N}$ and $\omega, \omega' \in \mathbb{S}^3$. Let $j < k - C$, $k < \min\{k_0, k_1, \dots, k_3\} - C$ and $\mathcal{C}^{(1)}, \mathcal{C}^{(2)} \in \{\mathcal{C}_k(\ell)\}$, where $\ell = \frac{j-k}{2}$. We have*

$$\|P_{k_0} Q_{<j-C}\mathcal{O}'(\square^{-1} P_k Q_j \mathcal{O}(Q_{<j-C} u^{(1)}_{k_1,\mathcal{C}^{(1)}}, \partial_\alpha Q_{<j-C} u^{(2)}_{k_2,\mathcal{C}^{(2)}}), \partial^\alpha Q_{<j-C} u^{(3)}_{k_3})\|_{L^1L^2}$$
$$\lesssim 2^{-\delta_1(k_1-k)} 2^{-\delta_1(k-j)} \|Du^{(1)}_{k_1,\mathcal{C}^{(1)}}\|_{S_{k_1}[\mathcal{C}_k(\ell)]} \|Du^{(2)}_{k_2,\mathcal{C}^{(2)}}\|_{S_{k_2}[\mathcal{C}_k(\ell)]} \|Du^{(3)}_{k_3}\|_S, \tag{8-44}$$

$$\|P_{k_0} Q_{<j-C}\mathcal{O}'(\square^{-1}\Delta^{-1} P_k Q_j \partial_t \partial_\alpha \mathcal{O}(Q_{<j-C} u^{(1)}_{k_1,\mathcal{C}^{(1)}}, \partial^\alpha Q_{<j-C} u^{(2)}_{k_2,\mathcal{C}^{(2)}}), \partial_t Q_{<j-C} u^{(3)}_{k_3})\|_{L^1L^2}$$
$$\lesssim 2^{-\delta_1(k_1-k)} 2^{-\delta_1(k-j)} \|Du^{(1)}_{k_1,\mathcal{C}^{(1)}}\|_{S_{k_1}[\mathcal{C}_k(\ell)]} \|Du^{(2)}_{k_2,\mathcal{C}^{(2)}}\|_{S_{k_2}[\mathcal{C}_k(\ell)]} \|Du^{(3)}_{k_3}\|_S, \tag{8-45}$$

$$\|P_{k_0} Q_{<j-C}\mathcal{O}'(\square^{-1}\Delta^{-1} P_k Q_j \partial_\ell \partial_\alpha \mathcal{O}(Q_{<j-C} u^{(1)}_{k_1,\mathcal{C}^{(1)}}, \partial^\ell Q_{<j-C} u^{(2)}_{k_2,\mathcal{C}^{(2)}}), \partial^\alpha Q_{<j-C} u^{(3)}_{k_3})\|_{L^1L^2}$$
$$\lesssim 2^{-\delta_1(k_1-k)} 2^{-\delta_1(k-j)} \|Du^{(1)}_{k_1,\mathcal{C}^{(1)}}\|_{S_{k_1}[\mathcal{C}_k(\ell)]} \|Du^{(2)}_{k_2,\mathcal{C}^{(2)}}\|_{S_{k_2}[\mathcal{C}_k(\ell)]} \|Du^{(3)}_{k_3}\|_S. \tag{8-46}$$

8C. *Proof of the interval-localized estimates.* In this subsection, we prove all estimates claimed in Section 4 except Theorem 4.24 and Proposition 4.25, which are proved in the next section.

The key technical issue we address here is passage to interval-localized frequency envelope bounds (as stated in Section 4) from the global-in-time dyadic estimates stated in Section 8B.

In what follows, we denote by $\mathcal{O}$ and $\boldsymbol{O}$ disposable multilinear operators on $\mathbb{R}^{1+4}$ and $\mathbb{R}^4$, respectively, which may vary from line to line. Similarly, χ^k_I indicates a generalized time cutoff adapted to the scale 2^{-k}, which may vary from line to line.

8C1. *Estimates that do not involve any null forms.* Here we establish Propositions 4.12, 4.13, 4.14 and 4.18, whose proofs do not involve any null forms.

Proofs of Propositions 4.12 and 4.13. We introduce the shorthand $A' = \partial_t A$ and $B' = \partial_t B$. Using (4-25) and Lemma 8.1 we write

$$|D|^{-1} P_k \mathcal{M}_0^2(P_{k_1} A, P_{k_2} B) = 2^{-k} P_k \boldsymbol{O}(P_{k_1} A, P_{k_2} B'), \tag{8-47}$$

$$P_k \boldsymbol{Q}(P_{k_1} A, P_{k_2} B) = 2^k 2^{-k_{\max}} P_k \boldsymbol{O}(P_{k_1} A, P_{k_2} B), \tag{8-48}$$

$$|D|^{-1} P_k \boldsymbol{Q}(P_{k_1} A, P_{k_2} \partial_t B) = 2^{-k_{\max}} P_k \boldsymbol{O}(P_{k_1} A, P_{k_2} B'), \tag{8-49}$$

$$|D|^{-2} P_k \mathcal{D}\mathcal{M}_0^2(P_{k_1} A, P_{k_2} B) = 2^{-k} 2^{-k_{\max}} P_k \boldsymbol{O}(P_{k_1} A', P_{k_2} B'). \tag{8-50}$$

Step 1: fixed-time estimates. Applying Hölder and Bernstein (to one of the inputs or the output, whichever has the lowest frequency), we obtain

$$\| P_k \boldsymbol{O}(P_{k_1} u', P_{k_2} v') \|_{L^2} \lesssim 2^{2k_{\min}} \| u' \|_{L^2} \| v' \|_{L^2}. \tag{8-51}$$

Recalling (8-47)–(8-50), the fixed-time estimates (4-27), (4-28) and (4-35) follow.

Step 2: space-time estimates. Here, we prove the remaining estimates in Propositions 4.12 and 4.13. In this step, we simply extend A, B, A', B' by zero outside I. Furthermore, we define

$$\mathcal{M}^{\kappa,2}_{0,\text{small}}(A, B) = \sum_{|k_{\max} - k_{\min}| \geq \kappa} P_k \mathcal{M}_0^2(P_{k_1} A, P_{k_2} B), \tag{8-52}$$

$$\mathcal{M}^{\kappa,2}_{0,\text{large}}(A, B) = \sum_{|k_{\max} - k_{\min}| < \kappa} P_k \mathcal{M}_0^2(P_{k_1} A, P_{k_2} B), \tag{8-53}$$

so that $\mathcal{M}_0^{\kappa,2}(A, B) = \mathcal{M}^{\kappa,2}_{0,\text{small}}(A, B) + \mathcal{M}^{\kappa,2}_{0,\text{large}}(A, B)$.

Step 2.1: $L^2 \dot{H}^{1/2}$-norm estimates. We first verify (4-29)–(4-34), (4-36) and (4-38) with the $L^2 \dot{H}^{1/2}$-norm (instead of the Y-norm) on the left-hand side. All of these estimates follow from (8-15) and (8-47)–(8-50). The small factor in (4-31) arises from the exponential gain in (8-15) and the frequency gap κ in (8-52), whereas the factor $\varepsilon^{\delta_2} M$ in (4-33), (4-34) and (4-38) arises from (4-21).

Step 2.2: $L^1 L^\infty$-norm estimates. By Hölder's inequality, we have

$$\| P_k u \|_{L^{p_0} \dot{W}^{2-3/p_0, p_0'}} \lesssim \| P_k u \|^{1-\theta_0}_{L^2 \dot{H}^{1/2}} \| P_k u \|^{\theta_0}_{L^1 \dot{W}^{-1,\infty}}, \tag{8-54}$$

where $\theta_0 = 2\big(\frac{1}{p_0} - \frac{1}{2}\big) \in (0, 1)$. Therefore, (4-29), (4-31) and (4-33) follow by combining (8-17) with the $L^2 \dot{H}^{1/2}$-norm estimates from Step 2.1. On the other hand, for (4-32) we use (8-20) instead of (8-17), which allows us to use the DS^1-norm on the right-hand side at the expense of losing the exponential off-diagonal gain. Finally, for (4-37) and (4-38), observe that by (8-20), (8-48) and (8-49) we have

$$\| |D|^{-\sigma-1} \boldsymbol{Q}(P_{k_1} A, P_{k_2} B') \|_{L^1 L^\infty} \lesssim 2^{-\delta_1 (k_{\max} - k_{\min})} \| P_{k_1} A \|_{DS^1} \| |D|^{-\sigma} P_{k_2} B' \|_{DS^1}$$

for $\sigma = 0, 1$. Therefore, the $L^1 L^\infty$-norm bound in (4-37) follows directly, whereas the Y-norm bounds in (4-37) and (4-38) follow after interpolating with the $L^2 \dot{H}^{1/2}$-norm estimates from Step 2.1. □

Proofs of Proposition 4.14. For this proof we use the square function $L_x^{10/3} L_t^2$ component of the S_k norm, for which we have

$$\|u\|_{S_k^{sq}} = 2^{-\frac{3}{10}k} \|u\|_{L_x^{10/3} L_t^2}.$$

We recall that the symbol of $\Delta \boldsymbol{A}_0^2$ is

$$\Delta \boldsymbol{A}_0^2(\xi, \eta) = \frac{|\xi|^2}{|\xi|^2 + |\eta|^2}.$$

Then we use Bernstein at the lowest frequency to estimate

$$\|P_k \Delta \boldsymbol{A}_0^2(A_{k_1}, \partial_t A_{k_2})\|_{L^2 L^1} \lesssim 2^{-2(k_2-k_1)_+} 2^{-\frac{7}{10}k_1} 2^{\frac{3}{10}k_2} 2^{\frac{4}{10}k_{\min}} c_{k_1} c_{k_2} \lesssim 2^{-\frac{3}{10}(k_{\max}-k_{\min})} c_{k_1} c_{k_2}.$$

Now the bound (4-39) immediately follows due to the off-diagonal decay. □

Proof of Proposition 4.18. The bounds in this proposition are trivial consequences of Proposition 8.11, along with the observation that $\||D|u\|_{\mathrm{Str}^0} \lesssim \|\nabla u\|_{L^2 \dot{H}^{1/2}}$. We omit the details. □

8C2. *Estimates for $\boldsymbol{P}\mathcal{M}^2$, $\boldsymbol{P}^\perp \mathcal{M}^2$ and* $\mathrm{Rem}^{2,\kappa}$. We now present the proofs of Propositions 4.15 and 4.20, which require the bilinear null form estimates in Proposition 8.12, as well as the $X_r^{s,b,p}$-type norm estimates in Propositions 8.14, 8.15 and 8.16.

Proof of Proposition 4.15. Unless otherwise stated, we extend the inputs A, B by homogeneous waves outside I. For $k, k_1, k_2 \in \mathbb{Z}$, by Lemma 8.1, note that

$$P_k \boldsymbol{P}\mathcal{M}^2(P_{k_1} A, P_{k_2} B) = P_k \boldsymbol{PO}(P_{k_1} A, \partial_x P_{k_2} B), \tag{8-55}$$

$$P_k \boldsymbol{P}^\perp \mathcal{M}^2(P_{k_1} A, P_{k_2} B) = 2^{-k_{\max}} P_k \boldsymbol{O}(\partial_\alpha P_{k_1} A, \partial^\alpha P_{k_2} B) \tag{8-56}$$

for some disposable operator $\boldsymbol{O}$ on $\mathbb{R}^4$. Note also that, by Lemma 8.7, the right-hand sides are null forms.

<u>Step 0</u>: proofs of (4-40), (4-41). In view of (8-55) and (8-56), both follow easily using the standard Littlewood–Paley trichotomy and (8-51).

<u>Step 1</u>: proofs of (4-42), (4-43), (4-44) and (4-45). The N-norm bounds in (4-42) and (4-43) follow from the null form estimates (8-21)–(8-22). On the other hand, the $\Box \underline{X}^1$-norm bounds in (4-42) and (4-43) follow from (8-15), (8-16) and (8-32); we remark that the $\Box Z^1_{p_0}$-norm bound for $\boldsymbol{P}^\perp \mathcal{M}$ is unnecessary, since $\boldsymbol{P}\boldsymbol{P}^\perp \mathcal{M} = 0$. Estimates (4-44) and (4-45) immediately follow from (8-15), where we may simply extend A, $\partial_t A$, B, $\partial_t B$ by zero outside I as in the proofs of Propositions 4.12 and 4.13 above.

<u>Step 2</u>: proofs of (4-46), (4-47), (4-48) and (4-49). Since the case of $\boldsymbol{P}\mathcal{M}^2$ (i.e., estimates (4-46) and (4-48)) can be read off from [Oh and Tataru 2018, Proof of Proposition 4.1], we will only provide a detailed proof in the case of $\boldsymbol{P}^\perp \mathcal{M}^2$ (i.e., estimates (4-47), (4-49)).

Step 2.1: off-diagonal dyadic frequencies. If $\max\{|k-k_1|, |k-k_2|\} \geq \kappa$, then by (8-22) we have

$$\begin{aligned} \|P_k \boldsymbol{P}^\perp \mathcal{M}^2(P_{k_1} A, P_{k_2} B)\|_N &\lesssim 2^{-\delta_1(k_{\max}-k_{\min})} \|P_{k_1} A\|_{S^1} \|P_{k_2} B\|_{S^1} \\ &\lesssim 2^{-\frac{1}{2}\delta_1 \kappa} 2^{-\frac{1}{2}\delta_1(k_{\max}-k_{\min})} \|P_{k_1} A\|_{S^1} \|P_{k_2} B\|_{S^1}. \end{aligned}$$

Hence the contribution in the case $\max\{|k-k_1|, |k-k_2|\} \geq \kappa$ can always be put in $\boldsymbol{P}^\perp \mathcal{M}^{\kappa,2}_{\mathrm{small}}$.

Step 2.2: balanced dyadic frequencies, short time interval. Next, we consider the case when $|k-k_1|<\kappa$, $|k-k_2|<\kappa$ and $|I|\le 2^{-k+C\kappa}$. Then by Hölder and (8-56), we simply estimate

$$\begin{aligned}\|P_k \boldsymbol{P}^\perp \mathcal{M}^2(P_{k_1}A, P_{k_2}B)\|_{L^1L^2[I]} &\lesssim |I|^{\frac{1}{2}}\|P_k \boldsymbol{P}^\perp \mathcal{M}^2(P_{k_1}A, P_{k_2}B)\|_{L^2L^2[I]}\\ &\lesssim |I|^{\frac{1}{2}}2^{-k_{\max}}\|\mathcal{O}(\partial^\alpha P_{k_1}A, \partial_\alpha P_{k_2}B)\|_{L^2L^2}\\ &\lesssim 2^{C\kappa}\||D|^{-\frac{3}{4}}\nabla A_{k_1}\|_{L^4L^4[I]}\||D|^{-\frac{3}{4}}\nabla B_{k_2}\|_{L^4L^4[I]}.\end{aligned}$$

Therefore, when $|I|\le 2^{-k+C\kappa}$, the contribution in the case $\max\{|k-k_1|,|k-k_2|\}<\kappa$ can be put in $\boldsymbol{P}^\perp\mathcal{M}^{\kappa,2}_{\mathrm{large}}$.

Step 2.3: balanced dyadic frequencies, long time interval. Finally, we consider the case when $|k-k_1|<\kappa$, $|k-k_2|<\kappa$ and $|I|\ge 2^{-k+C\kappa}$. We define $\boldsymbol{P}^\perp\mathcal{M}^{\kappa,2}_{\mathrm{large}}$ by the relation

$$\begin{aligned}&\sum_{\max\{|k-k_1|,|k-k_2|\}<\kappa} P_k \boldsymbol{P}^\perp \mathcal{M}^2(P_{k_1}A, P_{k_2}B)\\ &\quad= \sum_{\max\{|k-k_1|,|k-k_2|\}<\kappa} P_k Q_{<k_{\min}-\kappa}\boldsymbol{P}^\perp \mathcal{M}^2(P_{k_1}Q_{<k_{\min}-\kappa}A, P_{k_2}Q_{<k_{\min}-\kappa}B) + \boldsymbol{P}^\perp\mathcal{M}^{\kappa,2}_{\mathrm{large}}(A,B).\end{aligned}$$

By (8-25), the first term on the right-hand side gains a factor of $2^{-c\delta_1\kappa}$, and therefore can be put in $\boldsymbol{P}^\perp\mathcal{M}^{\kappa,2}_{\mathrm{small}}$. Now it only remains to establish (4-49) for $\boldsymbol{P}^\perp\mathcal{M}^{\kappa,2}_{\mathrm{large}}$ defined as above.

By definition, $\boldsymbol{P}^\perp\mathcal{M}^{\kappa,2}_{\mathrm{large}}(A,B)$ is the sum over $\{(k,k_1,k_2):\max\{|k-k_1|,|k-k_2|\}<\kappa\}$ of

$$P_k \boldsymbol{P}^\perp \mathcal{M}^2(P_{k_1}A, P_{k_2}B) - P_k Q_{<k_{\min}-\kappa}\boldsymbol{P}^\perp \mathcal{M}^2(P_{k_1}Q_{<k_{\min}-\kappa}A, P_{k_2}Q_{<k_{\min}-\kappa}B).$$

Since we are allowed to lose an exponential factor in κ in (4-49), it suffices to freeze k,k_1,k_2 and estimate the preceding expression. At this point, we divide into three subcases:

Step 2.3a: output has high modulation. When the output has modulation $\ge 2^{k_{\min}-\kappa}$, we use the $X_1^{0,-1/2}$-component of the N-norm. Since the kernel of $P_kQ_{\ge k_{\min}-\kappa}$ decays rapidly in t on the scale $\simeq 2^{-k}2^{C\kappa}$, we have

$$\|P_kQ_{\ge k_{\min}-\kappa}\boldsymbol{P}^\perp \mathcal{M}^2(P_{k_1}A, P_{k_2}B)\|_{X_1^{0,-1/2}[I]} \lesssim 2^{C\kappa}2^{-\frac{1}{2}k}\|\chi_I^k\,\boldsymbol{P}^\perp \mathcal{M}^2(P_{k_1}A, P_{k_2}A)\|_{L^2L^2}$$

for some generalized cutoff function χ_I^k adapted to the scale 2^{-k}. Then, by Proposition 4.10,

$$\begin{aligned}2^{C\kappa}2^{-\frac{1}{2}k}\|\chi_I^k\,\boldsymbol{P}^\perp \mathcal{M}^2(P_{k_1}A, P_{k_2}A)\|_{L^2L^2} &\lesssim 2^{C\kappa}\|\chi_I^k|D|^{-\frac{3}{4}}\nabla P_{k_1}A\|_{L^4L^4}\|\chi_I^k|D|^{-\frac{3}{4}}\nabla P_{k_2}B\|_{L^4L^4}\\ &\lesssim 2^{C\kappa}\||D|^{-\frac{3}{4}}\nabla P_{k_1}A\|_{L^4L^4[I]}\||D|^{-\frac{3}{4}}\nabla P_{k_2}B\|_{L^4L^4[I]},\end{aligned}$$

which is acceptable.

Step 2.3b: A has high modulation. Next, we consider the case when the output has modulation $<2^{k_{\min}-\kappa}$, yet A has modulation $\ge 2^{k_{\min}-\kappa}$. The kernel of $P_kQ_{<k_{\min}-\kappa}$ again decays rapidly in t on the scale

$\simeq 2^{-k}2^{C\kappa}$. For any $2 \le q \le \infty$, we have

$$\begin{aligned}\|P_k Q_{<k_{\min}-\kappa} \boldsymbol{P}^{\perp}\mathcal{M}^2(Q_{\ge k_{\min}-\kappa} P_{k_1} A, P_{k_2} B)\|_{L^1 L^2[I]} \\ \lesssim 2^{C\kappa} \|\chi_I^k \, \boldsymbol{P}^{\perp}\mathcal{M}^2(Q_{\ge k_1-\kappa} P_{k_1} A, P_{k_2} B)\|_{L^1 L^2} \\ \lesssim 2^{C\kappa} \||D|^{-\frac{1}{q}} \Box P_{k_1} A\|_{L^{q'} L^2} \|\chi_I^k \, |D|^{2-\frac{1}{q}} \nabla P_{k_2} B\|_{L^q L^\infty} \\ \lesssim 2^{C\kappa} \||D|^{-\frac{1}{q}} \Box P_{k_1} A\|_{L^{q'} L^2[I]} \||D|^{2-\frac{1}{q}} \nabla P_{k_2} B\|_{L^q L^\infty[I]},\end{aligned}$$

where we used Proposition 4.10 on the last line. Taking $q = 2$, we see that the last line is bounded by $\lesssim 2^{C\kappa} \|\Box P_{k_1} A\|_{L^2 \dot{H}^{-1/2}[I]} \|P_{k_2} B\|_{DS^1[I]}$, which is acceptable.

Step 2.3c: B has high modulation. Finally, the only remaining case is when the output and A have modulation $< 2^{k_{\min}-\kappa}$, but B has modulation $\ge 2^{k_{\min}-\kappa}$. Proceeding as in Step 2.3b, and using the fact that the kernel of $P_{k_1} Q_{<k_{\min}-\kappa}$ decays rapidly in t on the scale $\simeq 2^{-k}2^{C\kappa}$, we have

$$\begin{aligned}\|P_k Q_{<k_{\min}-\kappa} \boldsymbol{P}^{\perp}\mathcal{M}^2(Q_{<k_{\min}-\kappa} P_{k_1} A, Q_{\ge k_{\min}-\kappa} P_{k_2} B)\|_{L^1 L^2[I]} \\ \lesssim 2^{C\kappa} \|\chi_I^k \, \boldsymbol{P}^{\perp}\mathcal{M}^2(Q_{<k_1-\kappa} P_{k_1} A, Q_{\ge k_2-\kappa} P_{k_2} B)\|_{L^1 L^2} \\ \lesssim 2^{C\kappa} \|\chi_I^k \, |D|^{-\frac{3}{2}} \nabla Q_{<k_{\min}-\kappa} P_{k_1} A\|_{L^2 L^\infty} \||D|^{-\frac{1}{2}} \Box P_{k_2} B\|_{L^2 L^2} \\ \lesssim 2^{C\kappa} \||D|^{-\frac{3}{2}} \nabla P_{k_1} A\|_{L^2 L^\infty[I]} \|\Box P_{k_2} B\|_{L^2 \dot{H}^{-1/2}[I]},\end{aligned}$$

which is acceptable.

Step 3: proofs of (4-50) and (4-51). Since the $L^2 \dot{H}^{-1/2}$-norm bounds follow from (4-21), (4-44) and (4-45), it remains to only consider the N-norm. The case of $\boldsymbol{P}\mathcal{M}^2$ can be read off from [Oh and Tataru 2018, Proof of Proposition 4.1]. Finally, for $\boldsymbol{P}^{\perp}\mathcal{M}^2$, we split into the small and large parts as in Step 2. For the small part, we already have

$$\|\boldsymbol{P}^{\perp}\mathcal{M}^{\kappa,2}_{\mathrm{small}}(A, B)\|_{N_c[I]} \lesssim 2^{-c\delta_1 \kappa} \|A\|_{S^1_c[I]} M.$$

For the large part, we proceed as in Step 2, except we choose $q = \frac{9}{4}$ in Step 2.3b. Then by (4-20), (4-21) and the embedding

$$\mathrm{Str}^1[I] \subseteq L^4 L^4[I] \cap L^{9/4} L^\infty[I],$$

it follows that

$$\|\boldsymbol{P}^{\perp}\mathcal{M}^{\kappa,2}_{\mathrm{large}}(A, B)\|_{N_c[I]} \lesssim 2^{C\kappa} \varepsilon^{\delta_1} \|A\|_{\underline{S}^1_c[I]} M.$$

Therefore, choosing $2^{-\kappa} = \varepsilon^c$ with $c > 0$ sufficiently small, (4-51) follows. □

Remark 8.19. As a corollary of the preceding proof in the case of $\boldsymbol{P}\mathcal{M}^2$, we obtain the following statement: let $\boldsymbol{O}$ be a disposable operator on $\mathbb{R}^4$, and let A, B be $\mathfrak{g}$-valued functions (or 1-forms) on I. Then we have

$$\begin{aligned}\|P_k(\boldsymbol{O}(\partial_i P_{k_1} A, \partial_j P_{k_2} B) - \boldsymbol{O}(\partial_j P_{k_1} A, \partial_i P_{k_2} B))\|_{N[I]} \\ \lesssim 2^{C(k_{\max}-k_{\min})} 2^k \|P_{k_1} A\|_{DS^1[I]} \|P_{k_2} B\|_{DS^1[I]}. \quad (8\text{-}57)\end{aligned}$$

Moreover, if (B, I) is (ε, M)-energy-dispersed, then

$$\| P_k(\boldsymbol{O}(\partial_i P_{k_1} A, \partial_j P_{k_2} B) - \boldsymbol{O}(\partial_j P_{k_1} A, \partial_i P_{k_2} B))\|_{N[I]} \lesssim 2^{C(k_{\max}-k_{\min})} 2^k \varepsilon^{c\delta_1} \| P_{k_1} A\|_{\underline{S}^1[I]} M. \tag{8-58}$$

Proof of Proposition 4.20. We decompose $\mathrm{Rem}^{\kappa,2}_A B$ into

$$\mathrm{Rem}^{\kappa,2}_A B = \mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}_x A} B + \mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}^\perp A} B + \mathrm{Rem}^{\kappa,2}_{A_0} B,$$

where

$$\mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}_x A} B = \sum_{\substack{k,k_1,k_2 \\ k_1 \geq k_2 - \kappa}} 2 P_k [\boldsymbol{P}_\ell P_{k_1} A, \partial^\ell P_{k_2} B], \tag{8-59}$$

$$\mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}^\perp A} B = \sum_{\substack{k,k_1,k_2 \\ k_1 \geq k_2 - \kappa}} 2 P_k [P_{k_1} \boldsymbol{P}^\perp_\ell A, \partial^\ell P_{k_2} B], \tag{8-60}$$

$$\mathrm{Rem}^{\kappa,2}_{A_0} B = - \sum_{\substack{k,k_1,k_2 \\ k_1 \geq k_2 - \kappa}} 2 P_k [P_{k_1} A_0, P_{k_2} \partial_t B]. \tag{8-61}$$

By Littlewood–Paley trichotomy, note that the summands on the right-hand sides of (8-59)–(8-61) vanish unless $k - k_1 \leq \kappa + C$.

Unless otherwise stated, we extend B by homogeneous waves outside I. For (8-59), we extend A by homogeneous waves outside I and for (8-60)–(8-61), we extend $\boldsymbol{P}^\perp_\ell A$ and A_0 by zero outside I. (Of course $\boldsymbol{P}^\perp$ of the extended A does not coincide with such an extension of $\boldsymbol{P}^\perp A$ outside I, but this will not be an issue.)

<u>Step 1</u>: proofs of (4-77) and (4-78). The N-norm bound in (4-77) follows from Lemma 8.7 and (8-21) for $\mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}_x A} B$, and (8-23) for $\mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}^\perp A} B$, $\mathrm{Rem}^{\kappa,2}_{A_0} B$. On the other hand, for the $\Box \underline{X}^1$-norm bound in (4-77), we apply (8-15), (8-16), (8-32) to $\mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}_x A} B$, and (8-18), (8-19) and (8-40) to $\mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}^\perp A} B$, $\mathrm{Rem}^{\kappa,2}_{A_0} B$. Finally, (4-78) follows from (8-15) and (8-18).

<u>Step 2</u>: proofs of (4-79), (4-80) and (4-81). The term $\mathrm{Rem}^{\kappa,2}_{A_0} B$ can be put in $\mathrm{Rem}^{\kappa,2}_{A,\mathrm{large}} B$, since for each triple (k, k_1, k_2) within the range $k_1 \geq k_2 - \kappa$, by (8-23) we have

$$\begin{aligned} \| P_k [P_{k_1} A_0, P_{k_2} \partial_t B]\|_{L^1 L^2[I]} &= \| P_k \mathcal{O}(\chi_I P_{k_1} A_0, \chi_I P_{k_2} \partial_t B)\|_{L^1 L^2} \\ &\lesssim 2^{k_2 - k_1} \| P_k \mathcal{O}(\chi_I |D| P_{k_1} A_0, \chi_I |D|^{-1} P_{k_2} \partial_t B)\|_{L^1 L^2} \\ &\lesssim 2^\kappa 2^{-\delta_1 (k_{\max} - k_{\min})} \| P_{k_1} A_0\|_{L^2 \dot{H}^{3/2}[I]} \| P_{k_2} B\|_{DS^1[I]}. \end{aligned}$$

Similarly, the term $\mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}^\perp A} B$ can be put in $\mathrm{Rem}^{\kappa,2}_{A,\mathrm{large}} B$. Moreover, the contributions of these two terms to (4-81) are clearly acceptable, since they need not gain any small factor.

It remains to handle the term $\mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}_x A} B$. We proceed differently according to the length of I. If $|I| \leq 2^{-k + C\kappa}$, we define

$$\mathrm{Rem}^{\kappa,2}_{A,\mathrm{small}} B = \sum_{\substack{k,k_1,k_2 : k_1 \geq k_2 - \kappa \\ \max\{|k_1 - k_2|, |k_1 - k|\} \geq C_0 \kappa}} 2 P_k [\boldsymbol{P}_\ell P_{k_1} A, \partial^\ell P_{k_2} B],$$

and if $|I| \geq 2^{-k+C\kappa}$, we define

$$\mathrm{Rem}^{\kappa,2}_{A,\mathrm{small}} B = \sum_{\substack{k,k_1,k_2: k_1 \geq k_2 - \kappa, \\ \max\{|k_1-k_2|,|k_1-k|\} \geq C_0\kappa}} 2P_k[\boldsymbol{P}_\ell P_{k_1} A, \partial^\ell P_{k_2} B] + \sum_{\substack{k,k_1,k_2 \\ \max\{|k_1-k_2|,|k_1-k|\} < C_0\kappa}} 2P_k Q_{<k_{\min}-C_0\kappa}[\boldsymbol{P}_\ell P_{k_1} Q_{<k_{\min}-C_0\kappa} A, \partial^\ell P_{k_2} Q_{<k_{\min}-C_0\kappa} B].$$

In both cases, we put the remainder $\mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}_x A} B - \mathrm{Rem}^{\kappa,2}_{A,\mathrm{small}} B$ in $\mathrm{Rem}^{\kappa,2}_{A,\mathrm{large}} B$.

Choosing $C_0 > 0$ large enough (depending on δ_1), it follows from Lemma 8.7, (8-21) and (8-24) that $\mathrm{Rem}^{\kappa,2}_{A,\mathrm{small}} B$ obeys the desired bound (4-79); this bound is also acceptable for (4-81). On the other hand, the contribution of $\mathrm{Rem}^{\kappa,2}_{\boldsymbol{P}_x A} B - \mathrm{Rem}^{\kappa,2}_{A,\mathrm{small}} B$ in (4-80) and (4-81) can be handled by proceeding as in Steps 2.2–2.3 and 3 in proof of Proposition 4.15; for the details, we refer to [Oh and Tataru 2018, Proof of Proposition 4.6]. □

8C3. *Estimates for* $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}^\perp A} B$ *and high-modulation estimates for* $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A} B$. Next, we prove Propositions 4.21 and 4.22, which mainly concern the $X^{-1/2+b_1,-b_1} \cap \Box \underline{X}^1$-norms of $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}^\perp A} B$ and $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A} B$.

Proof of Proposition 4.21. We extend B by homogeneous waves outside I, and $\boldsymbol{P}^\perp A$ by zero outside I. Note that

$$\|D\boldsymbol{P}^\perp A\|_Y \lesssim \|\boldsymbol{P}^\perp A\|_{Y^1[I]}, \quad \|B\|_{S^1} \lesssim \|B\|_{S^1[I]}. \tag{8-62}$$

To prove (4-82), we need to estimate the $X^{-1/2+b_1,-b_1} \cap \Box \underline{X}^1$-norm of $\chi_I \mathrm{Diff}^{\kappa}_{\boldsymbol{P}^\perp A} B$. We may write

$$\chi_I \mathrm{Diff}^{\kappa}_{\boldsymbol{P}^\perp A} B = \sum_k 2[P_{<k-\kappa} \boldsymbol{P}^\perp_\ell A, \chi_I \partial^\ell P_k A] = \sum_k 2^k \mathcal{O}(P_{<k-\kappa} \boldsymbol{P}^\perp A, \chi_I P_k A).$$

Then by (8-18), (8-19), (8-40) and (8-42), as well as (8-62), we obtain (4-82). On the other hand, (4-83) simply follows from Hölder's inequality $L^1 L^\infty \times L^\infty L^2 \to L^1 L^2$. □

Proof of Proposition 4.22. We extend A, B by homogeneous waves outside I, and A_0 by zero outside I. In addition to $\|A\|_{S^1} \lesssim \|A\|_{S^1[I]}$, observe that we have

$$\|DA_0\|_Y \lesssim \|A_0\|_{Y^1[I]}, \quad \|\boldsymbol{P}A\|_{Z^1_{p_0}} \lesssim \|\boldsymbol{P}A\|_{Z^1_{p_0}[I]}, \quad \|\boldsymbol{P}A\|_{\tilde{Z}^1_{p_0}} \lesssim \|\boldsymbol{P}A\|_{\tilde{Z}^1_{p_0}[I]}. \tag{8-63}$$

Moreover, by (4-10), we have

$$\|\chi_I \nabla A\|_S \lesssim \|\nabla A\|_S \lesssim \|A\|_{S^1[I]}, \quad \|\chi_I \nabla B\|_S \lesssim \|\nabla B\|_S \lesssim \|B\|_{S^1[I]}. \tag{8-64}$$

We first prove (4-84), for which we need to estimate the $X^{-1/2+b_1,-b_1} \cap \Box \underline{X}^1$-norm of $\chi_I \mathrm{Diff}^{\kappa}_{A_0} B$. We may write

$$\chi_I \mathrm{Diff}^{\kappa}_{A_0} B = -\sum_k 2[P_{<k-\kappa} A_0, \chi_I \partial_t P_k B] = \sum_k \mathcal{O}(P_{<k-\kappa} A_0, \chi_I P_k \partial_t B).$$

Then by (8-18), (8-19), (8-40) and (8-42), as well as (8-63)–(8-64), we obtain (4-84).

For (4-85), (4-86) and (4-87), by Lemma 8.7, we may write

$$\chi_I \mathrm{Diff}^{\kappa}_{\boldsymbol{P}_x A} B = -\sum_k 2[P_{<k-\kappa} \boldsymbol{P}_\ell A, \chi_I \partial^\ell P_k B] = \sum_k \mathcal{N}(|D|^{-1} P_{<k-\kappa} \boldsymbol{P} A, \chi_I P_k B).$$

By (8-36), (8-37) and (8-39), combined with (8-15), (8-16) and the extension relations (8-63)–(8-64), we obtain the desired estimates. □

8C4. *Estimates for* $\mathrm{Diff}^{\kappa}_{\boldsymbol{P}A} B$. Here we prove Propositions 4.23, 4.26, 4.27, 4.28 and 4.30. Note that, by the estimates proved so far in this subsection, we may now use Proposition 5.4 (see also Remark 5.5).

Before we embark on the proofs, we first establish some bilinear Z^1-norm bounds that will be used multiple times below.

Lemma 8.20. *We have*

$$\|P_k \boldsymbol{P} \mathcal{M}^2(\chi_I P_{k_1} A, P_{k_2} B)\|_{\square Z^1} \lesssim 2^{-\delta_1 |k_1 - k_2|} \|P_{k_1} A\|_{S^1[I]} \|P_{k_2} B\|_{S^1[I]}, \tag{8-65}$$

$$\|P_k \mathcal{M}^2_0(\chi_I P_{k_1} A, P_{k_2} B)\|_{L^1 L^\infty} \lesssim 2^{-\delta_1 |k_1 - k_2|} \|P_{k_1} A\|_{S^1[I]} \|P_{k_2} B\|_{S^1[I]}, \tag{8-66}$$

$$\|P_k [P_{k_1} \boldsymbol{P}_\ell A, \chi_I \partial^\ell P_{k_2} B]\|_{\square Z^1} \lesssim 2^{-\delta_1 (k_{\max} - k_{\min})} \|P_{k_1} A\|_{\underline{S}^1[I]} \|P_{k_2} B\|_{S^1[I]}, \tag{8-67}$$

$$\|P_k [P_{k_1} G, \chi_I \nabla P_{k_2} B]\|_{\square Z^1} \lesssim 2^{-\delta_1 (k_{\max} - k_{\min})} \|P_{k_1} G\|_{Y^1[I]} \|P_{k_2} B\|_{S^1[I]}. \tag{8-68}$$

Moreover, for $k < k_1 - 10$, *we have*

$$\|(1 - \mathcal{H}_k) P_k \boldsymbol{P} \mathcal{M}^2(\chi_I P_{k_1} A, P_{k_2} B)\|_{\square Z^1} \lesssim 2^{-\delta_1 (k_{\max} - k_{\min})} \|P_{k_1} A\|_{S^1[I]} \|P_{k_2} B\|_{S^1[I]}, \tag{8-69}$$

$$\|(1 - \mathcal{H}_k) P_k \mathcal{M}^2_0(\chi_I P_{k_1} A, P_{k_2} B)\|_{\Delta^{1/2} \square^{1/2} Z^1} \lesssim 2^{-\delta_1 (k_{\max} - k_{\min})} \|P_{k_1} A\|_{S^1[I]} \|P_{k_2} B\|_{S^1[I]}. \tag{8-70}$$

These bounds follow from Lemma 8.7, (8-17), (8-34), (8-35), (8-38) and (8-41), where we use (8-63) and (8-64) to absorb χ_I and return to interval-localized norms. We omit the straightforward details.

Proof of Proposition 4.23. As in the proof of Proposition 4.22, we extend A, B by homogeneous waves outside I, and A_0 by zero outside I. Furthermore, we extend $\boldsymbol{P}^\perp A$ by zero outside I, and denote the extension by G (we emphasize that, in general, G does not coincide with $\boldsymbol{P}^\perp A$ outside I). In addition to (8-63) and (8-64), by Proposition 5.4 (see also Remark 5.5) we have

$$\|A\|_{\underline{S}^1} \lesssim_M 1, \quad \|DA_0\|_{\ell^1 Y} \lesssim_M 1, \quad \|DG\|_{\ell^1 Y} \lesssim_M 1. \tag{8-71}$$

In the case of the $L^2 \dot{H}^{-1/2}$-norm on the left-hand side, (4-89) now follows easily from (8-15) and (8-18). It remains to estimate the N-norm of $\mathrm{Diff}^{\kappa}_{P_{k_0} \boldsymbol{P} A} B$.

By our extension procedure, note that $P_{k_0} A_0$ and $P_{k_0} \boldsymbol{P}_x A$ obey the equations

$$\begin{aligned} \Delta P_{k_0} A_0 &= P_{k_0}\big([\chi_I A^\ell, \partial_t A_\ell] + 2\boldsymbol{Q}(A, \chi_I \partial_t A) + \chi_I \Delta \boldsymbol{A}^3_0(A)\big), \\ \square P_{k_0} \boldsymbol{P}_x A &= P_{k_0} \boldsymbol{P}\big(\boldsymbol{P}\mathcal{M}^2(\chi_I A, A) + 2[A_0, \chi_I \partial_t A] - 2[G_\ell, \chi_I \partial^\ell A] - 2[\boldsymbol{P}_\ell A, \chi_I \partial^\ell A]\big) \\ &\qquad + P_{k_0} \boldsymbol{P}(\chi_I R(A) - \chi_I \mathrm{Rem}^3(A) A). \end{aligned}$$

For the cubic and higher-order nonlinearities, by Theorem 3.5 and Proposition 4.19, we have

$$\|\chi_I P_{k_0} \Delta \boldsymbol{A}_0^3(A)\|_{L^1 L^2} \lesssim_M 1, \tag{8-72}$$

$$\|\chi_I P_{k_0} R(A)\|_{L^1 L^2} \lesssim_M 1, \tag{8-73}$$

$$\|\chi_I P_{k_0} \mathrm{Rem}^3(A) A\|_{L^1 L^2} \lesssim_M 1. \tag{8-74}$$

For the quadratic nonlinearities, we use (8-17) for $[\chi_I A^\ell, \partial_t A_\ell]$ and $\boldsymbol{Q}(A, \chi_I \partial_t A)$, Lemma 8.7 and (8-33) for $\boldsymbol{P}\mathcal{M}^2[\chi_I A, A)$, Lemma 8.7 and (8-38) for $-[\boldsymbol{P}_\ell A, \chi_I \partial^\ell A]$, and (8-41) for $[A_0, \chi_I \partial_t A]$ and $[G_\ell, \chi_I \partial^\ell A]$. Combining these with the cubic and higher-order estimates and the embedding $L^1 L^2 \subseteq \Box Z^1 \cap \Delta^{-1/2} \Box^{1/2} Z^1$, we arrive at

$$\|P_{k_0} A_0\|_{L^1 L^\infty + L^2 \dot{H}^{3/2} \cap \Delta^{-1/2} \Box^{1/2} Z^1} \lesssim_M 1, \tag{8-75}$$

$$\|P_{k_0} \boldsymbol{P}_x A\|_{Z^1} \lesssim_M 1. \tag{8-76}$$

By Lemma 8.7, (8-26), (8-27), (8-28), (8-29) and Hölder's inequality $L^1 L^\infty \times L^\infty L^2 \to L^1 L^2$, it follows that

$$\|P_k \mathrm{Diff}^\kappa_{P_{k_0} A_0} P_{k_2} B\|_N \lesssim \|P_{k_0} A_0\|_{L^1 L^\infty + L^2 \dot{H}^{3/2} \cap \Delta^{-1/2} \Box^{1/2} Z^1} \|DB\|_S,$$

$$\|P_k \mathrm{Diff}^\kappa_{P_{k_0} \boldsymbol{P}_x A} P_{k_2} B\|_N \lesssim \|P_{k_0} \boldsymbol{P}_x A\|_{S^1 \cap Z^1} \|DB\|_S.$$

Thanks to the frequency gap $\kappa \geq 5$, note furthermore that the left-hand sides vanish unless $k = k_2 + O(1)$. This completes the proof of Proposition 4.23. □

Proof of Proposition 4.26. Estimate (4-94) follows easily using Hölder and Bernstein. To prove (4-95), we extend $\boldsymbol{P}A$, B by homogeneous waves outside I, so that $\|P_{k_1} \Box \boldsymbol{P}A\|_{L^1 L^2} \leq \|P_{k_1} \Box \boldsymbol{P}A\|_{L^1 L^2[I]}$ and $\|P_{k_2} B\|_{S^1} \lesssim \|P_{k_2} B\|_{S^1[I]}$. Moreover, by the embedding $L^1 L^2 \subseteq N \cap \Box Z^1$, we have

$$\|P_{k_1} \boldsymbol{P}A\|_{S^1 \cap Z^1} \lesssim \|P_{k_1} \nabla \boldsymbol{P}A(t_0)\|_{L^2} + \|P_{k_1} \Box \boldsymbol{P}A\|_{L^1 L^2[I]}.$$

Then (4-95) follows by Lemma 8.7, (8-26) and (8-28). □

Proof of Proposition 4.27. Here, in addition to the bilinear null forms (Lemma 8.7), we need to use the secondary null structure (Lemma 8.10).

Without loss of generality, we set $t_0 = 0$. We extend B, $B^{(1)}$ and $B^{(2)}$ by homogeneous waves outside I, and then define A_0 and $\boldsymbol{P}A$ by solving (4-96) and (4-97), respectively.[9] In A_0 and $\boldsymbol{P}A$, we separate out the (high × high → low) interaction terms by defining

$$A_0^{hh} = \sum_{\substack{k, k_1, k_2 \\ k < k_1 - 10}} \Delta^{-1} P_k [P_{k_1} B^{(1)\ell}, P_{k_2} \partial_t B^{(2)}_\ell],$$

$$\boldsymbol{P}A^{hh} = \sum_{\substack{k, k_1, k_2 \\ k < k_1 - 10}} \Box^{-1} P_k \boldsymbol{P} [P_{k_1} B^{(1)\ell}, \partial_x P_{k_2} B^{(2)}_\ell],$$

[9] We may put in χ_I on the right-hand sides of (4-96) and (4-97), but it is not necessary.

where $\Box^{-1} f$ refers to the solution to the inhomogeneous wave equation $\Box u = f$ with $(u, \partial_t u)(0) = 0$. We also introduce

$$\mathcal{H}A_0^{hh} = \sum_{\substack{k,k_1,k_2 \\ k<k_1-10}} \Delta^{-1}\mathcal{H}_k P_k [P_{k_1} B^{(1)\ell}, P_{k_2}\partial_t B_\ell^{(2)}],$$

$$\mathcal{H}\boldsymbol{P}A^{hh} = \sum_{\substack{k,k_1,k_2 \\ k<k_1-10}} \Box^{-1}\mathcal{H}_k P_k \boldsymbol{P}[P_{k_1} B^{(1)\ell}, \partial_x P_{k_2} B_\ell^{(2)}].$$

Accordingly, we split

$$\mathrm{Diff}^\kappa_{\boldsymbol{P}A} B = \sum_k \big(2[P_{<k-\kappa}(A_0 - \mathcal{H}A_0^{hh}), \partial^0 P_k B] + 2[P_{<k-\kappa}(\boldsymbol{P}_\ell A - \mathcal{H}\boldsymbol{P}_\ell A^{hh}), \partial^\ell P_k B]\big) \tag{8-77}$$

$$+ \sum_k \big(2[P_{<k-\kappa}\mathcal{H}A_0^{hh}, \partial^0 P_k B] + 2[P_{<k-\kappa}\mathcal{H}\boldsymbol{P}_\ell A^{hh}, \partial^\ell P_k B]\big). \tag{8-78}$$

By Propositions 4.12, 4.15 and Lemma 8.20, we have

$$\|A_0\|_{Y^1_{cd}} + \|A_0 - A_0^{hh}\|_{L^1 L^\infty_{cd}} + \|A_0^{hh}\|_{Y^1_{cd}} + \|A_0^{hh} - \mathcal{H}A_0^{hh}\|_{\Delta^{-1/2}\Box^{1/2} Z^1_{cd}} \lesssim \|B^{(1)}\|_{S^1_c} \|B^{(2)}\|_{S^1_d},$$

$$\|\boldsymbol{P}A\|_{S^1_{cd}} + \|\boldsymbol{P}A^{hh} - \mathcal{H}\boldsymbol{P}A^{hh}\|_{Z^1_{cd}} \lesssim \|B^{(1)}\|_{S^1_c} \|B^{(2)}\|_{S^1_d}.$$

Combining these bounds with Lemma 8.7, (8-26), (8-27), (8-28), (8-29) and Hölder's inequality $L^1 L^\infty \times L^\infty L^2 \to L^1 L^2$, it follows that

$$\Big\| \sum_k [P_{<k-\kappa}(A_0 - \mathcal{H}A_0^{hh}), \partial^0 P_k B] \Big\|_{N_f} \lesssim \|B^{(1)}\|_{S^1_c} \|B^{(2)}\|_{S^1_d} \|B\|_{S^1_e},$$

$$\Big\| \sum_k [P_{<k-\kappa}(\boldsymbol{P}_\ell A - \mathcal{H}\boldsymbol{P}_\ell A^{hh}), \partial^\ell P_k B] \Big\|_{N_f} \lesssim \|B^{(1)}\|_{S^1_c} \|B^{(2)}\|_{S^1_d} \|B\|_{S^1_e},$$

which handles the contribution of (8-77). On the other hand, unraveling the definitions, we may rewrite (8-78) as

$$(8\text{-}78) = \sum \big(Q_{<j-C}\mathcal{O}'(\Delta^{-1} P_k Q_j \mathcal{O}(P_{k_1} Q_{<j-C} B^{(1)}, \partial_0 P_{k_2} Q_{<j-C} B^{(2)}), \partial^0 Q_{<j-C} P_{k_3} B)$$
$$+ Q_{<j-C}\mathcal{O}'(\Box^{-1} P_k Q_j \boldsymbol{P}_\ell \mathcal{O}(P_{k_1} Q_{<j-C} B^{(1)}, \partial_x P_{k_2} Q_{<j-C} B^{(2)}), \partial^\ell Q_{<j-C} P_{k_3} B)\big)$$

for some disposable operators $\mathcal{O}$ and $\mathcal{O}'$, where the summation is taken over the range $\{(k, k_1, k_2, k_3) : k < k_1 - 10,\ k < k_3 - \kappa + 5\}$. By (8-43), it follows that

$$\|(8\text{-}78)\|_{L^1 L^2_f} \lesssim \|B^{(1)}\|_{S^1_c} \|B^{(2)}\|_{S^1_d} \|B\|_{S^1_e},$$

which is acceptable. Finally, for the $L^2 \dot{H}^{-1/2}$-norm of $\mathrm{Diff}^\kappa_{\boldsymbol{P}A} B$, note that (8-15) and the preceding bounds imply

$$\|P_k(\mathrm{Diff}^\kappa_{\boldsymbol{P}A} B)\|_{L^2 \dot{H}^{-1/2}} \lesssim c_{k-\kappa} d_{k-\kappa} e_k,$$

which is better than what we need. □

Proof of Proposition 4.28. As in the preceding proof, we extend B, $B^{(1)}$ and $B^{(2)}$ by homogeneous waves outside I. This time, however, we also extend $\boldsymbol{P}A$ by homogeneous waves outside I. We moreover extend B_0 and $\boldsymbol{P}^{\perp}B^{(1)}$ by zero outside I, where the latter is denoted by $G^{(1)}$. Note that $\boldsymbol{P}A$ solves the equation

$$\Box \boldsymbol{P}A = \boldsymbol{P}\big([\boldsymbol{P}_\ell B^{(1)}, \chi_I \partial^\ell B^{(2)}] + [B_0^{(1)}, \chi_I \partial^0 B^{(2)}] + [G_\ell^{(1)}, \chi_I \partial^\alpha B^{(2)}]\big).$$

By Lemma 8.20 and the frequency envelope bounds (4-100)–(4-101), it follows that

$$\|\boldsymbol{P}A\|_{Z^1_{cd}} \lesssim (\|B^{(1)}\|_{\underline{S}^1_c[I]} + \|(B_0^{(1)}, G^{(1)})\|_{Y^1_c[I]})\|B^{(2)}\|_{S^1_d[I]} \leq 1. \tag{8-79}$$

On the other hand, recall that $\|\boldsymbol{P}A\|_{S^1_a} \leq 1$ by (4-101). Therefore, by Lemma 8.7, (8-26) and (8-28), we have

$$\|\mathrm{Diff}^{\kappa}_{\boldsymbol{P}_x A} B\|_{N_f} \lesssim 1.$$

On the other hand, by (8-15), we also have

$$\|P_k(\mathrm{Diff}^{\kappa}_{\boldsymbol{P}_x A} B)\|_{L^2\dot{H}^{-1/2}} \lesssim a_{k-\kappa} e_k,$$

which is better than what we need. The desired estimate (4-102) follows. □

Proof of Proposition 4.30. We move the problem to the entire real line using the free-wave extension for $\boldsymbol{P}A_x$ and B, and the zero extension for A_0.

The expression $|D|^{-1}[\nabla, \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}]B$ is a translation-invariant bilinear expression in $\boldsymbol{P}A$ and B, whose Littlewood–Paley pieces can be expressed in the form

$$|D|^{-1}[\nabla, \mathrm{Diff}^{\kappa}_{P_{k'}\boldsymbol{P}A}]P_k B = 2^{k'-k}\mathcal{O}(P_{k'}\boldsymbol{P}A_\alpha, \partial^\alpha P_k B), \quad k' < k-\kappa, \tag{8-80}$$

with $\mathcal{O}$ disposable. By (8-9) the spatial part is a null form, so we can rewrite the above expression as

$$2^{-k}\mathcal{N}(P_{k'}\boldsymbol{P}A_x, P_k B) + 2^{k'-k}\mathcal{O}(P_{k'}A_0, P_k\partial_t B).$$

We consider separately the spatial part and the temporal part. For the spatial part we use the bound (8-21) to estimate

$$\|2^{-k}\mathcal{N}(P_{k'}\boldsymbol{P}A_x, P_k B)\|_N \lesssim 2^{-\delta_1|k-k'|}\|P'_k \boldsymbol{P}A\|_{S^1}\|B\|_{S^1},$$

which suffices after summation in $k' < k-\kappa$.

For the temporal part we use instead the bound (8-23), which yields

$$\|2^{k'}\mathcal{O}(P_{k'}A_0, P_k B)\|_{L^1L^2} \lesssim 2^{-\delta_1|k-k'|}\|P'_k DA_0\|_{L^2\dot{H}^{1/2}}\|B\|_{S^1},$$

which again suffices.

The expression $\mathrm{Diff}^{\kappa}_{P_{k'}\boldsymbol{P}A}B - (\mathrm{Diff}^{\kappa}_{P_{k'}\boldsymbol{P}A})^* B$ is easily seen to have the same form as in (8-80), so the same estimate follows. □

8C5. *Estimates involving* $\boldsymbol{W}$. Here we prove Propositions 4.16, 4.17 and 4.29, which involve $\boldsymbol{w}_0^2$ and $\boldsymbol{w}_x^2$.

Proof of Proposition 4.16. By definition (3-29), we have

$$P_k \boldsymbol{w}_0^2(P_{k_1} A, P_{k_2} B, s) = -2 P_k \boldsymbol{W}(P_{k_1} \partial_t A, P_{k_2} \Delta B, s).$$

Applying Lemma 8.2 to the expression on the right-hand side, we have

$$P_k \boldsymbol{W}(P_{k_1} \partial_t A, P_{k_2} \Delta B, s) = -\langle s 2^{2k}\rangle^{-10} \langle s^{-1} 2^{-2k_{\max}}\rangle^{-1} 2^{-2k_{\max}} 2^{2k_2} P_k \boldsymbol{O}(P_{k_1} \partial_t A, P_{k_2} B) \quad (8\text{-}81)$$

for some disposable operator $\boldsymbol{O}$ on $\mathbb{R}^4$. The rest of the proof follows that of Proposition 4.12. First, by (8-51), it follows that

$$\| |D|^{-1} P_k \boldsymbol{w}_0^2(P_{k_1} A, P_{k_2} B, s)\|_{L^2} \lesssim \langle s 2^{2k}\rangle^{-10} \langle s^{-1} 2^{-2k_{\max}}\rangle^{-1} 2^{2(k_{\min}-k_{\max})} 2^{k_2-k} \|P_{k_1} \partial_t A\|_{L^2} \|P_{k_2} B\|_{\dot{H}^1}.$$

From this dyadic bound, the frequency envelope bound (4-52) follows. Indeed, for any $0 < \delta' < 4\delta$ and any δ'-admissible frequency envelopes c, d, we compute

$$\begin{aligned}\langle s 2^{2k}\rangle^{-10} \langle s^{-1} 2^{-2k_{\max}}\rangle^{-1} 2^{-\delta(k_{\max}-k_{\min})} c_{k_1} d_{k_2} &\lesssim \langle s 2^{2k}\rangle^{-10} \langle s^{-1} 2^{-2k_{\max}}\rangle^{-1} 2^{-\frac{1}{2}\delta(k_{\max}-k_{\min})} c_k d_k \\ &\lesssim \langle s 2^{2k}\rangle^{-10} \langle s^{-1} 2^{-2k}\rangle^{-\frac{1}{4}\delta} c_k d_k, \end{aligned} \quad (8\text{-}82)$$

which proves (4-52). The estimate (4-53) follows in a similar manner from (8-51).

Next, extending $\partial_t A$ and B by zero outside I, then applying (8-15) and (8-17), it follows that

$$\| |D|^{-1} P_k \boldsymbol{w}_0^2(P_{k_1} A, P_{k_2} B, s)\|_{L^2 \dot{H}^{-1/2}[I]} \lesssim \langle s 2^{2k}\rangle^{-10} \langle s^{-1} 2^{-2k_{\max}}\rangle^{-1} 2^{-\delta_1(k_{\max}-k_{\min})} 2^{2(k_1-k_{\max})} \|P_{k_1} A\|_{\mathrm{Str}^1[I]} \|P_{k_2} B\|_{\mathrm{Str}^1[I]},$$

$$\| |D|^{-2} P_k \boldsymbol{w}_0^2(P_{k_1} A, P_{k_2} B, s)\|_{L^1 L^\infty[I]} \lesssim \langle s 2^{2k}\rangle^{-10} \langle s^{-1} 2^{-2k_{\max}}\rangle^{-1} 2^{2(k_1-k_{\max})} \|P_{k_1} A\|_{S^1[I]} \|P_{k_2} B\|_{S^1[I]}.$$

Using (4-21) and (8-54), these two bounds imply (4-54) and (4-55), as in the proof of Proposition 4.12, Step 2. ☐

Proof of Proposition 4.17. We begin with algebraic observations. By (3-30), we have

$$\begin{aligned} P_k \boldsymbol{P}_j \boldsymbol{w}^2(P_{k_1} A, P_{k_2} B, s) = &-2 P_k \boldsymbol{P}_j \boldsymbol{W}(P_{k_1} \partial_t A^\ell, \partial_x P_{k_2} \partial_t B_\ell, s) \\ &+ 4 P_k \boldsymbol{P}_j \boldsymbol{W}(P_{k_1} \boldsymbol{P} \partial_t A^\ell, \partial_\ell P_{k_2} \partial_t B, s) \\ &+ 4 P_k \boldsymbol{P}_j \boldsymbol{W}(P_{k_1} \boldsymbol{P}^\perp \partial_t A^\ell, \partial_\ell P_{k_2} \partial_t B, s), \end{aligned} \quad (8\text{-}83)$$

where, by Lemma 8.2, we may write

$$P_k \boldsymbol{P}_j \boldsymbol{W}(P_{k_1} \partial_t A^\ell, \partial_x P_{k_2} \partial_t B_\ell, s) = \langle s 2^{2k}\rangle^{-10} \langle s^{-1} 2^{-2k_{\max}}\rangle^{-1} 2^{-2k_{\max}} P_k \boldsymbol{P}_j \boldsymbol{O}(P_{k_1} \partial_t A^\ell, \partial_x P_{k_2} \partial_t B_\ell), \quad (8\text{-}84)$$

$$P_k \boldsymbol{P}_j \boldsymbol{W}(P_{k_1}\partial_t \boldsymbol{P}A^{\ell}, \partial_{\ell} P_{k_2}\partial_t B, s) = -2\langle s2^{2k}\rangle^{-10}\langle s^{-1}2^{-2k_{\max}}\rangle^{-1}2^{-2k_{\max}} P_k \boldsymbol{O}(\boldsymbol{P}_{\ell} P_{k_1}\partial_t A, \partial^{\ell} P_{k_2}\partial_t B), \quad (8\text{-}85)$$

$$P_k \boldsymbol{P}_j \boldsymbol{W}(P_{k_1}\partial_t \boldsymbol{P}^{\perp}A^{\ell}, \partial_{\ell} P_{k_2}\partial_t B, s) = -2\langle s2^{2k}\rangle^{-10}\langle s^{-1}2^{-2k_{\max}}\rangle^{-1}2^{-2k_{\max}} P_k \boldsymbol{O}(P_{k_1}\partial_t \boldsymbol{P}_{\ell}^{\perp} A, \partial^{\ell} P_{k_2}\partial_t B) \quad (8\text{-}86)$$

for some disposable operator $\boldsymbol{O}$ on $\mathbb{R}^4$. Note that (8-84) and (8-85) are null forms according to Lemma 8.7, and (8-86) is favorable since $\partial_t \boldsymbol{P}^{\perp} A$ is controlled in the $L^2 \dot{H}^{1/2}$-norm.

Given the above formulas for $\boldsymbol{w}_x$, the proof of the estimates (4-56) and (4-57) is almost identical to the proof of (4-52) and (4-53), using the dyadic bounds (8-51), (8-51) and (8-82).

We now prove (4-58). We extend A, B by homogeneous waves outside I. By (8-15), (8-16), Lemma 8.7, (8-21) and (8-32), it follows that

$$\|P_k \boldsymbol{P}_j \boldsymbol{W}(P_{k_1}\partial_t A, \partial_x P_{k_2}\partial_t B, s)\|_{N\cap \Box \underline{X}^1} \lesssim \langle s2^{2k}\rangle^{-10}\langle s^{-1}2^{-2k_{\max}}\rangle^{-1}2^{-\delta_1(k_{\max}-k_{\min})}2^{k_1+k_2-2k_{\max}}\|P_{k_1}A\|_{S^1}\|P_{k_2}B\|_{S^1},$$

$$\|P_k \boldsymbol{P}_j \boldsymbol{W}(P_{k_1}\partial_t \boldsymbol{P}A, \partial_x P_{k_2}\partial_t B, s)\|_{N\cap \Box \underline{X}^1} \lesssim \langle s2^{2k}\rangle^{-10}\langle s^{-1}2^{-2k_{\max}}\rangle^{-1}2^{-\delta_1(k_{\max}-k_{\min})}2^{k+k_2-2k_{\max}}\|P_{k_1}A\|_{S^1}\|P_{k_2}B\|_{S^1},$$

$$\|P_k \boldsymbol{P}_j \boldsymbol{W}(P_{k_1}\partial_t \boldsymbol{P}^{\perp}A, \partial_x P_{k_2}\partial_t B, s)\|_{N\cap \Box \underline{X}^1} \lesssim \langle s2^{2k}\rangle^{-10}\langle s^{-1}2^{-2k_{\max}}\rangle^{-1}2^{-\delta_1(k_{\max}-k_{\min})}2^{2k_2-2k_{\max}}\|P_{k_1}\partial_t \boldsymbol{P}^{\perp}A\|_{L^2\dot{H}^{1/2}}\|P_{k_2}B\|_{S^1}.$$

Clearly, $2^{k_1+k_2-2k_{\max}}$, $2^{k+k_2-2k_{\max}}$ and $2^{2k_2-2k_{\max}}$ are bounded, so they may be safely discarded. By the same frequency envelope computation (8-82) as before, we obtain (4-58).

In the energy-dispersed case (4-59), we proceed as in the proofs of Propositions 4.15 and 4.20. The contribution of (8-86) is already acceptable, since we need not gain any smallness factor. Moreover, for the contribution of (8-84) and (8-85), the case of $L^2\dot{H}^{-1/2}$ on the left-hand side can be easily handled using (8-15) and (4-21); we omit the details.

It remains to consider only the N-norm of (8-84) and (8-85). For a parameter $\kappa > 0$ to be chosen below, the preceding proof of (4-58) implies that in the case $k_{\max} - k_{\min} \geq \kappa$, we have

$$\|(8\text{-}84)\|_N + \|(8\text{-}85)\|_N \lesssim \langle s2^{2k}\rangle^{-10}\langle s^{-1}2^{-2k_{\max}}\rangle^{-1}2^{-\frac{1}{2}\delta_1\kappa}2^{-\frac{1}{2}\delta_1(k_{\max}-k_{\min})}\|P_{k_1}A\|_{S^1}\|P_{k_2}B\|_{S^1}.$$

On the other hand, when $k_{\max} - k_{\min} \leq \kappa$, we may apply Lemma 8.7 (in particular, (8-13) and (8-14)) and Remark 8.19, which implies

$$\|(8\text{-}84)\|_N + \|(8\text{-}85)\|_N \lesssim \langle s2^{2k}\rangle^{-10}\langle s^{-1}2^{-2k_{\max}}\rangle^{-1}2^{C\kappa}\varepsilon^{c\delta_1}\|P_{k_1}A\|_{\underline{S}^1}M.$$

Choosing $2^{\kappa} = \varepsilon^{c}$ for a sufficiently small $c > 0$, and performing a similar frequency envelope computation as in (8-82), we arrive at (4-59). ☐

Proof of Proposition 4.29. We first note that both $\boldsymbol{w}_0$ and $\boldsymbol{w}_x$ depend on $\partial_t B_1$, for which we control $\|\partial_t B_1\|_{S_c}$ and $\|\boldsymbol{P}^{\perp}\partial_t B_1\|_{Y_c}$. We may assume that

$$\|\partial_t B^{(1)}\|_{S_c[I]},\ \|\boldsymbol{P}^{\perp}\partial_t B^{(1)}\|_{Y_c[I]},\ \|B^{(2)}\|_{S^1_d[I]},\ \|B\|_{S^1_e[I]} \leq 1.$$

We can now extend $\partial_t B_1$ by zero outside I, and $B^{(2)}$ and B by free waves. Then the problem is reduced to the similar problem on the real line. We begin with the simpler $L^2 \dot{H}^{-1/2}$ bound. For that we use (4-53) and (4-58) to obtain

$$\|P_k \boldsymbol{w}_0\|_{L^2 \dot{H}^{-1/2}} + \|P_k \boldsymbol{w}_x\|_{N \cap \Box \underline{X}^1} \lesssim \langle s 2^{2k'} \rangle^{-10} \langle s^{-1} 2^{-2k_{\max}} \rangle^{-\delta_2} c_k d_k \tag{8-87}$$

and then conclude with (8-15) and (8-18).

It remains to prove the N bound. We define

$$\begin{aligned} &\mathcal{I}(k', k_1, k_2, k, s) \\ &\quad = \big(-[\Delta^{-1} P_{k'} \boldsymbol{w}_0^2(P_{k_1} B^{(1)}, P_{k_2} B^{(2)}, s), \partial_t P_k B] + [\Box^{-1} P_{k'} \boldsymbol{P}_\ell \boldsymbol{w}_x^2(P_{k_1} B^{(1)}, P_{k_2} B^{(2)}, s), \partial^\ell P_k B]\big), \end{aligned}$$

so that

$$\mathrm{Diff}^{\kappa}_{\boldsymbol{P} A} B = \sum_{k', k_1, k_2, k : k' < k - \kappa} \mathcal{I}(k', k_1, k_2, k)$$

on I. Introducing the shorthand

$$k_{\max} = \max\{k', k_1, k_2\}, \quad k_{\min} = \min\{k', k_1, k_2\}$$

and

$$\alpha(k', k_1, k_2, s) = \langle s 2^{2k'} \rangle^{-10} \langle s^{-1} 2^{-2k_{\max}} \rangle^{-1} 2^{-c\delta_1 (k_{\max} - k_{\min})},$$

we claim that

$$\|\mathcal{I}(k', k_1, k_2, k, s)\|_N \lesssim \alpha(k', k_1, k_2, s) c_{k_1} d_{k_2} e_k. \tag{8-88}$$

This would conclude the proof of the proposition after summation with respect to k_1 and k_2.

We start with a simple observation, namely that we can easily dispense with the high modulations of $\partial_t B_1$ and B_2 using Lemma 8.2, combined with Hölder and Bernstein inequalities and also (8-26) and (8-30). Thus from here on we assume that

$$P_{k_1} \partial_t B^{(1)} = P_{k_1} Q_{<k_1} \partial_t B^{(1)}, \quad P_{k_2} \partial_t B^{(2)} = P_{k_2} Q_{<k_2} \partial_t B^{(2)}.$$

In view of (8-83) and the identity

$$\boldsymbol{w}_0^2(A, B, s) = -2\boldsymbol{W}(\partial_t A, \partial_t^2 B, s) - 2\boldsymbol{W}(\partial_t A, \Box B, s),$$

we may expand

$$\begin{aligned} \mathcal{I}(k', k_1, k_2, k, s) = {} & 2[P_{k'} \Delta^{-1} \boldsymbol{W}(P_{k_1} \partial_t B^{(1)}, \Box P_{k_2} B^{(2)}, s), \partial_t P_k B] \\ & + 4[\Box^{-1} P_{k'} \boldsymbol{P}_\ell \boldsymbol{W}(P_{k_1} \boldsymbol{P} \partial_t B^{(1),m}, \partial_m P_{k_2} \partial_t B^{(2)}, s), \partial^\ell P_k B] \\ & + 4[\Box^{-1} P_{k'} \boldsymbol{P}_\ell \boldsymbol{W}(P_{k_1} \boldsymbol{P}^\perp \partial_t B^{(1),m}, \partial_m P_{k_2} \partial_t B^{(2)}, s), \partial^\ell P_k B] \\ & + 2[\Delta^{-1} P_{k'} \boldsymbol{W}(P_{k_1} \partial_t B^{(1)}, \partial_t P_{k_2} \partial_t B^{(2)}, s), \partial_t P_k B] \\ & - 2[\Box^{-1} P_{k'} \boldsymbol{P}_\ell \boldsymbol{W}(P_{k_1} \partial_t B^{(1),m}, \partial_x P_{k_2} \partial_t B^{(2)}_m, s), \partial^\ell P_k B] \\ = {} & \mathcal{I}_{(1)} + \mathcal{I}_{(2)} + \mathcal{I}_{(3)} + \mathcal{I}_{(4)} + \mathcal{I}_{(5)}. \end{aligned} \tag{8-89}$$

The first term is easily estimated in L^1L^2 using Lemma 8.2 and Hölder and Bernstein inequalities by

$$\begin{aligned}\|\mathcal{I}_{(1)}\|_{L^1L^2} &\lesssim \|P_{k'}\Delta^{-1}\boldsymbol{W}(P_{k_1}\partial_t B^{(1)}, \Box P_{k_2}B^{(2)}, s)\|_{L^1L^\infty}\|\partial_t P_k B]\|_{L^\infty L^2}\\ &\lesssim \langle s2^{2k'}\rangle^{-10}\langle s^{-1}2^{-2k_{\max}}\rangle^{-1}2^{\frac{1}{2}(k_{\min}-k_{\max})}\|\partial_t P_{k_1}B^{(1)}\|_{L^2\dot{W}^{1,8}}\|\Box P_{k_2}B^{(2)}\|_{L^2\dot{H}^{-1/2}}e_k,\end{aligned}$$

which suffices.

To continue, we use (8-23), (8-33) and the embedding $L^1L^2 \subseteq \Box Z^1$, and we have

$$\begin{aligned}\|P_{k'}\boldsymbol{P}_\ell\boldsymbol{W}(P_{k_1}\boldsymbol{P}\partial_t B^{(1)}, \partial_x P_{k_2}\partial_t B^{(2)}, s)\|_{N\cap\Box Z^1} &\lesssim \alpha(k', k_1, k_2, s)c_{k_1}d_{k_2},\\ \|P_{k'}\boldsymbol{P}_\ell\boldsymbol{W}(P_{k_1}\boldsymbol{P}^\perp\partial_t B^{(1)}, \partial_x P_{k_2}\partial_t B^{(2)}, s)\|_{N\cap\Box Z^1} &\lesssim \alpha(k', k_1, k_2, s)c_{k_1}d_{k_2}.\end{aligned}$$

This yields

$$\begin{aligned}\|\Box^{-1}P_{k'}\boldsymbol{P}_\ell\boldsymbol{W}(P_{k_1}\boldsymbol{P}\partial_t B^{(1)}, \partial_x P_{k_2}\partial_t B^{(2)}, s)\|_{S\cap Z^1} &\lesssim \alpha(k', k_1, k_2, s)c_{k_1}d_{k_2},\\ \|\Box^{-1}P_{k'}\boldsymbol{P}_\ell\boldsymbol{W}(P_{k_1}\boldsymbol{P}^\perp\partial_t B^{(1)}, \partial_x P_{k_2}\partial_t B^{(2)}, s)\|_{S\cap Z^1} &\lesssim \alpha(k', k_1, k_2, s)c_{k_1}d_{k_2}.\end{aligned}$$

We use this directly for the next two terms $\mathcal{I}_{(2)}$ and $\mathcal{I}_{(3)}$, arguing in a bilinear fashion. The desired N bound for both is obtained using both (8-26) and (8-30) with $\kappa = 0$.

The final two terms are combined together in a trilinear null form,

$$\mathcal{I}_{(4)} + \mathcal{I}_{(5)} = \mathrm{Diff}^{\kappa}_{\boldsymbol{P}\tilde{A}}B,$$

where

$$\begin{aligned}\tilde{A}_0 &= \Delta^{-1}P_{k'}\boldsymbol{W}(P_{k_1}\partial_t B^{(1)}, \partial_t P_{k_2}\partial_t B^{(2)}, s),\\ \boldsymbol{A}_x &= \Box^{-1}P_{k'}\boldsymbol{P}_\ell\boldsymbol{W}(P_{k_1}\partial_t B^{(1),m}, \partial_x P_{k_2}\partial_t B^{(2)}_m, s).\end{aligned}$$

At this point we have placed ourselves in the same setting as in the proof of Proposition 4.27. Then the same argument applies, with the only difference that, due to Lemma 8.2, we obtain an additional factor of

$$\langle s2^{2k'}\rangle^{-10}\langle s^{-1}2^{-2k_{\max}}\rangle^{-1}2^{-2k_{\max}}2^{k_1+k_2}$$

as needed. Here the factors 2^{k_1} and 2^{k_2} come from one time derivative on $B^{(1)}$ and $B^{(2)}$, respectively, at low modulation. Thus the N bound for $\mathcal{I}_{(4)} + \mathcal{I}_{(5)}$ follows. $\Box$

8C6. *Estimates for* $\mathrm{Rem}^3(A)B$ *and* $\mathrm{Rem}^3_s(A)B$. Finally, we sketch the proof of Proposition 4.19.

Proof of Proposition 4.19. By Hölder and Bernstein inequalities, it suffices to show that the following nonlinear maps are Lipschitz and envelope-preserving:

$$\begin{aligned}&\mathrm{Str}^1 \ni A \to (\boldsymbol{D}\boldsymbol{A}_0, \boldsymbol{D}\boldsymbol{A}) \in L^{2-}\dot{H}^{\frac{1}{2}+} \cap L^{2+}\dot{H}^{\frac{1}{2}-},\\ &\mathrm{Str}^1 \ni A \to \boldsymbol{A}_0 \in L^2\dot{H}^{\frac{3}{2}}.\end{aligned}$$

The same applies for the maps

$$\begin{aligned}&\mathrm{Str}^1 \ni A \to \boldsymbol{D}\boldsymbol{A}_{0,s} \in L^{2-}\dot{H}^{\frac{1}{2}+} \cap L^{2+}\dot{H}^{\frac{1}{2}-},\\ &\mathrm{Str}^1 \ni A \to \boldsymbol{A}_{0;s} \in L^2\dot{H}^{\frac{3}{2}},\end{aligned}$$

with the addition that now the output has to be also concentrated at frequency $k(s)$.

The $\boldsymbol{A}_0$ property is a consequence of (4-30) for the quadratic term, and (3-23) for the cubic part $\boldsymbol{A}_0^3$. Similarly, the $A_{0;s}$ property is a consequence of (4-53) for the quadratic term, and (3-36) for the cubic part $\boldsymbol{A}_{0;s}^3$.

The $\boldsymbol{DA}$ property follows from (a minor variation of) (4-36) for the quadratic part, and (3-18) for the cubic part $\boldsymbol{DA}^3$.

Finally, the $\boldsymbol{DA}_0$ property is a consequence of (a small variation of) (4-30) for the quadratic part and of (3-24) for the cubic part. Similarly, for $\boldsymbol{DA}_0^s$ we need (a small variation of) (4-53) and of (3-37). □

8D. *Proof of the global-in-time dyadic estimates.* In this subsection, we prove the global-in-time dyadic estimates stated in Section 8B.

8D1. *Preliminaries on orthogonality.* Let $\mathcal{O}$ be a translation-invariant bilinear operator on $\mathbb{R}^{1+4}$. Consider the expression

$$\iint u^{(0)}\mathcal{O}(u^{(1)}, u^{(2)})\, dt\, dx. \tag{8-90}$$

Our general strategy for proving the dyadic estimates stated in Section 8B will be as follows: decompose $u^{(i)}$ by frequency projection into various sets, estimate each such piece, and exploit vanishing (or orthogonality) properties of (8-90), which depend on the relative configuration of the frequency supports of $u^{(i)}$'s, to sum up. Some simple examples of orthogonality properties of (8-90) that we will use are as follows:

Littlewood–Paley trichotomy: If $u^{(i)} = P_{k_1} u^{(i)}$, then (8-90) vanishes unless the largest two numbers of k_0, k_1, k_2 are part by at most (say) 5. This property has already been used freely.

Cube decomposition: If $u^{(i)} = P_{k_i} P_{\mathcal{C}^i} u^{(i)}$ with $\mathcal{C}^i = \mathcal{C}_{k_{\min}}(0)$ (i.e., is a cube of dimension $2^{k_{\min}} \times \cdots \times 2^{k_{\min}}$) situated in $\{|\xi| \simeq 2^{k_i}\}$, then (8-90) vanishes unless $\mathcal{C}^0 + \mathcal{C}^1 + \mathcal{C}^2 \ni 0$.

To obtain more useful statements, let $\mathcal{C}^{\max}$, $\mathcal{C}^{\mathrm{med}}$ and $\mathcal{C}^{\min}$ denote the reindexing of the cubes $\mathcal{C}^0$, $\mathcal{C}^1$ and $\mathcal{C}^2$, which are situated at the annuli $\{|\xi| \simeq 2^{k_{\max}}\}$, $\{|\xi| \simeq 2^{k_{\mathrm{med}}}\}$ and $\{|\xi| \simeq 2^{k_{\min}}\}$, respectively. Then for every fixed $\mathcal{C}^{\min}$ and $\mathcal{C}^{\max}$ (resp. $\mathcal{C}^{\mathrm{med}}$), there are only $O(1)$-many cubes $\mathcal{C}^{\mathrm{med}}$ (resp. $\mathcal{C}^{\max}$) satisfying $\mathcal{C}^{\min} + \mathcal{C}^{\mathrm{med}} + \mathcal{C}^{\max} \ni 0$. Moreover, we have

$$|\angle(\mathcal{C}^{\max}, -\mathcal{C}^{\mathrm{med}})| \lesssim 2^{k_{\max} - k_{\min}}.$$

Geometrically, such cubes $\mathcal{C}^{\max}$ and $\mathcal{C}^{\mathrm{med}}$ are "nearly antipodal".

We will also exploit the relationship between modulation localization and angular restriction for (8-90). In the proofs below, we will only need the following simple statement. For a more complete discussion, see, e.g., [Tao 2001].

Lemma 8.21 (geometry of the cone). *Consider integers $k_0, k_1, k_2, j_0, j_1, j_2 \in \mathbb{Z}$ such that $|k_{\mathrm{med}} - k_{\max}| \le 5$. For $i = 0, 1, 2$, let $\omega_i \subseteq \mathbb{S}^3$ be an angular cap of radius $r_i < 2^{-5}$, $\pm_i \in \{+, -\}$, and $u^{(i)} \in \mathcal{S}(\mathbb{R}^{1+4})$ have frequency support in the region $\{|\xi| \simeq 2^{k_i},\ \xi/|\xi| \in \omega_i,\ |\tau - \pm_i |\xi|| \simeq 2^{j_i}\}$. Suppose that $j_{\max} \le k_{\min}$, and define $\ell = \frac{1}{2}\min\{j_{\max} - k_{\min}, 0\}$.*

Then the expression (8-90) vanishes unless

$$|\angle(\pm_i \omega_i, \pm_{i'} \omega_{i'})| \lesssim 2^{k_{\min} - \min\{k_i, k_{i'}\}} 2^{\ell} + \max\{r_i, r_{i'}\}$$

for every pair $i, i' \in \{0, 1, 2\}$ $(i \neq i')$.

Finally, we collect some often-used estimates. For $k' \leq k$ and $\ell' < -5$, note that

$$2^{-\frac{5}{6}k} \| P_{\mathcal{C}_{k'}(\ell')} u_k \|_{L^2 L^6} + 2^{-k' - \frac{1}{2}k} 2^{-\frac{1}{2}\ell'} \| P_{\mathcal{C}_{k'}(\ell')} u_k \|_{L^2 L^\infty} \lesssim \| P_{\mathcal{C}_{k'}(\ell')} u_k \|_{S_k[\mathcal{C}_{k'}(\ell')]},$$

where, by (4-1), we have

$$\sum_{\mathcal{C} \in \{\mathcal{C}_{k'}(\ell')\}} \| P_{\mathcal{C}} u_k \|_{S_k[\mathcal{C}]}^2 \lesssim \| u_k \|_{S_k}^2 \simeq \| u_k \|_S^2.$$

Also note that, for any $j \leq k + 2\ell$, we have

$$\sum_{\omega} \| P_\ell^\omega Q_{<j} u_k \|_{L^\infty L^2}^2 \lesssim \| u_k \|_{S_k}^2 \simeq \| u_k \|_S^2,$$

by disposing of $Q_{<j}$ (using boundedness on $L^\infty L^2$) and using $S_k^{\mathrm{ang}} \supseteq S_k$.

8D2. *Bilinear estimates that do not involve any null forms.* We first prove Proposition 8.11, which does not involve any null forms.

Proof of Proposition 8.11. In this proof, we adopt the convention of writing $L^p L^{q+}$ for $L^p L^{\tilde{q}}$ with $\tilde{q}^{-1} = q^{-1} - \delta_0$. In particular, if (p, q) is a sharp Strichartz exponent with $\delta_0 \leq p^{-1} \leq \frac{1}{2} - \delta_0$, then $2^{(1/p + 4/q - 2 - 4\delta_0)k} \mathrm{Str}_k^0 \subseteq L^p L^{q+}$.

To prove (8-15), we apply Hölder and Bernstein (on the lowest-frequency factor), where we put u_{k_1} in $L^{9/4} L^{(54/11)+}$ and v_{k_2} in $L^{18} L^{(27/13)+}$. The proof of (8-16) is similar, except we put v_{k_2} in $L^9 L^{(54/23)+}$. The proofs of (8-18) and (8-19) are similar; for (8-18), we apply Hölder and Bernstein with u_{k_1} in $L^2 L^\infty$ and v_{k_2} in $L^\infty L^2$, and for (8-19) we put v_{k_2} in $L^{18} L^{27/13}$ instead.

It only remains to establish (8-17) and (8-20). First, (8-20) follows simply by applying Hölder and Bernstein (on the lowest-frequency factor), where we put u_{k_1}, v_{k_2} in $L^2 L^6$. To prove (8-17), we divide into two cases. When $k \geq k_1 - 10$, the desired bound follows by Hölder, where we put both u_{k_1} and v_{k_2} in $L^2 L^\infty$. On the other hand, when $k < k_1 - 10$, we have $k = k_{\min}$ and $k_1 = k_2 + O(1)$ by Littlewood–Paley trichotomy. We decompose the inputs and the output by frequency projections to cubes of the form $\mathcal{C}_k(0)$, i.e.,

$$P_k \mathcal{O}(u_{k_1}, v'_{k_2}) = \sum_{\mathcal{C}, \mathcal{C}^1, \mathcal{C}^2} P_k P_{\mathcal{C}} \mathcal{O}(P_{\mathcal{C}^1} u_{k_1}, P_{\mathcal{C}^2} v'_{k_2}),$$

where $\mathcal{C}, \mathcal{C}^1, \mathcal{C}^2 \in \{\mathcal{C}_k(0)\}$. The summand on the right-hand side vanishes except when $-\mathcal{C} + \mathcal{C}^1 + \mathcal{C}^2 \ni 0$. For a pair $\mathcal{C}$ and $\mathcal{C}^1$ (resp. $\mathcal{C}^2$), there are only $O(1)$-many $\mathcal{C}^2$ (resp. $\mathcal{C}^1$) such that the preceding condition holds. Moreover, there are only $O(1)$-many $\mathcal{C}$ in the annulus $\{|\xi| \simeq 2^k\}$. Therefore, by Hölder and

Cauchy–Schwarz (in $\mathcal{C}^1$ and $\mathcal{C}^2$), we have

$$2^{-2k}\|P_k\mathcal{O}(u_{k_1},v'_{k_2})\|_{L^1L^\infty}\lesssim 2^{-2k}\Big(\sum_{\mathcal{C}^1}\|P_{\mathcal{C}^1}u_{k_1}\|^2_{L^2L^\infty}\Big)^{\frac12}\Big(\sum_{\mathcal{C}^2}\|P_{\mathcal{C}^2}v'_{k_2}\|^2_{L^2L^\infty}\Big)^{\frac12} \lesssim \|Du_{k_1}\|_S\|v'_{k_2}\|_S,$$

which completes the proof. □

8D3. *Bilinear null form estimates for the N-norm.* We now prove Proposition 8.12. We start with a lemma quantifying the gain from the null form $\mathcal{O}(\partial^\alpha(\cdot),\partial_\alpha(\cdot))$, which is a quick consequence of Lemmas 8.7 and 8.21.

Lemma 8.22. *Let k,k_1,k_2,j,j_1,j_2 satisfy $k_{\max}-k_{\text{med}}\le 5$, $j,j_1,j_2\le k_{\min}+C_0$, $j_1=j+O(1)$ and $j_2=j+O(1)$. Define $\ell=\min\{(j-k_{\min})/2,0\}$, and let $\mathcal{C},\mathcal{C}^1,\mathcal{C}^2$ be rectangular boxes of the form $\mathcal{C}_{k_{\min}}(\ell)$. Then we have*

$$P_kQ_{<j}P_{\mathcal{C}}\mathcal{O}(\partial^\alpha Q_{<j_1}P_{\mathcal{C}^1}u_{k_1},\partial_\alpha Q_{<j_2}P_{\mathcal{C}^2}v_{k_2})=C2^{2\ell}P_{\mathcal{C}}\widetilde{\mathcal{O}}(\nabla P_{\mathcal{C}^1}u_{k_1},\nabla P_{\mathcal{C}^2}v_{k_2}) \tag{8-91}$$

for some universal constant C and a disposable operator $\widetilde{\mathcal{O}}$.

Proof. By disposability of $P_kQ_{<j}P_{\mathcal{C}}$, $P_{k_1}Q_{<j_1}P_{\mathcal{C}^1}$ and $P_{k_2}Q_{<j_2}P_{\mathcal{C}^2}$, we may harmlessly assume that (say) $j,j_1,j_2<k_{\min}-5$. Then we can take the decomposition

$$P_kQ_{<j}P_{\mathcal{C}}\mathcal{O}(\partial^\alpha Q_{<j_1}P_{\mathcal{C}^1}u_{k_1},\partial_\alpha Q_{<j_2}P_{\mathcal{C}^2}v_{k_2})=\sum_{\pm,\pm_1,\pm_2}P_kQ^{\mp}_{<j}P_{\mathcal{C}}\mathcal{O}(\partial^\alpha Q^{\pm_1}_{<j_1}P_{\mathcal{C}^1}u_{k_1},\partial_\alpha Q^{\pm_2}_{<j_2}P_{\mathcal{C}^2}v_{k_2}).$$

By Lemma 8.21, the summand on the right-hand side vanishes (and thus (8-91) holds trivially) unless $|\angle(\pm_1\mathcal{C}^1,\pm_2\mathcal{C}^2)|\lesssim 2^\ell$. In such a case, (8-91) follows from the decompositions (8-11) in Lemma 8.7 and the schematic identities

$$\begin{aligned}\mathcal{N}_{0,\pm_1\pm_2}(Q^{\pm_1}_{<j_1}P_{\mathcal{C}^1}u_{k_1},Q^{\pm_2}_{<j_2}P_{\mathcal{C}^2}v_{k_2})&=C2^{k_1+k_2}2^{2\ell}\widetilde{\mathcal{O}}(P_{\mathcal{C}^1}u_{k_1},P_{\mathcal{C}^2}v_{k_2}),\\ \mathcal{R}_0(Q^{\pm_1}_{<j_1}P_{\mathcal{C}^1}u_{k_1},Q^{\pm_2}_{<j_2}P_{\mathcal{C}^2}v_{k_2})&=C2^j2^{-\min\{k_1,k_2\}}\widetilde{\mathcal{O}}(\nabla P_{\mathcal{C}^1}u_{k_1},\nabla P_{\mathcal{C}^2}v_{k_2}),\end{aligned}$$

which in turn follow from Definition 8.3 (see also Remark 8.4) and (8-12), respectively. □

Proof of Proposition 8.12. Estimates (8-21) and (8-24) were proved in [Oh and Tataru 2018, Proposition 7.1]. Estimate (8-23) is a simple consequence of Hölder and Bernstein for u'_{k_1}, v_{k_2} or the output, depending on which has the lowest frequency. In the remainder of the proof, we prove (8-22) and (8-25) simultaneously.

Step 1: high-modulation inputs/output. The goal of this step is to prove

$$\|P_k\mathcal{O}(\partial^\alpha u_{k_1},\partial_\alpha v_{k_2})-P_kQ_{<k_{\min}}\mathcal{O}(\partial^\alpha Q_{<k_{\min}}u_{k_1},\partial_\alpha Q_{<k_{\min}}v_{k_2})\|_N \lesssim 2^{\frac{k_{\min}+k_{\max}}{2}}\|\nabla u_{k_1}\|_S\|\nabla v_{k_2}\|_S. \tag{8-92}$$

Note that this step is vacuous for (8-25). Here we do not need the null form, and simply view $\mathcal{O}(\partial^\alpha u_{k_1},\partial^\alpha v_{k_2})$ as $\widetilde{\mathcal{O}}(\nabla u_{k_1},\nabla v_{k_2})$ for some disposable $\widetilde{\mathcal{O}}$.

We begin by reducing (8-92) into an atomic form. For $j, j_1, j_2 \geq k_{\min}$, we claim that

$$\left| \int Q_j w_k \widetilde{\mathcal{O}}(Q_{<j_1} u'_{k_1}, Q_{<j_2} v'_{k_2})\, dt\, dx \right| \lesssim 2^{-\frac{1}{2}j} 2^{k_{\min}} 2^{\frac{1}{2}k_1} \|w_k\|_{X^{0,1/2}_\infty} \|u'_{k_1}\|_S \|v'_{k_2}\|_{L^\infty L^2}. \tag{8-93}$$

Once we prove (8-93), by duality (recall that $N^* = L^\infty L^2 \cap X^{0,1/2}_\infty$) we would have

$$\sum_{j \geq k_{\min}} \|P_k Q_j \mathcal{O}(\partial^\alpha u_{k_1}, \partial_\alpha v_{k_2})\|_N \lesssim 2^{\frac{1}{2}k_{\min}} 2^{\frac{1}{2}k_1} \|\nabla u_{k_1}\|_S \|\nabla v_{k_2}\|_{L^\infty L^2},$$

$$\sum_{j \geq k_{\min}} \|P_k Q_{<k_{\min}} \mathcal{O}(\partial^\alpha Q_j u_{k_1}, \partial_\alpha v_{k_2})\|_N \lesssim 2^{\frac{1}{2}k_{\min}} 2^{\frac{1}{2}k_2} \|\nabla u_{k_1}\|_{X^{0,1/2}_\infty} \|\nabla v_{k_2}\|_S,$$

$$\sum_{j \geq k_{\min}} \|P_k Q_{<k_{\min}} \mathcal{O}(\partial^\alpha Q_{<k_{\min}} u_{k_1}, \partial_\alpha Q_j v_{k_2})\|_N \lesssim 2^{\frac{1}{2}k_{\min}} 2^{\frac{1}{2}k_1} \|\nabla u_{k_1}\|_S \|\nabla v_{k_2}\|_{X^{0,1/2}_\infty},$$

from which (8-92) would follow.

To prove (8-93), we decompose u', v', w by frequency projection to cubes of the form $\mathcal{C}_{k_{\min}}(0)$, i.e.,

$$\int Q_j w_k \widetilde{\mathcal{O}}(Q_{<j_1} u'_{k_1}, Q_{<j_2} v'_{k_2})\, dt\, dx = \sum_{\mathcal{C}^0, \mathcal{C}^1, \mathcal{C}^2} \int Q_j P_{\mathcal{C}^0} w_k \widetilde{\mathcal{O}}(Q_{<j_1} P_{\mathcal{C}^1} u'_{k_1}, Q_{<j_2} P_{\mathcal{C}^1} v'_{k_2})\, dt\, dx,$$

where $\mathcal{C}, \mathcal{C}^1, \mathcal{C}^2 \in \{\mathcal{C}_{k_{\min}}(0)\}$.

Let $\mathcal{C}^{\max}$, $\mathcal{C}^{\text{med}}$ and $\mathcal{C}^{\min}$ denote the reindexing of the boxes $\mathcal{C}^0, \mathcal{C}^1, \mathcal{C}^2$, which are situated at the frequency annuli $\{|\xi| \simeq 2^{k_{\max}}\}$, $\{|\xi| \simeq 2^{k_{\text{med}}}\}$ and $\{|\xi| \simeq 2^{k_{\min}}\}$, respectively. The summand on the right-hand side vanishes unless $\mathcal{C}^{\max} + \mathcal{C}^{\text{med}} + \mathcal{C}^{\min} \ni 0$. For a fixed pair $\mathcal{C}^{\min}$ and $\mathcal{C}^{\max}$ (resp. $\mathcal{C}^{\text{med}}$), this happens only for $O(1)$-many $\mathcal{C}^{\text{med}}$ (resp. $\mathcal{C}^{\max}$). Moreover, note that each $\mathcal{C}^i$ lies within an angular sector of size $O(2^{k_{\min} - k_i})$; hence, $Q_{<j_i} P_{\mathcal{C}^i}$ is disposable ($i = 1, 2$). Thus, by Hölder, Cauchy–Schwarz (in $\mathcal{C}^{\max}$ and $\mathcal{C}^{\text{med}}$) and the fact that there are only $O(1)$-many cubes $\mathcal{C}^{\min}$ situated in $\{|\xi| \simeq 2^{k_{\min}}\}$ (so any ℓ^r-sums over $\mathcal{C}^{\min}$ are equivalent), we have

$$\begin{aligned}
&\left| \sum_{\mathcal{C}^0, \mathcal{C}^1, \mathcal{C}^2} \int Q_j P_{\mathcal{C}^0} w_k \widetilde{\mathcal{O}}(Q_{<j_1} P_{\mathcal{C}^1} u'_{k_1}, Q_{<j_2} P_{\mathcal{C}^2} v'_{k_2})\, dt\, dx \right| \\
&\quad \lesssim \left\| \Big(\sum_{\mathcal{C}^0} \|Q_j P_{\mathcal{C}^0} w_k(t,\cdot)\|_{L^2}^2 \Big)^{\frac{1}{2}} \right\|_{L^2_t} \left\| \Big(\sum_{\mathcal{C}^1} \|P_{\mathcal{C}^1} u'_{k_1}(t,\cdot)\|_{L^\infty}^2 \Big)^{\frac{1}{2}} \right\|_{L^2_t} \left\| \Big(\sum_{\mathcal{C}^2} \|P_{\mathcal{C}^2} v'_{k_2}(t,\cdot)\|_{L^2}^2 \Big)^{\frac{1}{2}} \right\|_{L^\infty_t} \\
&\quad \lesssim \left\| Q_j w_k \right\|_{L^2 L^2} \Big(\sum_{\mathcal{C}^1} \|P_{\mathcal{C}^1} u'_{k_1}\|_{L^2 L^\infty}^2 \Big)^{\frac{1}{2}} \|v'_{k_2}\|_{L^\infty L^2} \\
&\quad \lesssim 2^{-\frac{1}{2}j} 2^{k_{\min}} 2^{\frac{1}{2}k_1} \|w_k\|_{X^{0,1/2}_\infty} \|u'_{k_1}\|_S \|v'_{k_2}\|_{L^\infty L^2},
\end{aligned}$$

as desired.

Step 2: proofs of (8-22) and (8-25). For $j < k_{\min}$ and $\ell = (j - k_{\min})/2$, we claim that

$$\|P_k Q_j \mathcal{O}(\partial^\alpha Q_{<j} u_{k_1}, \partial_\alpha Q_{<j} v_{k_2})\|_N \lesssim 2^{-\frac{1}{2}(j - k_{\min})} 2^{\frac{5}{2}\ell} 2^{\frac{1}{2}k_{\min}} 2^{\frac{1}{2}k_1} \|\nabla u_{k_1}\|_S \|\nabla v_{k_2}\|_S, \tag{8-94}$$

$$\|P_k Q_{\leq j}\mathcal{O}(\partial^\alpha Q_j u_{k_1}, \partial_\alpha Q_{<j} v_{k_2})\|_N \lesssim 2^{-\frac{1}{2}(j-k_{\min})} 2^{\frac{5}{2}\ell} 2^{\frac{1}{2}k_{\min}} 2^{\frac{1}{2}k_2} \|\nabla u_{k_1}\|_S \|\nabla v_{k_2}\|_S, \tag{8-95}$$

$$\|P_k Q_{\leq j}\mathcal{O}(\partial^\alpha Q_{\leq j} u_{k_1}, \partial_\alpha Q_j v_{k_2})\|_N \lesssim 2^{-\frac{1}{2}(j-k_{\min})} 2^{\frac{5}{2}\ell} 2^{\frac{1}{2}k_{\min}} 2^{\frac{1}{2}k_1} \|\nabla u_{k_1}\|_S \|\nabla v_{k_2}\|_S. \tag{8-96}$$

Assuming that these estimates hold, we first conclude the proofs of (8-22) and (8-25). We start with (8-22). By Step 1, it suffices to estimate $P_k Q_{<k_{\min}}\mathcal{O}(\partial^\alpha Q_{<k_{\min}} u_{k_1}, Q_{<k_{\min}} v_{k_2})$. Decomposing the inputs and the output using $Q_{<k_{\min}} = \sum_{j<k_{\min}} Q_j$, and dividing cases according to which has dominant modulation (corresponding to j in the above estimates), (8-22) follows by summing (8-94)–(8-96) over j. To prove (8-25), observe simply that the modulation restrictions of the inputs and the output restricts the j-summation to $j < k_{\min} - \kappa$ in the preceding argument.

It remains to establish (8-94)–(8-96).

Step 2.1: proof of (8-94). Here we provide a detailed proof of (8-94); similar arguments involving orthogonality and the null form gain will be used repeatedly in the remainder of this subsection.

We expand

$$P_k Q_j \mathcal{O}(\partial^\alpha Q_{<j} u_{k_1}, \partial_\alpha Q_{<j} v_{k_2}) = \sum_{\pm_0,\pm_1,\pm_2} \sum_{\mathcal{C}^0,\mathcal{C}^1,\mathcal{C}^2} P_k Q_j^{\mp_0} P_{-\mathcal{C}^0}\mathcal{O}(\partial^\alpha Q^{\pm_1}_{<j} P_{\mathcal{C}^1} u_{k_1}, \partial_\alpha Q^{\pm_2}_{<j} P_{\mathcal{C}^2} v_{k_2}),$$

where $\mathcal{C}^0, \mathcal{C}^1, \mathcal{C}^2 \in \{\mathcal{C}_{k_{\min}}(\ell)\}$. By duality, in order to estimate the summand on the right-hand side, it suffices to bound

$$\int P_k Q_j^{\pm_0} P_{\mathcal{C}^0} w\, \mathcal{O}(\partial^\alpha Q^{\pm_1}_{<j} P_{\mathcal{C}^1} u_{k_1}, \partial_\alpha Q^{\pm_2}_{<j} P_{\mathcal{C}^2} v_{k_2})\, dt\, dx. \tag{8-97}$$

Let $\mathcal{C}^{\max}$, $\mathcal{C}^{\mathrm{med}}$ and $\mathcal{C}^{\min}$ denote the reindexing of the boxes $-\mathcal{C}$, $\mathcal{C}^1$, $\mathcal{C}^2$, which are situated at the frequency annuli $\{|\xi| \simeq 2^{k_{\max}}\}$, $\{|\xi| \simeq 2^{k_{\mathrm{med}}}\}$ and $\{|\xi| \simeq 2^{k_{\min}}\}$, respectively.

Note that (8-97) vanishes unless $\mathcal{C}^0 + \mathcal{C}^1 + \mathcal{C}^2 \ni 0$. Combined with the geometry of the cone (Lemma 8.21) we see that for a fixed $\mathcal{C}^{\max}$ (resp. $\mathcal{C}^{\mathrm{med}}$), (8-97) vanishes except for $O(1)$-many $\mathcal{C}^{\min}$ and $\mathcal{C}^{\mathrm{med}}$ (resp. $\mathcal{C}^{\max}$). By Hölder, Cauchy–Schwarz (in $\mathcal{C}^{\max}$ and $\mathcal{C}^{\mathrm{med}}$) and Lemma 8.22, we obtain

$$\begin{aligned}
\Big|\sum_{\pm_0,\pm_1,\pm_2} \sum_{\mathcal{C}^0,\mathcal{C}^1,\mathcal{C}^2} (8\text{-}97)\Big| &\lesssim \sum_{\pm_0} 2^{2\ell} \Big\| \Big(\sum_{\mathcal{C}^0} \|P_k Q_j^{\pm_0} P_{\mathcal{C}^0} w(t,\cdot)\|_{L^2}^2\Big)^{\frac{1}{2}} \Big\|_{L^2_t} \\
&\qquad \times \Big\| \Big(\sum_{\mathcal{C}^1} \|\nabla P_{\mathcal{C}^1} u_{k_1}(t,\cdot)\|_{L^\infty}^2\Big)^{\frac{1}{2}} \Big\|_{L^2_t} \Big\| \Big(\sum_{\mathcal{C}^2} \|\nabla P_{\mathcal{C}^2} v_{k_2}(t,\cdot)\|_{L^2}^2\Big)^{\frac{1}{2}} \Big\|_{L^\infty_t} \\
&\lesssim \sum_{\pm_0} 2^{2\ell} \|P_k Q_j^{\pm_0} w\|_{L^2 L^2} \Big(\sum_{\mathcal{C}^1} \|\nabla P_{\mathcal{C}^1} u_{k_1}\|_{L^2 L^\infty}^2\Big)^{\frac{1}{2}} \|\nabla v_{k_2}\|_{L^\infty L^2} \\
&\lesssim 2^{-\frac{1}{2}j} 2^{\frac{5}{2}\ell} 2^{k_{\min}} 2^{\frac{1}{2}k_1} \|w\|_{X^{0,1/2}_\infty} \|\nabla u_{k_1}\|_S \|\nabla v_{k_2}\|_{L^\infty L^2}.
\end{aligned}$$

By duality, (8-94) follows.

Steps 1.2–1.3: proofs of (8-95)–(8-96). We now sketch the proofs of (8-95) and (8-96), which are very similar to Step 2.1. As before, we expand each modulation projection to the $\pm$-parts, and decompose the output, u, v by frequency projection to $-\mathcal{C}^0, \mathcal{C}^1, \mathcal{C}^2 \in \{\mathcal{C}_{k_{\min}}(\ell)\}$, respectively.

We proceed as in Step 1.1 but put the test function w in $L^\infty L^2$ and the input with the dominant modulation in $L^2 L^2$. Then we obtain

$$\Big| \sum_{\pm_0,\pm_1,\pm_2} \sum_{\mathcal{C}^0,\mathcal{C}^1,\mathcal{C}^2} \iint P_k Q^{\pm_0}_{\leq j} P_{\mathcal{C}^0} \mathcal{O}(\partial^\alpha Q^{\pm_1}_{j} P_{\mathcal{C}^1} u_{k_1}, \partial_\alpha Q^{\pm_2}_{<j} P_{\mathcal{C}^2} v_{k_2}) \Big| \lesssim 2^{-\frac{1}{2}j} 2^{\frac{5}{2}\ell} 2^{k_{\min}} 2^{\frac{1}{2}k_2} \|w\|_{L^\infty L^2} \|\nabla u_{k_1}\|_{X^{0,1/2}_\infty} \|\nabla v_{k_2}\|_S,$$

$$\Big| \sum_{\pm_0,\pm_1,\pm_2} \sum_{\mathcal{C}^0,\mathcal{C}^1,\mathcal{C}^2} \iint P_k Q^{\pm_0}_{\leq j} P_{\mathcal{C}^0} \mathcal{O}(\partial^\alpha Q^{\pm_1}_{\leq j} P_{\mathcal{C}^1} u_{k_1}, \partial_\alpha Q^{\pm_2}_{j} P_{\mathcal{C}^2} v_{k_2}) \Big| \lesssim 2^{-\frac{1}{2}j} 2^{\frac{5}{2}\ell} 2^{k_{\min}} 2^{\frac{1}{2}k_1} \|w\|_{L^\infty L^2} \|\nabla u_{k_1}\|_S \|\nabla v_{k_2}\|_{X^{0,1/2}_\infty}.$$

By duality, (8-95) and (8-96) follow. □

8D4. *Bilinear estimates for the $X^{s,b,p}_r$-type norms.* Next, we prove Propositions 8.13, 8.14, 8.15 and 8.16.

Proof of Proposition 8.13. Estimates (8-26) and (8-27) were proved in [Krieger et al. 2015, equations (132) and (133)]; note that the slightly stronger S^1-norm is used on the right-hand side in [Krieger et al. 2015, equations (132) and (133)], but the proofs in fact lead to (8-26) and (8-27). Estimates (8-28) and (8-29) follow from slight modifications of the proofs of [Krieger et al. 2015, equations (134) and (140)] (the Z-norm in that paper is stronger than ours), as we outline below.

For (8-28), we first recall the definition of $\mathcal{H}^*$. For each $j < k_1 - C$, we introduce $\ell = \frac{1}{2}(j - k_1)$ and take the decomposition

$$P_k Q_{<j-C} \mathcal{N}(|D|^{-1} Q_j u_{k_1}, Q_{<j-C} v_{k_2}) = \sum_{\omega,\omega'} P_k Q_{<j-C} \mathcal{N}(|D|^{-1} P^\omega_\ell Q_j u_{k_1}, P^\ell_{\omega'} Q_{<j-C} v_{k_2}).$$

By the geometry of the cone (Lemma 8.21), the summand vanishes unless $|\angle(\omega, \pm\omega')| \lesssim 2^\ell$ for some sign $\pm$. In this case, the null form $\mathcal{N}$ gains $2^{k_1+k_2} 2^\ell$ (see Definition 8.3), and hence we have

$$\begin{aligned}
&\|P_k Q_{<j-C} \mathcal{N}(|D|^{-1} Q_j u_{k_1}, Q_{<j-C} v_{k_2})\|_{L^1 L^2} \\
&\quad \lesssim \sum_{\omega,\omega' : \min_\pm |\angle(\omega,\pm\omega')| \lesssim 2^\ell} 2^{k_2} 2^\ell \|P^\omega_\ell Q_j u_{k_1}\|_{L^1 L^\infty} \|P^{\omega'}_\ell Q_{<j-C} v_{k_2}\|_{L^\infty L^2} \\
&\quad \lesssim 2^{k_2} 2^{(\frac{1}{2}-2b_0)\ell} \Big(\sum_\omega (2^{(\frac{1}{2}+2b_0)\ell} \|P^\omega_\ell Q_{k+2\ell} u_{k_1}\|_{L^1 L^\infty})^2 \Big)^{\frac{1}{2}} \Big(\sum_{\omega'} \|P^{\omega'}_\ell Q_{<j-C} v_{k_2}\|^2_{L^\infty L^2} \Big)^{\frac{1}{2}} \\
&\quad \lesssim 2^{(\frac{1}{2}-2b_0)\ell} \Big(\sum_\omega (2^{(\frac{1}{2}+2b_0)\ell} \|P^\omega_\ell Q_{k+2\ell} u_{k_1}\|_{L^1 L^\infty})^2 \Big)^{\frac{1}{2}} \|D v_{k_2}\|_S.
\end{aligned}$$

In the second inequality, we used Cauchy–Schwarz (or Schur's test) with the fact that the ω, ω' is essentially diagonal (i.e., for a fixed ω, there are only $O(1)$ many ω''s such that the sum is nonvanishing, and vice versa). Summing up in $j < k_1 - C$, then using the definition of the Z^1-norm, (8-28) follows.

Next, (8-29) is proved by essentially the same argument (with the same numerology) as above. Here we do not gain 2^ℓ from the null form $\mathcal{N}$, but rather from the extra factor $\Delta^{-1/2}\Box^{1/2}$ in the norm $\Delta^{-1/2}\Box^{1/2} Z^1$. Finally, (8-30) and (8-31) follow from the preceding proofs, once we observe that the modulation localization of u_{k_1} restricts the j-summation to $j < k_1 - \kappa$, which then leads to the small factor $2^{-(1/2-2b_0)\kappa}$. □

Proof of Proposition 8.14. In view of the embedding $N \cap \Box Z^1 \subseteq \Box Z^1_{p_0}$, (8-32) would follow once (8-33) is proved. Estimates (8-34) and (8-35) follow from (134) and (141) in [Krieger et al. 2015], respectively. Moreover, when $k \geq k_1 - C$, (8-33) follows from (134) and (135) in [loc. cit.]. In using the estimates from [loc. cit.], we remind the reader that the Z-norm in [loc. cit.] (which is equal to $\sum_k \| P_k Q_{<k} u\|_{X_\infty^{-1/4,1/4,1}}$) is stronger the Z-norm in this work. Moreover, although (134), (135) and (141) in [loc. cit.] are stated with the S^1-norm on the right-hand side, an inspection of the proof reveals that only the S-norm is used.

It remains to establish (8-33) in the case $k < k_1 - C$. By Littlewood–Paley trichotomy, note that the left-hand side vanishes unless $k = k_{\min}$ and $k_1 = k_2 + O(1)$. By (8-34), we are only left to show that the $\Box Z^1$-norm of

$$P_k \mathcal{H}_k \mathcal{N}(u_{k_1}, v_{k_2}) = \sum_{j<k+C} P_k Q_j \mathcal{N}(Q_{<j-C} u_{k_1}, Q_{<j-C} v_{k_2}) \tag{8-98}$$

is bounded by $\lesssim 2^k \| D u_{k_1}\|_S \| D v_{k_2}\|_S$.

Consider the summand of (8-98). We decompose the inputs and the output by frequency projections to rectangular boxes of the form $\mathcal{C}_k(\ell)$, where $\ell = \min\{(j-k)/2, 0\}$. Then we need to consider the expression

$$P_k Q_j P_{\mathcal{C}} \mathcal{N}(Q_{<j-C} P_{\mathcal{C}^1} u_{k_1}, Q_{<j-C} P_{\mathcal{C}^2} v_{k_2}),$$

where $\mathcal{C}, \mathcal{C}^1, \mathcal{C}^2 \in \{\mathcal{C}_k(\ell)\}$. This expression is nonvanishing only when $-\mathcal{C} + \mathcal{C}^1 + \mathcal{C}^2 \ni 0$. In fact, combined with the geometry of the cone (Lemma 8.21), we see that for each fixed $\mathcal{C}^1$ (resp. $\mathcal{C}^2$), it is nonvanishing only for $O(1)$-many $\mathcal{C}$ and $\mathcal{C}^2$ (resp. $\mathcal{C}^1$). The null form gains the factor $2^{k_1+k_2} 2^{\ell}$. By Hölder and Cauchy–Schwarz (in $\mathcal{C}^1$ and $\mathcal{C}^2$), we have

$$\begin{aligned}
&\| P_k Q_j \mathcal{N}(Q_{<j-C} u_{k_1}, Q_{<j-C} v_{k_2})\|_{\Box Z^1} \\
&\qquad = 2^{-\frac{3}{2}k} 2^{-\frac{1}{2}j} \Big\| \sum_{\mathcal{C},\mathcal{C}^1,\mathcal{C}^2} P_k Q_j P_{\mathcal{C}} \mathcal{N}(Q_{<j-C} P_{\mathcal{C}^1} u_{k_1}, Q_{<j-C} P_{\mathcal{C}^2} v_{k_2}) \Big\|_{L^1 L^\infty} \\
&\qquad \lesssim 2^{-\frac{3}{2}k} 2^{-\frac{1}{2}j} 2^{k_1+k_2} 2^{\ell} \Big(\sum_{\mathcal{C}^1} \| Q_{<j-C} P_{\mathcal{C}^1} u_{k_1}\|^2_{L^2 L^\infty}\Big)^{\frac{1}{2}} \Big(\sum_{\mathcal{C}^2} \| Q_{<j-C} P_{\mathcal{C}^2} v_{k_2}\|^2_{L^2 L^\infty}\Big)^{\frac{1}{2}} \\
&\qquad \lesssim 2^{-\frac{1}{2}(k-j)} 2^k \| D u_{k_1}\|_S \| D v_{k_2}\|_S.
\end{aligned}$$

Summing up in $j < k + C$, the desired estimate follows. □

Proof of Proposition 8.15. For all the estimates, the most difficult case is when $k_1 < k - 10$ (low-high interaction) and when u_{k_1} has the dominant modulation, i.e., the expression $P_k \mathcal{H}^*_{k_1} \mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2})$.

Step 1: proof of (8-36), (8-37) and (8-38). We divide into three cases: (1) $k_1 \geq k - 10$, (2) $k_1 < k - 10$ but either the output or v_{k_2} has the dominant modulation, or (3) $k_1 < k - 10$ and u_{k_1} has the dominant modulation.

Step 1.1: $k_1 \geq k-10$. In this case, all three bounds can be proved simultaneously. The idea is to apply Propositions 8.12 and 8.14. Indeed, by (8-33) and the fact that the left-hand side vanishes unless $k_1 = k_{\max} + O(1)$ (Littlewood–Paley trichotomy), we see that

$$\begin{aligned}\|P_k \mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2})\|_{\Box Z^1} &\lesssim 2^{k-k_1} \|P_k |D|^{-1} \mathcal{N}(u_{k_1}, v_{k_2})\|_{\Box Z^1} \\ &\lesssim 2^{-C\delta_1(k_{\max}-k_{\min})} \|D u_{k_1}\|_S \|D v_{k_2}\|_S.\end{aligned}$$

Combined with (8-21), it follows that

$$\|P_k \mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2})\|_{N \cap \Box Z^1} \lesssim 2^{-C\delta_1(k_{\max}-k_{\min})} \|D u_{k_1}\|_S \|D v_{k_2}\|_S.$$

By the chain of embeddings $N \cap \Box Z^1 \subseteq \Box Z^1_{p_0} \subseteq \Box \widetilde{Z}^1_{p_0}$, the desired bounds follow.

Step 1.2: $k_1 < k-10$, contribution of $1 - \mathcal{H}^*_{k_1}$. By Littlewood–Paley trichotomy, $P_k \mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2})$ vanishes unless $k_1 = k_{\min}$ and $k = k_{\max} + O(1)$. In Steps 1.2a–1.2c below, we estimate the $\Box Z^1$-norm of $P_k(1-\mathcal{H}^*_{k_1})\mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2})$. Then in Step 1.2d, we conclude the proof by interpolating with (8-26).

Step 1.2a: High modulation inputs/output. The goal of this step is to prove

$$\begin{aligned}\|P_k \mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2}) - P_k Q_{<k_1} \mathcal{N}(|D|^{-1} Q_{<k_1+C} u_{k_1}, Q_{<k_1} v_{k_2})\|_{\Box Z^1} \\ \lesssim 2^{-\frac{1}{4}(k-k_1)} \|D u_{k_1}\|_S \|D v_{k_2}\|_S. \qquad \text{(8-99)}\end{aligned}$$

Here there is no need for null structure, so we simply write $\mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2}) = \mathcal{O}(u_{k_1}, D v_{k_2})$. We begin by proving

$$\|P_k Q_{\geq k_1} \mathcal{O}(u_{k_1}, D v_{k_2})\|_{\Box Z^1} \lesssim 2^{-b_0(k-k_1)} \||D|^{-\frac{1}{2}} u_{k_1}\|_{L^2 L^\infty} \|D v_{k_2}\|_S. \qquad \text{(8-100)}$$

For $j \geq k_1$, we take the decomposition

$$P_k Q_j P^{\omega}_{\frac{j-k}{2}} \mathcal{O}(u_{k_1}, D v_{k_2}) = \sum_{\omega'} P_k Q_j P^{\omega}_{\frac{j-k}{2}} \mathcal{O}(u_{k_1}, D P^{\omega'}_{\frac{j-k}{2}} v_{k_2}).$$

Since $(j-k)/2 \geq k_1 - k$, for each fixed ω there are only $O(1)$-many ω' such that the summand on the right-hand side is (possibly) nonvanishing, and vice versa. Therefore, by Hölder, Bernstein and Cauchy–Schwarz, we have

$$\begin{aligned}&2^{(-\frac{3}{4}+b_0)(j-k)} 2^{-2k} \Big(\sum_{\omega} \|P_k Q_j P^{\omega}_{\frac{j-k}{2}} \mathcal{O}(u_{k_1}, D P^{\omega'}_{\frac{j-k}{2}} v_{k_2})\|^2_{L^1 L^\infty}\Big) \\ &\quad\lesssim 2^{(-\frac{1}{2}+b_0)(j-k)} 2^{-\frac{1}{2}(k-k_1)} (2^{-\frac{1}{2}k_1} \|u_{k_1}\|_{L^2 L^\infty}) \Big(\sum_{\omega'} (2^{\frac{1}{6}k_2} \|P^{\omega'}_{\frac{j-k}{2}} v_{k_2}\|_{L^2 L^6})^2\Big)^{\frac{1}{2}} \\ &\quad\lesssim 2^{(-\frac{1}{2}+b_0)(j-k_1)} 2^{-b_0(k-k_1)} \||D|^{-\frac{1}{2}} u_{k_1}\|_{L^2 L^\infty} \|D v_{k_2}\|_S.\end{aligned}$$

Summing up in $j \geq k_1$, we obtain (8-100).

Next, we prove

$$\|P_k Q_{<k_1} \mathcal{O}(u_{k_1}, D Q_{\geq k_1} v_{k_2})\|_{\Box Z^1} \lesssim 2^{-b_0(k-k_1)} \||D|^{-\frac{1}{2}} u_{k_1}\|_{L^2 L^\infty} \|D v_{k_2}\|_S. \qquad \text{(8-101)}$$

By (4-6) and (uniform-in-j) boundedness of Q_j on L^1L^2, we have

$$\|P_k Q_{<k_1} f\|_{\Box Z^1} \lesssim 2^{-b_0(k-k_1)} \|f\|_{L^1L^2}. \tag{8-102}$$

Therefore,

$$\begin{aligned}\|P_k Q_{<k_1}\mathcal{O}(u_{k_1}, DQ_j v_{k_2})\|_{\Box Z^1} &\lesssim 2^{-b_0(k-k_1)}\|P_k Q_{<k_1}\mathcal{O}(u_{k_1}, DQ_j v_{k_2})\|_{L^1L^2} \\ &\lesssim 2^{-\frac{1}{2}(j-k_1)}2^{-b_0(k-k_1)}(2^{-\frac{1}{2}k_1}\|u_{k_1}\|_{L^2L^\infty})\|DQ_j v_{k_2}\|_{X^{0,1/2}_\infty} \\ &\lesssim 2^{-\frac{1}{2}(j-k_1)}2^{-b_0(k-k_1)}\||D|^{-\frac{1}{2}}u_{k_1}\|_{L^2L^\infty}\|Dv_{k_2}\|_S.\end{aligned}$$

Then summing up in $j \geq k_1$, (8-101) follows.

To conclude the proof of (8-99), note that $\||D|^{-1/2}u_{k_1}\|_{L^2L^\infty} \lesssim \|Du_{k_1}\|_S$. Moreover, observe that

$$P_k Q_{<k_1}\mathcal{O}(Q_j u_{k_1}, DQ_{<k_1} v_{k_2})$$

vanishes unless $j < k_1 + 10$.

Step 1.2b: output has dominant modulation. Here we prove

$$\sum_{j<k_1} \|P_k Q_j \mathcal{N}(|D|^{-1}Q_{<j_1}u_{k_1}, Q_{<j_2}v_{k_2})\|_{\Box Z^1} \lesssim 2^{-b_0(k-k_1)}\|Du_{k_1}\|_S \|Dv_{k_2}\|_S, \tag{8-103}$$

where $j_1, j_2 = j + O(1)$.

Let $\ell = \frac{1}{2}(j-k_1)$. After taking the decompositions $u_{k_1} = \sum_{\omega'} P^{\omega'}_\ell u_{k_1}$ and $v_{k_2} = \sum_{\omega''} P^{\omega''}_{(j-k)/2} v_{k_2}$, consider the expression

$$P_k Q_j P^{\omega}_{\frac{j-k}{2}} \mathcal{N}(|D|^{-1}Q_{<j_1}P^{\omega'}_\ell u_{k_1}, Q_{<j_2}P^{\omega''}_{\frac{j-k}{2}} v_{k_2}).$$

Using the geometry of the cone (Lemma 8.21), observe that for every fixed ω (resp. ω''), the preceding expression vanishes except for $O(1)$-many ω' and ω'' (resp. ω). Moreover, for such a triple $\omega, \omega', \omega''$, the null form $\mathcal{N}$ gains a factor of 2^ℓ. By Hölder, Bernstein (for $P^{\omega}_{(j-k)/2}v_{k_2}$) and Cauchy–Schwarz (in ω, ω''), we have

$$\begin{aligned}&\|P_k Q_j \mathcal{N}(|D|^{-1}Q_{<j_1}u_{k_1}, Q_{<j_2}v_{k_2})\|_{\Box Z^1} \\ &\lesssim 2^{(-\frac{3}{4}+b_0)(j-k)}2^{-2k}\Big(\sum_\omega \|P_k Q_j P^\omega_{\frac{j-k}{2}}\mathcal{N}(|D|^{-1}Q_{<j_1}u_{k_1}, Q_{<j_2}v_{k_2})\|^2_{L^1L^\infty}\Big)^{\frac{1}{2}} \\ &\lesssim 2^{(-\frac{1}{2}+b_0)(j-k)}2^\ell 2^{-\frac{1}{2}(k-k_1)}\big(\sup_{\omega'} 2^{-\frac{1}{2}k_1}\|Q_{<j_1}P^{\omega'}_\ell u_{k_1}\|_{L^2L^\infty}\big)\Big(\sum_\omega (2^{\frac{1}{6}k_2}\|Q_{<j_2}P^\omega_{\frac{j-k}{2}}v_{k_2}\|_{L^2L^6})^2\Big)^{\frac{1}{2}} \\ &\lesssim 2^{-b_0(k_1-j)}2^{-b_0(k-k_1)}\|Du_{k_1}\|_S\|Dv_{k_2}\|_S.\end{aligned}$$

Summing up in $j < k_1$, (8-103) follows.

Step 1.2c: v has dominant modulation. Next, we prove

$$\sum_{j<k_1} \|P_k Q_{<j_0} \mathcal{N}(|D|^{-1}Q_{<j_1}u_{k_1}, Q_j v_{k_2})\|_{\Box Z^1} \lesssim 2^{-b_0(k-k_1)}\|Du_{k_1}\|_S\|Dv_{k_2}\|_S, \tag{8-104}$$

where $j_0, j_1 = j + O(1)$. As before, let $\ell = (j - k_1)/2$. By (4-6) and (uniform-in-j) boundedness of Q_j on $L^1 L^2$, we have

$$\| P_k Q_{<j} f \|_{\Box Z^1} \lesssim 2^{-b_0(k-j)} \| f \|_{L^1 L^2}.$$

Hence it suffices to estimate the $L^1 L^2$ norm of the output. This time, we take the decompositions $u_{k_1} = \sum_\omega P_\ell^\omega u_{k_1}$ and $v_{k_2} = \sum_{\omega'} P_\ell^{\omega'} v_{k_2}$. By the geometry of the cone, for a fixed ω, the expression

$$P_k Q_{<j_0} \mathcal{N}(|D|^{-1} Q_{<j_1} P_\ell^\omega u_{k_1}, Q_j P_\ell^{\omega'} v_{k_1})$$

vanishes except for $O(1)$-many ω' and vice versa. Moreover, the null form $\mathcal{N}$ gains a factor of 2^ℓ. By Hölder and Cauchy–Schwarz (in ω, ω'), we have

$$\begin{aligned}
&2^{-b_0(k-j)} \| P_k Q_{<j_0} \mathcal{N}(|D|^{-1} Q_{<j_1} P_\ell^\omega u_{k_1}, Q_j P_\ell^{\omega'} v_{k_2}) \|_{L^1 L^2} \\
&\lesssim 2^{-b_0(k-j)} 2^{\frac{3}{2}\ell} 2^{\frac{1}{2}k_1} 2^{-\frac{1}{2}j} \Big(\sum_\omega (2^{-\frac{1}{2}k_1} 2^{-\frac{1}{2}\ell} \| Q_{<j_1} P_\ell^\omega u_{k_1} \|_{L^2 L^\infty})^2 \Big)^{\frac{1}{2}} \Big(\sum_{\omega'} (2^{k_2} \| Q_j P_\ell^{\omega'} v_{k_2} \|_{X_\infty^{0,1/2}})^2 \Big)^{\frac{1}{2}} \\
&\lesssim 2^{(-\frac{1}{4} - b_0)(k_1 - j)} 2^{-b_0(k-k_1)} \| D u_{k_1} \|_S \| D v_{k_2} \|_S.
\end{aligned}$$

Summing up in $j < k_1$, (8-104) is proved.

Step 1.2d: interpolation with (8-26). Combining (8-99), (8-103) and (8-104), we obtain

$$\| P_k (1 - \mathcal{H}_{k_1}^*) \mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2}) \|_{\Box Z^1} \lesssim 2^{-b_0(k-k_1)} \| D u_{k_1} \|_S \| D v_{k_2} \|_S.$$

On the other hand, (8-26) and the embedding $N \subseteq X_\infty^{0,-1/2}$ yields a similar bound for the $X_\infty^{0,-1/2}$-norm without the exponential gain. Nevertheless, since we have

$$\| f \|_{\Box Z^1_{p_0}} \lesssim \| f \|_{\Box Z^1}^{\theta_0} \| f \|_{X_\infty^{0,-1/2}}^{1-\theta_0},$$

where $\theta_0 = 2\big(\frac{1}{p_0} - \frac{1}{2}\big) > 0$,

$$\| P_k (1 - \mathcal{H}_{k_1}^*) \mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2}) \|_{\Box Z^1_{p_0}} \lesssim 2^{-\theta_0 b_0 (k-k_1)} \| D u_{k_1} \|_S \| D v_{k_2} \|_S.$$

Then the desired estimate for $\Box \widetilde{Z}^1_{p_0}$ follows as well, thanks to the embedding $\Box Z^1_{p_0} \subseteq \Box \widetilde{Z}^1_{p_0}$.

Step 1.3: $k_1 < k - 10$, contribution of $\mathcal{H}_{k_1}^*$. This is the most difficult case. We consider

$$P_k \mathcal{H}_{k_1}^* \mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2}) = \sum_{j < k_1 + C} P_k Q_{<j-C} \mathcal{N}(|D|^{-1} Q_j u_{k_1}, Q_{<j-C} v_{k_2}).$$

As before, by Littlewood–Paley trichotomy, this expression vanishes unless $k_1 = k_{\min}$ and $k = k_{\max} + O(1)$.

Recall that all three norms $\Box \widetilde{Z}^1_{p_0}$, $\Box Z^1_{p_0}$ and $\Box Z^1$ are of the type $X_1^{s,b,p}$. To ensure the ℓ^2-summability in ω in the definition (4-3), we go through the $L^p L^2$ norm. More precisely, by Bernstein and L^2-orthogonality of $P^\omega_{(j-k)/2}$, note that

$$\| P_k Q_j f \|_{X_1^{s,b,p}} \lesssim 2^{sk} 2^{\frac{5}{2}(\frac{1}{p} - \frac{1}{2})k} 2^{bj} 2^{\frac{3}{2}(\frac{1}{p} - \frac{1}{2})j} \| f \|_{L^p L^2}.$$

Since $b + \frac{3}{2}\big(\frac{1}{p} - \frac{1}{2}\big) > 0$ in all of these cases by (4-4), we have

$$\| P_k Q_{<j} f \|_{X_1^{s,b,p}} \lesssim 2^{sk} 2^{\frac{5}{2}(\frac{1}{p} - \frac{1}{2})k} 2^{bj} 2^{\frac{3}{2}(\frac{1}{p} - \frac{1}{2})j} \| f \|_{L^p L^2}. \tag{8-105}$$

Hereafter, the proofs of the three bounds differ.

Step 1.3a: proof of (8-36). We decompose the inputs and the output by frequency projections to rectangular boxes of the form $\mathcal{C}_{k_1}(\ell)$. Then we need to consider the expression

$$P_k Q_{<j-C} P_{\mathcal{C}} \mathcal{N}(|D|^{-1} Q_j P_{\mathcal{C}^1} u_{k_1}, Q_{<j-C} P_{\mathcal{C}^2} v_{k_2}),$$

where $\mathcal{C}, \mathcal{C}^1, \mathcal{C}^2 \in \{\mathcal{C}_{k_1}(\ell)\}$. Note that the above expression is nonvanishing only when $-\mathcal{C} + \mathcal{C}^1 + \mathcal{C}^2 \ni 0$. Moreover, by the geometry of the cone (Lemma 8.21), for each fixed $\mathcal{C}$ (resp. $\mathcal{C}^2$), this expression is nonvanishing only for $O(1)$-many $\mathcal{C}^1$ and $\mathcal{C}^2$ (resp. $\mathcal{C}$), and the null form gains the factor $2^{k_1+k_2} 2^{\ell}$.

For exponents $p_1, p_2, q_1, q_2 \geq 2$ such that $p_1^{-1} + p_2^{-1} = p^{-1}$ and $q_1^{-1} + q_2^{-1} = 2^{-1}$, proceeding carefully to exploit spatial orthogonality in L^2, we have

$$\begin{aligned}
&\| P_k Q_{<j-C} \mathcal{N}(|D|^{-1} Q_j u_{k_1}, Q_{<j-C} v_{k_2}) \|_{L^p L^2} \\
&\quad = \Big\| \sum_{\mathcal{C}, \mathcal{C}^1, \mathcal{C}^2} P_k Q_{<j-C} P_{\mathcal{C}} \mathcal{N}(|D|^{-1} Q_j P_{\mathcal{C}^1} u_{k_1}, Q_{<j-C} P_{\mathcal{C}^2} v_{k_2}) \Big\|_{L^p L^2} \\
&\quad \lesssim \bigg\| \Big(\sum_{\mathcal{C}} \Big\| \sum_{\mathcal{C}^1, \mathcal{C}^2} P_k Q_{<j-C} P_{\mathcal{C}} \mathcal{N}(|D|^{-1} Q_j P_{\mathcal{C}^1} u_{k_1}, Q_{<j-C} P_{\mathcal{C}^2} v_{k_2})(t, \cdot) \Big\|_{L^2}^2 \Big)^{\frac{1}{2}} \bigg\|_{L^p_t} \\
&\quad \lesssim 2^{\ell} 2^{k_2} \Big\| \sup_{\mathcal{C}_1} \| Q_j P_{\mathcal{C}^1} u_{k_1}(t, \cdot) \|_{L^{q_1}} \Big\|_{L^{p_1}_t} \bigg\| \Big(\sum_{\mathcal{C}^2} \| Q_{<j-C} P_{\mathcal{C}^2} v_{k_2}(t, \cdot) \|_{L^{q_2}}^2 \Big)^{\frac{1}{2}} \bigg\|_{L^{p_2}_t} \\
&\quad \lesssim 2^{\ell} 2^{k_2} \| Q_j u_{k_1} \|_{L^{p_1} L^{q_1}} \Big(\sum_{\mathcal{C}^2} \| Q_{<j-C} P_{\mathcal{C}^2} v_{k_2} \|_{L^{p_2} L^{q_2}}^2 \Big)^{\frac{1}{2}}. \qquad (8\text{-}106)
\end{aligned}$$

We now apply (8-105) and (8-106) with

$$(s, b, p, p_1, q_1, p_2, q_2) = \Big(\frac{5}{4} - \frac{3}{p_0} + \Big(\frac{1}{4} - b_0\Big)\theta_0, -\frac{1}{4} - \Big(\frac{1}{4} - b_0\Big)\theta_0, p_0, 2, 2, \frac{2p_0}{2-p_0}, \infty \Big),$$

where $\theta_0 = 2(\frac{1}{p_0} - \frac{1}{2})$. We then obtain

$$\begin{aligned}
&\| P_k Q_{<j-C} \mathcal{N}(|D|^{-1} Q_j u_{k_1}, Q_{<j-C} v_{k_2}) \|_{\Box \tilde{Z}^1_{p_0}} \\
&\quad \lesssim 2^{-(1-\frac{1}{p_0})k} 2^{k} 2^{(-\frac{1}{4} - (\frac{1}{4} - b_0)\theta_0)(j-k)} 2^{-\frac{3}{2}(1-\frac{1}{p_0})(j-k)} 2^{\frac{3}{4}(j-k)} 2^{\ell} \\
&\qquad\qquad \times \| Q_j u_{k_1} \|_{L^2 L^2} \Big(\sum_{\mathcal{C}^2} \| P_{\mathcal{C}^2} Q_{<j-C} v_{k_2} \|_{L^{p_2} L^\infty}^2 \Big)^{\frac{1}{2}} \\
&\quad \lesssim 2^{(-\frac{3}{4} + \frac{1}{2}(1-\frac{1}{p_0}) + (\frac{1}{4} - b_0)\theta_0)(k_1 - j)} 2^{(-\frac{1}{2}(1-\frac{1}{p_0}) + (\frac{1}{4} - b_0)\theta_0)(k-k_1)} \\
&\qquad\qquad \times \| Q_j u_{k_1} \|_{X^{1,1/2}_\infty} \Big(\sum_{\mathcal{C}^2} \| D P_{\mathcal{C}^2} v_{k_2} \|_{S_{k_2}[\mathcal{C}_{k_1}(\ell)]}^2 \Big)^{\frac{1}{2}}.
\end{aligned}$$

On the last line, we used

$$\| Q_{<j-C} P_{\mathcal{C}^2} v_{k_2} \|_{L^{p_2} L^\infty} \lesssim 2^{(\frac{3}{2} - \theta_0)\ell} 2^{(2-\theta_0)(k_1 - k_2)} 2^{(2 - \frac{1}{2}\theta_0)k_2} \| P_{\mathcal{C}^2} v_{k_2} \|_{S_{k_2}[\mathcal{C}_{k_1(\ell)}]},$$

which follows from interpolation. By (4-4), the factors in front of $(k_1 - j)$ and $(k - k_1)$ are both negative. Summing up in $j < k_1 + C$, we obtain (8-36).

Step 1.3b: proof of (8-37). As in the proof of (8-104) (Step 1.2c), we take the decompositions $u_{k_1} = \sum_\omega P_\ell^\omega u_{k_1}$ and $v_{k_2} = \sum_{\omega'} P_\ell^{\omega'} v_{k_2}$, where $\ell = (j-k_1)/2$. By the geometry of the cone (Lemma 8.21), the null form gain, Hölder, Cauchy–Schwarz (in ω, ω') and Bernstein (for u_{k_1}), we have

$$\begin{aligned}\|P_k Q_{<j} \mathcal{N}(|D|^{-1} Q_j P_\ell^\omega u_{k_1}, Q_{<j-C} P_\ell^{\omega'} v_{k_2})\|_{L^p L^2} \\ \lesssim 2^{(1+3(1-\frac{1}{p}))\ell} 2^{4(1-\frac{1}{p})k_1} 2^{k_2} \Big(\sum_\omega \|P_\ell^\omega Q_j u_{k_1}\|_{L^p L^{p'}}^2\Big)^{\frac12} \Big(\sum_{\omega'} \|P_\ell^{\omega'} Q_{<j-C} v_{k_2}\|_{L^\infty L^2}^2\Big)^{\frac12}. \end{aligned} \tag{8-107}$$

Applying (8-105) and (8-107) with $(s, b, p) = \big(\frac32 - \frac{3}{p_0} + \big(\frac14 - b_0\big)\theta_0, -\frac12 - \big(\frac14 - b_0\big)\theta_0, p_0\big)$, where $\theta_0 = 2\big(\frac{1}{p_0} - \frac12\big)$, we obtain

$$\begin{aligned}&\|P_k Q_{<j} \mathcal{N}(|D|^{-1} Q_j P_\ell^\omega u_{k_1}, Q_{<j} P_\ell^{\omega'} v_{k_2})\|_{\Box Z^1_{p_0}} \\ &\lesssim 2^{-(1-\frac{1}{p_0})k} 2^{(\frac14 - (\frac14 - b_0)\theta_0)(j-k)} 2^{-\frac32(1-\frac{1}{p_0})(j-k)} \|P_k Q_{<j} \mathcal{N}(|D|^{-1} Q_j P_\ell^\omega u_{k_1}, Q_{<j} P_\ell^{\omega'} v_{k_2})\|_{L^{p_0} L^2} \\ &\lesssim 2^{(-\frac14 + (\frac14 - b_0)\theta_0 + \frac12(1-\frac{1}{p_0}))(k-k_1)} \|Q_j u_{k_1}\|_{X_1^{9/4-3/p_0+(1/4-b_0)\theta_0, 3/4-(1/4-b_0)\theta_0, p_0}} \|D v_{k_2}\|_S.\end{aligned}$$

By our choices of b_0 and p_0, the overall factor in front of $(k - k_1)$ is negative. Summing up in $j < k_1 + C$, we obtain the desired conclusion.

Step 1.3c: proof of (8-38). We again take the decompositions $u_{k_1} = \sum_\omega P_\ell^\omega u_{k_1}$ and $v_{k_2} = \sum_{\omega'} P_\ell^{\omega'} v_{k_2}$, where $\ell = (j-k_1)/2$. We use (8-105) with $(s, b, p) = \big(-\frac54 - b_0, -\frac34 + b_0, 1\big)$. By the geometry of the cone (Lemma 8.21), the null form gain, Hölder and Cauchy–Schwarz (in ω, ω'), we have

$$\begin{aligned}&2^{b_0(j-k)} \|P_k Q_{<j} \mathcal{N}(|D|^{-1} Q_j P_\ell^\omega u_{k_1}, Q_{<j-C} P_\ell^{\omega'} v_{k_2})\|_{L^1 L^2} \\ &\quad \lesssim 2^{b_0(j-k)} 2^\ell 2^{k_2} \Big(\sum_\omega \|Q_j P_\ell^\omega u_{k_1}\|_{L^{p_0} L^{p_0'}}^2\Big)^{\frac12} \Big(\sum_{\omega'} \|Q_{<j-C} P_\ell^{\omega'} v_{k_2}\|_{L^{p_0'} L^{q_0}}^2\Big)^{\frac12} \\ &\quad \lesssim 2^{(b_0 + (\frac14 - b_0)\theta_0)(k_1 - j)} 2^{-b_0(k-k_1)} 2^{3(1-\frac{1}{p_0})(k-k_1)} \\ &\qquad\qquad \times \|u_{k_1}\|_{X_\infty^{3(1-1/p_0)-1/2+(1/4-b_0)\theta_0, 1/2-(1/4-b_0)\theta_0, p_0}} \|D v_{k_2}\|_S,\end{aligned}$$

where $q_0^{-1} = 2^{-1} - (p_0')^{-1}$ and $\theta_0 = 2\big(\frac{1}{p_0} - \frac12\big)$. By our choices of p_0 and b_0, the overall factors in front of $(k_1 - j)$ and $(k - k_1)$ are both negative. Summing up in $j < k_1$, the proof is complete.

Step 2: proof of (8-39). As in Step 1, we divide into three cases.

Step 2.1: $k_1 \geq k - 10$. In view of the embedding $N \cap L^2 \dot H^{-1/2} \subseteq X^{-1/2+b_1, -b_1}$ for any $0 < b_1 < \frac12$, the desired bound follows from (8-15) and (8-21).

Step 2.2: $k_1 < k - 10$, contribution of $1 - \mathcal{H}_{k_1}^*$. Consider the expression

$$P_k (1 - \mathcal{H}_{k_1}^*) \mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2}).$$

Interpolating the N-norm bound (8-26) (recall that $N \subseteq X_\infty^{0, -1/2}$) with an $L^2 \dot H^{-1/2}$-norm bound (which is a minor modification of (8-15)), the desired estimate for this expression follows for $0 < b_1 < \frac12$.

Step 2.3: $k_1 < k-10$, contribution of $\mathcal{H}^*_{k_1}$. Finally, we estimate

$$P_k \mathcal{H}^*_{k_1} \mathcal{N}(|D|^{-1} u_{k_1}, v_{k_2}) = \sum_{j<k_1+C} P_k Q_{<j-C} \mathcal{N}(|D|^{-1} Q_j u_{k_1}, Q_{<j-C} v_{k_2}).$$

By (8-107), we have

$$\begin{aligned}
&2^{(\frac{1}{p_0}-1)k} \| P_k Q_{<j-C} \mathcal{N}(|D|^{-1} Q_j u_{k_1}, Q_{<j-C} v_{k_2}) \|_{L^{p_0} L^2} \\
&\lesssim 2^{(\frac{1}{2}+\frac{3}{2}(1-\frac{1}{p_0}))(j-k_1)} 2^{(\frac{1}{p_0}-1)k} 2^{4(1-\frac{1}{p_0})k_1} 2^{-3(1-\frac{1}{p_0})k_1} 2^{(-\frac{1}{2}+(\frac{1}{4}-b_0)\theta_0)(j-k_1)} \|u_{k_1}\|_{Z^1_{p_0}} \|Dv_{k_2}\|_{L^\infty L^2} \\
&\lesssim 2^{(-\frac{3}{2}(1-\frac{1}{p_0})-(\frac{1}{4}-b_0)\theta_0)(k_1-j)} 2^{-(1-\frac{1}{p_0})(k-k_1)} \|u_{k_1}\|_{Z^1_{p_0}} \|Dv_{k_2}\|_S.
\end{aligned}$$

Summing up in $j < k_1 + C$ and using the embedding $2^{(1-1/p_0)k} P_k Q_{<k} L^{p_0} L^2 \subseteq X^{-1/2+b_1,-b_1}$, which holds by Bernstein since $b_1 < \frac{1}{p_0} - \frac{1}{2}$, the proof of (8-39) is complete. □

Proof of Proposition 8.16. As in Proposition 8.15, we divide the proof into two cases: $k_1 \geq k-10$ and $k_1 < k-10$.

Step 1: $k_1 \geq k-10$. In this case, by (8-18), (8-23) and the embeddings $L^1 L^2 \subseteq \Box Z^1_{p_0} \cap \Box Z^1$ and $L^1 L^2 \cap L^2 \dot{H}^{-1/2} \subseteq X^{-1/2+b_1,-b_1}$, the three bounds follow simultaneously.

Step 2: $k_1 < k-10$. We begin with (8-40) and (8-42). By Hölder and Bernstein, we have

$$2^{(\frac{1}{p_0}-1)k} \| P_k \mathcal{O}(u_{k_1}, v'_{k_2}) \|_{L^{p_0} L^2} \lesssim 2^{-(1-\frac{1}{p_0})(k-k_1)} \|u_{k_1}\|_{L^{p_0} \dot{W}^{2-\frac{3}{p_0}, p'_0}} \|v'_{k_2}\|_{L^\infty L^2}$$

By (8-105) with $(s, b, p) = \big(\frac{3}{2} - \frac{3}{p_0}, -\frac{1}{2}, p_0\big)$, (8-40) follows. Moreover, by the $L^2 \dot{H}^{-1/2}$-norm estimate (8-15) and the embedding $P_k Q_{<k} L^{p_0} L^2 \subseteq X^{-1/2+b_1,-b_1}$, (8-42) follows as well.

It remains to prove (8-41). Applying (8-100) (from Step 1.2a of the proof of Proposition 8.15) with $Dv_{k_2} = v'_{k_2}$ and the embedding $2^{-3k_1/2} P_{k_1} Y \subseteq L^2 L^\infty$, we have

$$\| P_k Q_{\geq k_1} \mathcal{O}(u_{k_1}, v'_{k_2}) \|_{\Box Z^1} \lesssim 2^{-b_0(k-k_1)} \|Du_{k_1}\|_Y \|v'_{k_2}\|_S.$$

On the other hand, by (8-102) and Hölder, we have

$$\begin{aligned}
\| P_k Q_{<k_1} \mathcal{O}(u_{k_1}, v'_{k_2}) \|_{\Box Z^1} &\lesssim 2^{-b_0(k-k_1)} \| P_k \mathcal{O}(u_{k_1}, v'_{k_2}) \|_{L^1 L^2} \\
&\lesssim 2^{-b_0(k-k_1)} 2^{3(1-\frac{1}{p_0})(k-k_1)} \|Du_{k_1}\|_Y (2^{(\frac{3}{p_0}-3)k_2} \|v'_{k_2}\|_{L^{p'_0} L^{q_0}}) \\
&\lesssim 2^{-b_0(k-k_1)} 2^{3(1-\frac{1}{p_0})(k-k_1)} \|Du_{k_1}\|_Y \|v'_{k_2}\|_S,
\end{aligned}$$

where $q_0^{-1} = 2^{-1} - (p'_0)^{-1}$. By our choice of p_0, the overall factor in front of $(k-k_1)$ is negative; hence, (8-41) follows. □

8D5. *Trilinear null form estimates.*

Proofs of Propositions 8.17 and 8.18. Estimate (8-43) would follow from Lemma 8.10 and the core estimates (8-44), (8-45) and (8-46), combined with Lemma 8.21 and (4-1).

Estimates (8-44), (8-45) and (8-46) can be established by repeating the proofs of (136), (137) and (138) in [Krieger et al. 2015] with the following modifications:

- Thanks to the frequency localization of the inputs and the output to rectangular boxes of the type $\mathcal{C}_k(\ell)$, the bilinear operators $\mathcal{O}$ and $\mathcal{O}'$ can be safely disposed of.

- Moreover, for any disposable multilinear operator $\mathcal{M}$ and rectangular boxes $\mathcal{C}, \mathcal{C}'$ of the type $\mathcal{C}_k(\ell)$ situated in the annuli $\{|\xi| \simeq 2^{k_1}\}$ and $\{|\xi| \simeq 2^{k_2}\}$, respectively, note that (by Lemma 8.7)

$$\mathcal{M}(\partial^\alpha Q^{\pm}_{<j-C} P_{\mathcal{C}} u_{k_1}, \partial_\alpha Q^{\pm'}_{<j-C} P_{\mathcal{C}'} v_{k_2}, \dots) \\ = C 2^{k_1+k_2} \max\{|\angle(\pm\mathcal{C}, \pm'\mathcal{C}')|^2, 2^{j-\min\{k_1,k_2\}}\} \widetilde{\mathcal{M}}(P_{\mathcal{C}} u_{k_1}, P_{\mathcal{C}'} v_{k_2}, \dots)$$

for some disposable $\widetilde{\mathcal{M}}$, which suffices for the proofs in [Krieger et al. 2015].

We also note that although (136)–(138) in [Krieger et al. 2015] are stated with the factor $2^{\delta(k-\min\{k_i\})}$ on the right-hand side, an inspection of the proofs reveals that the actual gain is $2^{\delta(k-k_1)}$, as claimed in (8-44)–(8-46). We omit the straightforward details. □

9. The paradifferential wave equation

Sections 9, 10 and 11 are devoted to the proofs of Theorem 4.24 and Proposition 4.25. In this section, we first reduce the task of proving these results to that of constructing an appropriate parametrix (Section 9A). Parametrix construction, in turn, is reduced to constructing a renormalization operator that roughly conjugates $\Box + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}$ to $\Box$. Sections 10 and 11 are devoted proofs of the desired properties of the renormalization operator.

9A. ***Reduction to parametrix construction.*** We start with a quick reduction of the problem (4-90). After peeling off perturbative terms using commutator estimates (which will be sketched in more detail below), we are led to consideration of the frequency-localized problem

$$\begin{cases} \Box u_k + 2[P_{<k-\kappa} \boldsymbol{P}_\alpha A, \partial^\alpha u_k] = f_k, \\ (u_k, \partial_t u_k)(0) = (u_{0,k}, u_{1,k}), \end{cases} \tag{9-1}$$

for each $k \in \mathbb{Z}$. By scaling, we may normalize $k = 0$.

Our goal is to construct a parametrix to (9-1). We summarize the main properties of the parametrix in this case, as well as the precise hypotheses on A_α that we need, in the following theorem.

Theorem 9.1 (parametrix construction). *Let A_α be a $\mathfrak{g}$-valued 1-form on $I \times \mathbb{R}^4$ such that*

$$\|A\|_{S^1[I]} + \|\Box A\|_{\ell^1 X^{-1/2+b_1,-b_1}[I]} \le M \tag{9-2}$$

for some $M > 0$ and $b_1 > \frac{1}{4}$. Let $\varepsilon > 0$. Assume that $\kappa > \kappa_1(\varepsilon, M)$ and

$$\|A\|_{DS^1[I]} + \|\Box A\|_{\ell^1 L^2 \dot{H}^{-1/2}} < \delta_p(\varepsilon, M, \kappa_1) \tag{9-3}$$

for some functions $\kappa_1(\varepsilon, M) \gg 1$, $0 < \delta_p(\varepsilon, M, \kappa_1) \ll 1$ *independent of* A_α. *Moreover, assume that there exists* $\tilde{A}_\alpha$ *such that*

$$\|\tilde{A}\|_{\underline{S}^1[I]} + \|(D\tilde{A}_0, D\boldsymbol{P}^\perp \tilde{A})\|_{Y[I]} \leq M, \tag{9-4}$$

$$\|\tilde{A}\|_{DS^1[I]} + \|(\tilde{A}_0, \boldsymbol{P}^\perp \tilde{A})\|_{L^2 \dot{H}^{3/2}[I]} < \delta_p(\varepsilon, M, \kappa_1), \tag{9-5}$$

and

$$\|\Delta A_0 - \boldsymbol{O}(\tilde{A}^\ell, \partial_0 \tilde{A}_\ell)\|_{\ell^1(\Delta L^1 L^\infty \cap L^2 \dot{H}^{-1/2})[I]} < \delta_p^2(\varepsilon, M, \kappa_1), \tag{9-6}$$

$$\|\Box \boldsymbol{P} A - \boldsymbol{P}\boldsymbol{O}(\tilde{A}^\ell, \partial_x \tilde{A}_\ell)) - \boldsymbol{P}\boldsymbol{O}'(\tilde{A}^\alpha, \partial_\alpha \tilde{A})\|_{\ell^1(L^1 L^2 \cap L^2 \dot{H}^{-1/2})[I]} < \delta_p^2(\varepsilon, M, \kappa_1), \tag{9-7}$$

where $\boldsymbol{O}(\cdot,\cdot)$ *and* $\boldsymbol{O}'(\cdot,\cdot)$ *are disposable bilinear operators on* $\mathbb{R}^4$. *Then the following statements hold*:

(1) *Given any* $(u_0, u_1) \in \dot{H}^1 \times L^2$ *and* $f \in N \cap L^2 \dot{H}^{-1/2}$ *such that* u_0, u_1, f *are all frequency-localized in* $\{C^{-1} \leq |\xi| \leq C\}$, *there exists a* $\mathfrak{g}$*-valued function* $u(t)$ *on* I *which obeys*

$$\|u\|_{S^1[I]} \lesssim_M \|(u_0, u_1)\|_{\dot{H}^1 \times L^2} + \|f\|_{N \cap L^2 \dot{H}^{-1/2}[I]}, \tag{9-8}$$

$$\|\Box u + 2[P_{<-\kappa} \boldsymbol{P}_\alpha A, \partial^\alpha u] - f\|_{N \cap L^2 \dot{H}^{-1/2}[I]} \leq \varepsilon(\|(u_0, u_1)\|_{\dot{H}^1 \times L^2} + \|f\|_{N \cap L^2 \dot{H}^{-1/2}[I]}), \tag{9-9}$$

$$\|u[0] - (u_0, u_1)\|_{\dot{H}^1 \times L^2} \leq \varepsilon(\|(u_0, u_1)\|_{\dot{H}^1 \times L^2} + \|f\|_{N \cap L^2 \dot{H}^{-1/2}[I]}). \tag{9-10}$$

Moreover, u *is frequency-localized in* $\{(2C)^{-1} \leq |\xi| \leq 2C\}$.

(2) *Assume furthermore that*

$$\|A_x\|_{\ell^\infty S^1[I]} + \|A_0\|_{\ell^\infty L^2 \dot{H}^{3/2}[I]} < \delta_o(M) \tag{9-11}$$

for some $\delta_o(M) \ll 1$ *independent of* A_α. *Then the approximate solution* u *constructed above obeys* (9-8) *with a universal constant*, *i.e.*,

$$\|u\|_{S^1[I]} \lesssim \|(u_0, u_1)\|_{\dot{H}^1 \times L^2} + \|f\|_{N \cap L^2 \dot{H}^{-1/2}[I]}. \tag{9-12}$$

In the remainder of this subsection, we sketch the proofs of Theorem 4.24 and Proposition 4.25 assuming Theorem 9.1. Then in the rest of this section, as well as in Sections 10 and 11, our goal will be to establish Theorem 9.1.

Lemma 9.2. (a) *Let* $A_{t,x}$ *and* $\tilde{A}_{t,x}$ *be* $\mathfrak{g}$*-valued* 1*-forms on* $I \times \mathbb{R}^4$, *which satisfy* (9-2), (9-3), (9-4), (9-5), (9-6) *and* (9-7). *Then for* $\varepsilon > 0$ *sufficiently small* (*depending on* M) *and* κ *sufficiently large* (*depending on* ε, M), *given any* $(u_0, u_1) \in \dot{H}^1 \times L^2$ *and* $f \in N \cap L^2 \dot{H}^{-1/2}[I]$, *there exists a unique solution* $u \in S^1[I]$ *to the IVP*

$$\begin{cases} (\Box + \mathrm{Diff}^\kappa_{\boldsymbol{P}A}) u = f, \\ u[0] = (u_0, u_1), \end{cases} \tag{9-13}$$

which obeys

$$\|u\|_{S^1[I]} \lesssim_M \|(u_0, u_1)\|_{\dot{H}^1 \times L^2} + \|f\|_{N \cap L^2 \dot{H}^{-1/2}[I]}. \tag{9-14}$$

(b) *If*, *in addition*, $\|A\|_{\ell^\infty S^1[I]}$ *obeys* (9-11), *then the solution* u *constructed above obeys* (9-14) *with a universal constant*, *i.e.*,

$$\|u\|_{S^1[I]} \lesssim \|(u_0, u_1)\|_{\dot{H}^1 \times L^2} + \|f\|_{N \cap L^2 \dot{H}^{-1/2}[I]}. \tag{9-15}$$

Proof. Let u_k be the function given by (the rescaled) Theorem 9.1 which is determined by the initial data $(P_k u_0, P_k u_1, P_k f)$. We set

$$u_{\mathrm{app}} = \sum_{k'} u_{k'}.$$

We claim u is a good approximate solution to (9-13) in the sense that in any subinterval $J \subset I$ we have

$$\|u_{\mathrm{app}}\|_{S^1[J]} \lesssim_M \|(u_0, u_1)\|_{\dot H^1 \times L^2} + \|f\|_{N \cap L^2 \dot H^{-1/2}[J]}, \tag{9-16}$$

$$\|u_{\mathrm{app}}[0] - (u_0, u_1)\|_{\dot H^1 \times L^2} \lesssim \varepsilon(\|(u_0, u_1)\|_{\dot H^1 \times L^2} + \|f\|_{N \cap L^2 \dot H^{-1/2}[J]}), \tag{9-17}$$

and

$$\begin{aligned}\|(\Box + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}) u_{\mathrm{app}} - f\|_{N \cap L^2 \dot H^{-1/2}[J]} &\lesssim_M (\varepsilon + 2^{-\delta_2 \kappa} + 2^{C\kappa}(\|\boldsymbol{P}A\|_{\ell^\infty DS^1[I]} + \|A_0\|_{\ell^\infty L^2 \dot H^{3/2}[J]})) \\ &\qquad \times (\|(u_0, u_1)\|_{\dot H^1 \times L^2} + \|f\|_{N \cap L^2 \dot H^{-1/2}[J]}).\end{aligned} \tag{9-18}$$

Assume that we have these bounds. Then the solution u to (9-13) is obtained as follows:

(i) We choose κ large enough so that $2^{-\delta_2} \ll_M 1$.

(ii) We divide the interval I into subintervals J_j so that

$$2^{C\kappa} \|\boldsymbol{P}A\|_{DS^1[I]} + \|A_0\|_{L^2 \dot H^{3/2}[I_j]} \ll_M 1.$$

(iii) Within the interval J_1 we now have small errors for the approximate solution u_{app}; hence we can obtain an exact solution by reiterating.

(iv) We successively repeat the previous step on each of the subintervals I_j.

It remains to prove the bounds (9-16), (9-17) and (9-18). The first two follow directly from (9-8) and (9-9) for u_k after summation in k. We now consider (9-18), where we write

$$(\Box + \mathrm{Diff}^{\kappa}_{\boldsymbol{P}A}) u - f = \sum_k (\Box u_k + 2[P_{<k-\kappa} \boldsymbol{P}A_\alpha, \partial^\alpha u_k] - P_k f) + \sum_k g_k,$$

where

$$g_k = 2[P_{<k-\kappa} \boldsymbol{P}A_\alpha, \partial^\alpha u_k] - \sum_{k'} [P_{-k'-\kappa} \boldsymbol{P}A_\alpha, \partial^\alpha P_{k'} u_k]$$

The first sum is estimated directly via (9-9), so it remains to estimate g_k. We write

$$g_k = g_k^1 + g_k^2,$$

where

$$g_k^1 = \sum_{k'=k+O(1)} P_{k'}[P_{-k'-\kappa} \boldsymbol{P}A_\alpha, \partial^\alpha P_{k'} u_k] - [P_{-k'-\kappa} \boldsymbol{P}A_\alpha, \partial^\alpha P_{k'} u_k],$$

$$g_k^2 = \sum_{k'=k+O(1)} [P_{[-k'-\kappa, k-\kappa)} \boldsymbol{P}A_\alpha, \partial^\alpha P_{k'} u_k].$$

Here g_k^1 has a commutator structure, so we can estimate it as in Proposition 4.30, yielding a $2^{-\delta_2 \kappa}$ factor. For the expression g_k^2, on the other hand, we can apply Proposition 4.20 to split it into a small part and a large part which uses only divisible norms. Thus (9-18) follows, and the proof of the lemma is concluded.

(b) The same iterative construction applies, but no we no longer need to subdivide the interval as (9-11) ensure that the divisible norms in (9-18) are actually small. □

Proof of Theorem 4.24 assuming Theorem 9.1. We prove the theorem by repeatedly applying the preceding lemma in successive intervals. To achieve this, we begin by choosing ε and κ depending only on M so that Lemma 9.2 holds. It remains to ensure that we can divide the interval I into subintervals J_j where the conditions (9-2), (9-3), (9-4), (9-5), (9-6) and (9-7) hold.

We choose $\tilde{A} = A$. We carefully observe that we cannot use Theorem 5.1 here, as Theorem 4.24 is used in the proof of Theorem 5.1. However, we can use the weaker result in Proposition 5.4, which immediately gives (9-2) and (9-4) from Theorem 5.1.

The remaining bounds are for divisible norms, so it suffices to establish them with a large constant depending on M; then we gain smallness by subdividing. Indeed, for (9-3) and (9-5) this still follows from Proposition 5.4.

For (9-6) we choose $\boldsymbol{O}(A, \partial_0 A) = [A, \partial_0 A]$. Then we can use (3-23) and (4-37). Finally for (9-7) we choose in addition $\boldsymbol{O}(A_\alpha, \partial^\alpha A) = -2[A_\alpha, \partial^\alpha A]$. Then by Theorem 9.1 we have

$$\Box A - \boldsymbol{O}(A, \partial_x A) - \boldsymbol{O}(A_\alpha, \partial^\alpha A) = R(A) + \mathrm{Rem}^3(A) A$$

and it suffices to use (3-21) and (4-74). □

Proof of Proposition 4.25 assuming Theorem 9.1. We write

$$A_{t,x} = A_{t,x}^{\mathrm{pert}} + A_{t,x}^{\mathrm{nonpert}},$$

where

$$A_{t,x}^{\mathrm{pert}} = \sum_{k \in K} P_k A_{t,x},$$

with $|K| = O_{\delta_o(M)^{-1} M}(1)$ and

$$\| A^{\mathrm{nonpert}} \|_{\ell^\infty S^1[I]} < \delta_o(M).$$

By Proposition 4.23, it follows that the contribution of any finite number of dyadic pieces of $A_{t,x}$ in $\mathrm{Diff}^\kappa_{\boldsymbol{P}A}$ is perturbative. More precisely, for A^{pert}, we have

$$\| \mathrm{Diff}^\kappa_{\boldsymbol{P}A^{\mathrm{pert}}} B \|_{N \cap L^2 \dot{H}^{-1/2}[I]} \lesssim_{|K|, M} \| B \|_{S^1[I]}. \tag{9-19}$$

Thus B solves also

$$(\Box + \mathrm{Diff}^\kappa_{\boldsymbol{P}A^{\mathrm{nonpert}}}) B = \widetilde{G},$$

where

$$\| \widetilde{G} \|_{N \cap L^2 \dot{H}^{-1/2}[I]} \lesssim_M \| G \|_{N \cap L^2 \dot{H}^{-1/2}[I]} + \| B \|_{S^1[I]}.$$

We now claim that Theorem 9.1 and thus Lemma 9.2 apply for A^{nonpert}. If that were true, then the conclusion of the proposition is achieved by subdividing the interval I into finitely many subintervals J_j, depending only on M, so that

(i) Lemma 9.2 applies in J_j,

(ii) the size of the inhomogeneous term $\| \widetilde{G} \|_{N \cap L^2 \dot{H}^{-1/2}[I]}$ is small in J_j.

Indeed, to verify the hypothesis of Theorem 9.1 with A replaced by A^{nonpert} it suffices to leave $\tilde{A} = A$, unchanged, but instead replace the operators $\boldsymbol{O}$ and $\boldsymbol{O}'$ by $\big(1 - \sum_{k \in K} P_k\big)\boldsymbol{O}$ and $\big(1 - \sum_{k \in K} P_k\big)\boldsymbol{O}'$, respectively, which are still disposable. □

9B. *Extension and space-time Fourier projections.* As in [Krieger and Tataru 2017], our parametrix will be constructed by conjugating the usual Fourier representation formula for the ±-half-wave equations by a *renormalization operator* $\mathrm{Op}(\mathrm{Ad}(O_{\pm})_{<0})$; see (9-50). The renormalization operator is designed so that it cancels the most dangerous part of the paradifferential term $2[\boldsymbol{P} A_{\alpha,<-\kappa}, \partial^{\alpha} P_0 u]$ (Theorem 9.9), and furthermore enjoys nice mapping properties in functions spaces we use (Theorem 9.6).

9B1. *Extension to a global-in-time wave.* As in [Krieger and Tataru 2017], our parametrix construction for (9-1) involves fine space-time Fourier localization of $\boldsymbol{P}A$, which necessitates extension of $\boldsymbol{P}A$ outside I. Here we specify the extension procedure, and collect some of its properties that will be used later.

We extend $\boldsymbol{P}A$ by homogeneous waves outside I. By (9-2), this extension (still denoted by $\boldsymbol{P}A$) obeys the global-in-time bound

$$\|\boldsymbol{P}A\|_{S^1} + \|\Box \boldsymbol{P}A\|_{\ell^1 X^{-1/2+b_1,-b_1}} \lesssim M. \tag{9-20}$$

By Proposition 4.10, for any $p \geq 2$ note that

$$\|\chi_I^k P_k \boldsymbol{P}A\|_{L^p L^\infty} \lesssim \|P_k \boldsymbol{P}A\|_{L^p L^\infty[I]}. \tag{9-21}$$

Moreover, by (9-3), we have

$$\sum_k \|P_k \Box \boldsymbol{P}A\|_{L^2 \dot{H}^{-1/2}} = \|\Box \boldsymbol{P}A\|_{\ell^1 L^2 \dot{H}^{-1/2}[I]} < \delta_p. \tag{9-22}$$

Next, we specify the extension of A_0, and also of the relations (9-6) and (9-7) outside I. We first extend $\tilde{A}$ by homogeneous wave outside I and $\tilde{A}_0$ by zero outside I. These extensions (still denoted by $\tilde{A}$ and $\tilde{A}_0$, respectively) satisfy the global-in-time bound

$$\|\tilde{A}\|_{\underline{S}^1} + \|D\tilde{A}_0\|_Y \lesssim M. \tag{9-23}$$

In addition, we introduce the extension $\widetilde{G}$ of $\boldsymbol{P}^{\perp}\tilde{A}$ by zero outside I. It obeys

$$\|D\widetilde{G}\|_Y \lesssim M. \tag{9-24}$$

We emphasize that, in general, $\boldsymbol{P}^{\perp}\tilde{A}$ does *not* coincide with $\widetilde{G}$ outside I.

Define $\widetilde{R}_0$ and $\boldsymbol{P}\widetilde{R}$ as

$$\begin{aligned} \widetilde{R}_0(t) &= \Delta A_0(t) - \boldsymbol{O}(\tilde{A}^{\ell}(t), \partial_t \tilde{A}_{\ell}(t)) && \text{for } t \in I, \\ \boldsymbol{P}\widetilde{R}(t) &= \Box \boldsymbol{P}A(t) - \boldsymbol{P}\boldsymbol{O}(\tilde{A}^{\ell}(t), \partial_x \tilde{A}_{\ell}(t)) + \boldsymbol{P}\boldsymbol{O}'(\tilde{A}_{\alpha}, \partial^{\alpha}\tilde{A}) && \text{for } t \in I, \end{aligned}$$

and 0 for $t \notin I$. By the hypotheses (9-6) and (9-7), we have

$$\|\widetilde{R}_0\|_{\ell^1(\Delta L^1 L^\infty \cap L^2 \dot{H}^{-1/2})} < \delta_p^2, \tag{9-25}$$

$$\|\boldsymbol{P}\widetilde{R}\|_{\ell^1(L^1 L^2 \cap L^2 \dot{H}^{-1/2})} < \delta_p^2. \tag{9-26}$$

We extend A_0 outside I by solving the equation

$$\Delta A_0 = \boldsymbol{O}(\chi_I \tilde{A}^{\ell}, \partial_t \tilde{A}_{\ell}) + \chi_I \widetilde{R}_0. \tag{9-27}$$

By (8-15), (8-17), (9-5), (9-23) and (9-25), it follows that

$$\|DA_0\|_{\ell^1 Y} \lesssim M^2, \tag{9-28}$$

$$\|\Delta A_0\|_{\ell^1 L^2 \dot{H}^{-1/2}} \lesssim \delta_p^2. \tag{9-29}$$

Moreover, observe that the extension $\boldsymbol{P}A$ obeys the equation

$$\Box \boldsymbol{P}A = \boldsymbol{P}\boldsymbol{O}(\chi_I \tilde{A}^{\ell}, \partial_x \tilde{A}_{\ell}) + \boldsymbol{P}\boldsymbol{O}'(\boldsymbol{P}_{\ell}\tilde{A}, \chi_I \partial^{\ell}\tilde{A}) - \boldsymbol{P}\boldsymbol{O}'(\tilde{A}_0, \chi_I \partial_t \tilde{A}) + \boldsymbol{P}\boldsymbol{O}'(\widetilde{G}_{\ell}, \chi_I \partial^{\ell}\tilde{A}) + \chi_I \boldsymbol{P}\widetilde{R}. \tag{9-30}$$

9B2. *Space-time Fourier projections.* Here we introduce the space-time Fourier projections needed for definition of the renormalization operator. We denote by $(\tau, \xi) \in \mathbb{R} \times \mathbb{R}^4$ the Fourier variables for the input, and by $(\sigma, \eta) \in \mathbb{R} \times \mathbb{R}^4$ the Fourier variables for the symbol, which will be constructed from $\boldsymbol{P}A$. We remind the reader that our sign convention is such that the characteristic cone for a $\pm$-wave is $\{\tau \pm |\xi| = 0\}$.

Consider the following (overlapping) decomposition of $\mathbb{R}^{1+4}$, which is symmetric and homogeneous with respect to the origin:

$$\begin{aligned}
D^{\omega,\pm}_{\mathrm{cone}} &= \big\{\mathrm{sgn}(\sigma)(\sigma \pm \eta\cdot\omega) > \tfrac{1}{16}|\eta|^{-1}(|\eta_{\perp}|^2 + |\sigma \pm \eta\cdot\omega|^2)\big\} \\
&\qquad\cap \big\{\mathrm{sgn}(\sigma)(\sigma \pm \eta\cdot\omega) < \tfrac{4}{5}|\sigma|^{-1}(|\eta_{\perp}|^2 + |\sigma \pm \eta\cdot\omega|^2)\big\}, \\
D^{\omega,\pm}_{\mathrm{null}} &= \big\{|\sigma \pm \eta\cdot\omega| < \tfrac{1}{8}|\eta|^{-1}(|\eta_{\perp}|^2 + |\sigma \pm \eta\cdot\omega|^2)\big\}, \\
D^{\omega,\pm}_{\mathrm{out}} &= \big\{\mathrm{sgn}(\sigma)(\sigma \pm \eta\cdot\omega) < -\tfrac{1}{16}|\eta|^{-1}(|\eta_{\perp}|^2 + |\sigma \pm \eta\cdot\omega|^2)\big\} \\
&\qquad\cup \big\{\mathrm{sgn}(\sigma)(\sigma \pm \eta\cdot\omega) > \tfrac{2}{3}|\sigma|^{-1}(|\eta_{\perp}|^2 + |\sigma \pm \eta\cdot\omega|^2)\big\},
\end{aligned}$$

where $\eta_{\perp} = \eta - (\eta\cdot\omega)\omega$. See Figure 1 for a plot of these domains.

We construct a smooth partition of unity adapted to the decomposition $D^{\omega,\pm}_{\mathrm{cone}} \cup D^{\omega,\pm}_{\mathrm{null}} \cup D^{\omega,\pm}_{\mathrm{out}} = \mathbb{R}^{1+4}$ as follows. We begin with the preliminary definitions

$$\begin{aligned}
\widetilde{\Pi}^{\omega,\pm}_{\mathrm{in}}(\sigma, \eta) &= m_{>1}\bigg(\frac{4}{5}\frac{\sigma(\sigma \pm \eta\cdot\omega)}{(|\eta|^2 - (\eta\cdot\omega)^2) + |\sigma \pm \eta\cdot\omega|^2}\bigg), \\
\widetilde{\Pi}^{\omega,\pm}_{\mathrm{med}}(\sigma, \eta) &= m_{>1}\bigg(8\frac{\mathrm{sgn}(\sigma)|\eta|(\sigma \pm \eta\cdot\omega)}{(|\eta|^2 - (\eta\cdot\omega)^2) + |\sigma \pm \eta\cdot\omega|^2}\bigg), \\
\widetilde{\Pi}^{\omega,\pm}_{\mathrm{out}}(\sigma, \eta) &= m_{>1}\bigg(-8\frac{\mathrm{sgn}(\sigma)|\eta|(\sigma \pm \eta\cdot\omega)}{(|\eta|^2 - (\eta\cdot\omega)^2) + |\sigma \pm \eta\cdot\omega|^2}\bigg),
\end{aligned}$$

where $m_{>1}(z) : \mathbb{R} \to [0, 1]$ is a smooth cutoff to the region $\{z > 1\}$ (i.e., equals 1 there), which vanishes outside $\{z > \frac{5}{6}\}$. Then we define the symbols

$$\Pi^{\omega,\pm}_{\mathrm{cone}}(\sigma, \eta) = \widetilde{\Pi}^{\omega,\pm}_{\mathrm{med}}(\sigma, \eta) - \widetilde{\Pi}^{\omega,\pm}_{\mathrm{in}}(\sigma, \eta), \tag{9-31}$$

$$\Pi^{\omega,\pm}_{\mathrm{null}}(\sigma, \eta) = 1 - \widetilde{\Pi}^{\omega,\pm}_{\mathrm{med}}(\sigma, \eta) - \widetilde{\Pi}^{\omega,\pm}_{\mathrm{out}}(\sigma, \eta), \tag{9-32}$$

$$\Pi^{\omega,\pm}_{\mathrm{out}}(\sigma, \eta) = \widetilde{\Pi}^{\omega,\pm}_{\mathrm{out}}(\sigma, \eta) + \widetilde{\Pi}^{\omega,\pm}_{\mathrm{in}}(\sigma, \eta). \tag{9-33}$$

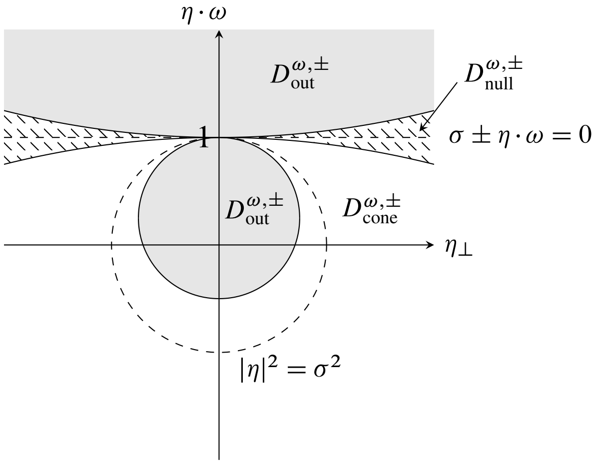

Figure 1. Sketch of $D^{\omega,\pm}_{\mathrm{cone}}$, $D^{\omega,\pm}_{\mathrm{med}}$ and $D^{\omega,\pm}_{\mathrm{out}}$ in the hyperplane $\{\sigma = 1\}$ with $\pm = -$. Note that the actual domains are defined to be slightly overlapping.

Observe that $1 = \Pi^{\omega,\pm}_{\mathrm{cone}} + \Pi^{\omega,\pm}_{\mathrm{null}} + \Pi^{\omega,\pm}_{\mathrm{out}}$, and $\operatorname{supp} \Pi^{\omega,\pm}_{*} \subseteq D^{\omega,\pm}_{*}$ for $* \in \{\mathrm{cone}, \mathrm{null}, \mathrm{out}\}$. Moreover, by symmetry, $\Pi^{\omega,\pm}_{*}$ preserves the real-valued property.

We also make use of a dyadic angular decomposition with respect to ω. Given $\theta > 0$, we define the symbol

$$\Pi^{\omega,\pm}_{>\theta}(\sigma, \eta) = m_{>1}\left(\frac{|\angle(\omega, -\operatorname{sgn}(\sigma)|\eta)}{\theta}\right).$$

Furthermore, we define

$$\Pi^{\omega,\pm}_{\leq\theta}(\sigma, \eta) = 1 - \Pi^{\omega,\pm}_{>\theta}(\sigma, \eta),$$
$$\Pi^{\omega,\pm}_{\theta}(\sigma, \eta) = (\Pi^{\omega,\pm}_{>\theta} - \Pi^{\omega,\pm}_{>\theta/2})(\sigma, \eta).$$

Since these symbols are real-valued and odd, the corresponding multipliers (which we simply denote by $\Pi^{\omega,\pm}_{>\theta}$, $\Pi^{\omega,\pm}_{\geq\theta}$ and $\Pi^{\omega,\pm}_{\theta}$, respectively) preserve the real-valued property.

The regularity of the symbols $\Pi^{\omega,\pm}_{\mathrm{cone}}$, $\Pi^{\omega,\pm}_{\mathrm{null}}$ and $\Pi^{\omega,\pm}_{\mathrm{out}}$ degenerates as $|\eta_\perp| \to 0$; however, they are well-behaved when composed with $\Pi^{\omega,\pm}_{\theta} P_h$. The following lemma will play a basic role for our construction.

Lemma 9.3. *For any fixed* $\pm$, $\omega \in \mathbb{S}^3$, $n \in \mathbb{N}$, $h \in 2^{\mathbb{R}}$ *and* $* \in \{\mathrm{cone}, \mathrm{null}, \mathrm{out}\}$, *the multiplier*[10] $\theta^n \partial^{(n)}_{\xi}(\Pi^{\omega,\pm}_{*} \Pi^{\omega,\pm}_{\theta} P_h)$ *is disposable.*

Proof. In this proof, we take $h = 0$ by scaling, and fix $\pm = +$. Let $* \in \{\mathrm{cone}, \mathrm{null}\}$.

We begin with some elementary reductions. First, since $1 = \Pi^{\omega,\pm}_{\mathrm{cone}} + \Pi^{\omega,\pm}_{\mathrm{null}} + \Pi^{\omega,\pm}_{\mathrm{out}}$, and $\theta^n \partial^{(n)}_{\xi} \Pi^{\omega,\pm}_{\theta} P_0$ is disposable, it suffices to prove the lemma for just $\Pi^{\omega,\pm}_{\mathrm{cone}}$ and $\Pi^{\omega,\pm}_{\mathrm{null}}$. In this case, note that the symbol $\Pi^{\omega,\pm}_{*} \Pi^{\omega,\pm}_{\theta} m_h(\eta)$ (where m_h is the symbol of P_h) is compactly supported. Furthermore, the lemma is obvious if $\theta \gtrsim 1$, since then the symbol is smooth in ξ, σ, η on the unit scale. Therefore, we may assume that $\theta \ll 1$.

[10] We quantize $(\sigma, \eta) \mapsto (D_t, D_x)$.

We now consider the case $n = 0$, when there is no ξ-differentiation. We fix $\omega \in \mathbb{S}^3$. To ease our computation, we introduce the null coordinate system $(\underline{\upsilon}, \upsilon, \tilde{\eta}_\perp)$, where

$$\underline{\upsilon} = \sigma - \eta \cdot \omega, \quad \upsilon = \sigma + \eta \cdot \omega,$$

and $\tilde{\eta}_\perp \in \mathbb{R}^3$ are the coordinates for the constant $\underline{\upsilon}, \upsilon$-spaces. Observe that

$$\frac{\sigma + \eta \cdot \omega}{|\eta_\perp|^2 + |\sigma + \eta \cdot \omega|^2} = \frac{\upsilon}{|\tilde{\eta}_\perp|^2 + \upsilon^2} \simeq 1, \quad |\eta_\perp| = |\tilde{\eta}_\perp| \simeq \theta, \quad |\upsilon| \simeq \theta^2, \quad |\underline{\upsilon}| \simeq 1 \tag{9-34}$$

on the support of $\Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm} m_0$. Moreover, $\sigma = \sigma(\underline{\upsilon}, \upsilon, \tilde{\eta}_\perp)$ and $|\eta| = |\eta|(\underline{\upsilon}, \upsilon, \tilde{\eta}_\perp)$ are comparable to 1, and are also smooth on the unit scale on the support of $\Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm} m_0$. Recalling the definition of $\Pi_*^{\omega,\pm}$, it can be computed from (9-34) that

$$|\partial_{\underline{\upsilon}}^\alpha \partial_\upsilon^\beta \partial_{\tilde{\eta}_\perp}^\gamma \Pi_*^{\omega,\pm}| \lesssim \theta^{-2|\beta|-|\gamma|} \quad \text{on } \operatorname{supp} \Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm} m_0.$$

On the other hand,

$$|\partial_{\underline{\upsilon}}^\alpha \partial_\upsilon^\beta \partial_{\tilde{\eta}_\perp}^\gamma (\Pi_\theta^{\omega,\pm} m_0)| \lesssim \theta^{-|\gamma|} \quad \text{on } \operatorname{supp} \Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm} m_0,$$

so it follows that

$$|\partial_{\underline{\upsilon}}^\alpha \partial_\upsilon^\beta \partial_{\tilde{\eta}_\perp}^\gamma (\Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm})| \lesssim \theta^{-2|\beta|-|\gamma|}. \tag{9-35}$$

Furthermore, from (9-34) we have

$$|\operatorname{supp} \Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm} m_0| \lesssim \theta^5. \tag{9-36}$$

From these bounds, we see that the multiplier $\Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm} P_0$ has a kernel with a universal bound on the mass, and thus is disposable.

Finally, we sketch the proof in the case $n \geq 1$. We first claim that

$$|\partial_\xi^{(n)} (\Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm} m_0)| \lesssim \theta^{-n}. \tag{9-37}$$

Clearly $|\partial_\xi^{(n)} \Pi_\theta^{\omega,\pm}| \lesssim_n \theta^{-n}$, so it suffices to verify that $|\partial_\xi^{(n)} \Pi_*^{\omega,\pm}| \lesssim_n \theta^{-n}$ on the support of $\Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm} m_0$. Note that

$$|\partial_\xi^\alpha (\eta \cdot \omega)| \lesssim_{|\alpha|} \begin{cases} \theta, & |\alpha| = 1, \\ 1, & |\alpha| \geq 2 \end{cases} \quad \text{on } \operatorname{supp} \Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm} m_0. \tag{9-38}$$

Then recalling the definition of $\Pi_*^{\omega,\pm}$ and using the chain rule, the claim (9-37) follows. We remark that a differentiation in $\sigma + \eta \cdot \omega$ loses θ^{-2}, but we gain back a factor of θ through the chain rule and (9-38).

Next, we fix $\omega \in \mathbb{S}^3$ and start differentiating in $(\underline{\upsilon}, \upsilon, \tilde{\eta}_\perp)$. Using the chain rule, (9-38) and (9-34), it can be proved that

$$|\partial_{\underline{\upsilon}}^\alpha \partial_\upsilon^\beta \partial_{\tilde{\eta}_\perp}^\gamma \partial_\xi^{(n)} (\Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm})| \lesssim \theta^{-2|\beta|-|\gamma|} \theta^{-n}. \tag{9-39}$$

We omit the details. Combined with (9-36), we see that $\theta^n \partial_\xi^{(n)} \Pi_*^{\omega,\pm} \Pi_\theta^{\omega,\pm} P_0$ is disposable. □

As a corollary of the proof of Lemma 9.3, we obtain the following disposability statement.

Corollary 9.4. *For any fixed* $\pm$, $\omega \in \mathbb{S}^3$, $h, k \in 2^{\mathbb{R}}$ *and* $* \in \{\mathrm{cone}, \mathrm{null}, \mathrm{out}\}$, *the translation-invariant bilinear operator on* $\mathbb{R}^{1+4}$ *with symbol*

$$\Pi_{*}^{|\xi|^{-1}\xi,\pm} \Pi_{2^{\ell}}^{|\xi|^{-1}\xi,\pm} P_h(\sigma,\eta) P_k P_{\ell}^{\omega}(\xi)$$

is disposable.

Clearly, the same corollary holds with any of the continuous Littlewood–Paley projections P_h, P_k replaced by the discrete analogue.

We also record a lemma which describes how the operator $\Box$ acts in the presence of $\Pi_{\mathrm{cone}}^{\omega,\pm} \Pi_{\theta}^{\omega,\pm} P_h$.

Lemma 9.5. *For any fixed* $\pm$, $\omega \in \mathbb{S}^3$, $n \in \mathbb{N}$ *and* $h \in 2^{\mathbb{R}}$, *the multiplier*

$$(2^{-2h}\theta^{-2}\Box)\theta^{n}\partial_{\xi}^{(n)}(\Pi_{\mathrm{cone}}^{\omega,\pm}\Pi_{\theta}^{\omega,\pm} P_h) \tag{9-40}$$

is disposable.

Proof. We set $h = 0$ by scaling. The symbol of $\Box$ is $-\sigma^2 + |\eta|^2$. For a fixed ω, we introduce the null coordinate system $(\underline{\upsilon}, \upsilon, \eta_{\perp})$ as before. Then observe that

$$|\partial_{\underline{\upsilon}}^{\alpha}\partial_{\upsilon}^{\beta}\partial_{\tilde{\eta}_{\perp}}^{\gamma}(-\sigma^2+|\eta|^2)| = |\partial_{\underline{\upsilon}}^{\alpha}\partial_{\upsilon}^{\beta}\partial_{\tilde{\eta}_{\perp}}^{\gamma}(-\underline{\upsilon}\upsilon+|\tilde{\eta}_{\perp}|^2)| \lesssim \theta^2\theta^{-2|\beta|-|\gamma|}$$

on the support of $\Pi_{\mathrm{cone}}^{\omega,\pm}\Pi_{\theta}^{\omega,\pm}P_0$. The lemma follows by combining this bound with the proof of Lemma 9.3. ☐

9C. *Pseudodifferential renormalization operator.* In this subsection we define the pseudodifferential renormalization operator, and describe its main properties.

9C1. *Definition of the pseudodifferential renormalization operator.* As mentioned before, the aim for our renormalization operator is not to remove all of $\boldsymbol{P}A$, but only the most harmful (nonperturbative) part of it. This part is defined as

$$A_{j,<h}^{\mathrm{main},\pm} = \Pi_{\geq|\eta|^{\delta}}^{\omega,\pm}\Pi_{\mathrm{cone}}^{\omega,\pm}P_{<h}(\boldsymbol{P}A)_j. \tag{9-41}$$

Precisely, given a direction ω, it selects the region which is both near the cone in a parabolic fashion near the direction ω, but also away from ω, on an angular scale that is slowly decreasing as the frequency η of A approaches 0. We emphasize that this decomposition depends on ω, which is what will make our renormalization operator a pseudodifferential operator.

To account for the fact that our gauge group is noncommutative, and also to better take advantage of previous work in this area, we divide the construction of the renormalization operator in two steps. The first step is microlocal but linear, and mirrors the renormalization construction in the (MKG) case; see [Krieger et al. 2015; Oh and Tataru 2018]. Precisely, we define the intermediate symbol

$$\Psi_{\pm,<h} = -L_{\mp}^{\omega}\Delta_{\omega^{\perp}}^{-1}A_{j,<h}^{\mathrm{main},\pm}\omega^{j}. \tag{9-42}$$

Here the operator $L_{\mp}^{\omega}\Delta_{\omega^{\perp}}^{-1}$ is chosen as a good approximate inverse for $L_{\pm}^{\omega}$, within the frequency-localization region for $A_{j,<h}^{\mathrm{main},\pm}$. In effect this frequency-localization region is chosen exactly so that this

property holds within. This is based on the decomposition

$$-L^{\omega}_{\pm} L^{\omega}_{\mp} + \Delta_{\omega^\perp} = \Box,$$

which gives

$$L^{\omega}_{\pm} L^{\omega}_{\mp} \Delta^{-1}_{\omega^\perp} = 1 - \Box \Delta^{-1}_{\omega^\perp}.$$

Given $A^{\mathrm{main},\pm}_{j,<h}$ and $\Psi_{\pm,<h}$ as above, we define their Littlewood–Paley pieces as

$$A^{\mathrm{main},\pm}_{j,h} = \frac{d}{dh} A^{\mathrm{main},\pm}_{j,<h}, \qquad \Psi_{\pm,h} = \frac{d}{dh} \Psi_{\pm,<h}.$$

Now we come to the second step in the construction of the renormalization operator. This step is nonlinear but local, and is based on the construction of the renormalization operator in [Sterbenz and Tataru 2010a] for the corresponding wave map problem. Precisely, we solve the ODE

$$\begin{aligned} \frac{d}{dh} O_{<h,\pm} O^{-1}_{<h,\pm} &= \Psi_{\pm,h}, \\ \lim_{h\to-\infty} \|\partial_x O_{<h,\pm}(t,x,\xi)\|_{L^\infty} &= 0. \end{aligned} \tag{9-43}$$

Thus our renormalization is achieved via the paradifferential operator

$$\mathrm{Ad}(O_{\pm})_{<0},$$

where the localization to small frequencies is so that this operator preserves the unit dyadic frequency shell.

The parameter $\delta > 0$ is a universal constant, which is chosen below so that the parametrix construction go through. In particular, we take $0 < \delta < \frac{1}{100}$. Logically, it is fixed at the end of Section 10.

9C2. *Properties of the pseudodifferential renormalization operator.* Now we state the key properties satisfied by the renormalization operator $\mathrm{Ad}(O_{\pm})_{<0}$ that we just defined; see Theorems 9.6 and 9.9. Proofs of these results are the subjects of Sections 10 and 11, respectively.

Theorem 9.6 (mapping properties of the pseudodifferential renormalization operator). *Let A be a Lie-algebra-valued spatial 1-form on $I \times \mathbb{R}^4$ such that $A = P_{<-\kappa} A$ and*

$$\|\boldsymbol{P} A\|_{S^1[I]} \le M_0$$

for some $\kappa, M_0 > 0$. Let $\Psi_{\pm,<h}$, $\Psi_{\pm,h}$ and $O_{<h,\pm}$ be defined on $\mathbb{R}^{1+4}$ as above from the homogeneous-wave extension of $\boldsymbol{P}A$. Let Z be any of the spaces L^2_x, N or N^*.

(1) *For $\kappa > 20$, the following bounds hold:*

- (*boundedness*)

$$\|\mathrm{Op}(\mathrm{Ad}(O_{\pm})_{<0})(t,x,D) P_0\|_{Z\to Z} \lesssim_{M_0} 1. \tag{9-44}$$

- (*dispersive estimates*)

$$\|\mathrm{Op}(\mathrm{Ad}(O_{\pm})_{<0})(t,x,D) P_0\|_{S^\sharp_0 \to S_0} \lesssim_{M_0} 1. \tag{9-45}$$

(2) *For any $\varepsilon > 0$, there exist $\kappa_0(\varepsilon, M_0) \gg 1$ (independent of A_x) such that if $\kappa > \kappa_0(\varepsilon, M_0)$, then*

- (*derivative bounds*)

$$\|[\partial_t, \mathrm{Op}(\mathrm{Ad}(O_{\pm})_{<0})(t,x,D)] P_0\|_{Z\to Z} \lesssim \varepsilon. \tag{9-46}$$

- (*approximate unitarity*)

$$\|(\mathrm{Op}(\mathrm{Ad}(O_\pm)_{<0})(t,x,D)\,\mathrm{Op}(\mathrm{Ad}(O_\pm^{-1})_{<0})(D,s,y)-I)P_0\|_{Z\to Z}\lesssim\varepsilon, \tag{9-47}$$

where the implicit constants are universal.

(3) *There exists* $0<\delta_o(M_0)\ll 1$ (*independent of* A_x) *such that if, in addition to the above hypothesis,*

$$\|\boldsymbol{P}A_x\|_{\ell^\infty S^1[I]}<\delta_o(M_0), \tag{9-48}$$

then (9-44) *and* (9-45) *hold with universal constants. That is, for* $\kappa>20$ *we have*:

- (*boundedness with a universal constant*)

$$\|\mathrm{Op}(\mathrm{Ad}(O_\pm)_{<0})(t,x,D)P_0\|_{Z\to Z}\lesssim 1. \tag{9-44$'$}$$

- (*dispersive estimates with a universal constant*)

$$\|\mathrm{Op}(\mathrm{Ad}(O_\pm)_{<0})(t,x,D)P_0\|_{S_0^\sharp\to S_0}\lesssim 1. \tag{9-45$'$}$$

Here the frequency-localization operator P_0 can easily be replaced by a more general localization to $\{|\xi|\simeq 1\}$.

Remark 9.7. As we will see in the proof below, $\kappa_0(\varepsilon,M_0)\simeq_\varepsilon\log M_0$ and $\delta_o(M_0)\ll_{M_0}1$.

Remark 9.8. Note that the symbol of each of the above PDOs is independent of $\tau=\xi_0$, and thus it defines a PDO on $\mathbb{R}^4$ for each fixed t. By the mapping property $Z\to Z$ with $Z=L^2_x$, we mean that the PDO maps $L^2_x\to L^2_x$ for each fixed t, with a constant uniform in t.

Theorem 9.9 (renormalization error). *Let* A_α *be a* $\mathfrak{g}$*-valued* 1*-form on* $I\times\mathbb{R}^4$ *such that* $A_\alpha=P_{<-\kappa}A_\alpha$ *and* $\|\boldsymbol{P}A_x\|_{S^1[I]}\le M$ *for some* $\kappa,M>0$. *Let* $\varepsilon>0$. *Assume that* $\kappa>\kappa_1(\varepsilon,M)$ *and* (9-3)–(9-7) *hold for some functions* $\kappa_1(\varepsilon,M)\gg 1$ *and* $0<\delta_p(\varepsilon,M,\kappa_1)\ll 1$ *independent of* A_α (*to be specified below*). *Let* $\Psi_{\pm,<h}$, $\Psi_{\pm,h}$ *and* $O_{<h,\pm}$ *be defined as above from the homogeneous-wave extension of* $\boldsymbol{P}A_x$. *Then we have*

$$\|(\Box^p_{\boldsymbol{P}A}\,\mathrm{Op}(\mathrm{Ad}(O_\pm)_{<0})-\mathrm{Op}(\mathrm{Ad}(O_\pm)_{<0})\Box)P_0\|_{S^\sharp_{0,\pm}[I]\to N_{0,\pm}[I]}<\varepsilon. \tag{9-49}$$

Remark 9.10. As we will see later, $\kappa_1(\varepsilon,M)\simeq_\varepsilon\log M$ and $\delta_p(\varepsilon,M,\kappa_1)\ll_{M,\kappa_1}\varepsilon$.

9D. *Definition of the parametrix and proof of Theorem 9.1.* Our parametrix is given by

$$u(t)=\sum_\pm\Big(\frac12\,\mathrm{Op}(\mathrm{Ad}(O_\pm)_{<0})(t,x,D)e^{\pm it|D|}\,\mathrm{Op}(\mathrm{Ad}(O_\pm^1)_{<0})(D,0,y)(u_0\pm i|D|^{-1}u_1)\\ +\mathrm{Op}(\mathrm{Ad}(O_\pm)_{<0})(t,x,D)\frac{1}{|D|}K^\pm\,\mathrm{Op}(\mathrm{Ad}(O_\pm^{-1})_{<0})(D,s,y)f\Big), \tag{9-50}$$

where

$$K^\pm g(t)=\int_0^t e^{\pm i(t-s)|D|}g(s)\,ds.$$

With this definition, the proof of Theorem 9.1 starting from Theorems 9.6 and 9.9 is essentially identical to the corresponding proof in [Oh and Tataru 2018] and is omitted.

10. Mapping properties of the renormalization operator

10A. *Fixed-time pointwise bounds for the symbols Ψ and O.* Here we state fixed-time pointwise bounds for Ψ and O. We borrow these estimates from [Krieger and Tataru 2017], while carefully noting dependence of constants on the frequency envelope of $A = A_x$ in S^1. The bounds below are stated using continuous Littlewood–Paley projections P_h, but we note that the same bounds hold for discrete Littlewood–Paley projections as well.

We begin with pointwise bounds for the $\mathfrak{g}$-valued symbol $\Psi_{h,\pm}(t,x,\xi)$.

Lemma 10.1. *The following bounds hold*:

(1) *For $m \geq 0$ and $0 \leq n < \delta^{-1}$, we have*

$$|\partial_\xi^{(n)} \partial_x^{(m-1)} \nabla \Psi_{\pm,h}^{(\theta)}(t,x,\xi)| \lesssim 2^{mh} \theta^{\frac{1}{2}-n} \|A_h\|_{S^1}. \tag{10-1}$$

When $m = 0$, we interpret the expression on the left-hand side as $\partial_\xi^n \Psi_{\pm,h}^{(\theta)}$.

(2) *Let $\langle t-s, x-y\rangle^2 = 1 + |t-s|^2 + |x-y|^2$. We have*

$$|\Psi_{\pm,h}(t,x,\xi) - \Psi_{\pm,h}(s,y,\xi)| \lesssim \min\{2^h \langle t-s, x-y\rangle, 1\} \|A_h\|_{S^1}. \tag{10-2}$$

(3) *Finally, for $1 \leq n < \delta^{-1}$ we have*

$$|\partial_\xi^{(n)}(\Psi_{\pm,h}(t,x,\xi) - \Psi_{\pm,h}(s,y,\xi))| \lesssim \min\{2^h \langle t-s, x-y\rangle, 1\} 2^{-(n-\frac{1}{2})\delta h} \|A_h\|_{S^1}. \tag{10-3}$$

For a proof, we refer to [Krieger and Tataru 2017, Section 7.3]. As a corollary of (10-1) we have

$$|\nabla \Psi_{\pm,h}| \lesssim 2^h \|A_h\|_{S^1}. \tag{10-4}$$

Next, we consider the $\boldsymbol{G}$-valued symbol $O_{<h,\pm}$.

Lemma 10.2. *Let c_h be an admissible frequency envelope for A in S^1. Then the following bounds hold*:

(1) *For $0 \leq n < \delta^{-1}$, we have*

$$|\partial_\xi^{(n)}(O_{<h,\pm})_{;t,x}(t,x,\xi)| \lesssim_{\|A\|_{S^1}} 2^{(1-n\delta)h} c_h. \tag{10-5}$$

(2) *We have*

$$d(O_{<h,\pm}(t,x,\xi) O_{<h,\pm}^{-1}(s,y,\xi), \mathrm{Id}) \lesssim_{\|A\|_{S^1}} \log(1 + 2^h \langle t-s, x-y\rangle) c_h. \tag{10-6}$$

(3) *Finally, for $1 \leq n < \delta^{-1}$, we have*

$$|\partial_\xi^{(n-1)}(O_{<h,\pm}(t,x,\xi) O_{<h,\pm}^{-1}(s,y,\xi))_{;\xi}| \lesssim_{\|A\|_{S^1}} \min\{2^h \langle t-s, x-y\rangle, 1\}^{1-(n-\frac{1}{2})\delta} (1 + \langle t-s, x-y\rangle)^{(n-\frac{1}{2})\delta} c_h. \tag{10-7}$$

For a proof, we refer to [Krieger and Tataru 2017, Section 7.7].

10B. ***Decomposability calculus.*** To handle symbol multiplications, we use the decomposability calculus introduced in [Rodnianski and Tao 2004; Krieger and Sterbenz 2013], which allows us to roughly regard these operations as multiplication by a function in $L^p L^q$. In the present work, we need an interval-localized version in order to exploit small divisible norms.

Given $\theta \in 2^{-\mathbb{N}}$, consider a covering of the unit sphere $\mathbb{S}^3 = \{\omega \in \mathbb{R}^4 : |\xi| = 1\}$ by solid angular caps of the form $\{\omega \in \mathbb{S}^3 : |\phi - \omega| < \theta\}$ with uniformly finite overlaps. We index these caps by their centers $\phi \in \mathbb{S}^3$, and denote by $\{(m_\theta^\phi)^2(\omega)\}$ the associated nonnegative smooth partition of unity on $\mathbb{S}^3$.

Let I be an interval. Consider a $\mathrm{End}(\mathfrak{g})$-valued symbol $c(t,x,\xi)$ on $I_t \times \mathbb{R}^4_x \times \mathbb{R}^4_\xi$, which is zero homogeneous in ξ, i.e., depends only on the angular variable $\omega = \xi/|\xi|$. We say that $c(t,x,\xi)$ is *decomposable in* $L^q L^r[I]$ if $c = \sum_\theta c^{(\theta)}$, $\theta \in 2^{-\mathbb{N}}$ and

$$\sum_\theta \|c^{(\theta)}\|_{D_\theta L^q L^r[I]} < \infty, \tag{10-8}$$

where

$$\|c^{(\theta)}\|_{D_\theta L^q L^r[I]} = \left\| \left(\sum_{n=0}^{40} \sum_\phi \sup_\omega (m_\theta^\phi(\omega) \|\theta^n \partial_\xi^{(n)} c^{(\theta)}\|_{L^r_x})^2 \right)^{\frac{1}{2}} \right\|_{L^q_t[I]}. \tag{10-9}$$

We define $\|c\|_{DL^q L^r[I]}$ to be the infimum of (10-8) over all possible decompositions $c = \sum_\theta c^{(\theta)}$. In what follows, we will use the convention of omitting $[I]$ when $I = \mathbb{R}$.

In the following lemma, we collect some basic properties of the symbol class $DL^q L^r[I]$.

Lemma 10.3. (1) *For any two intervals such that* $I \subset I'$, *we have*

$$\|c\|_{DL^q L^r[I]} \le \|c\|_{DL^q L^r[I']}.$$

(2) *For any symbols* $c \in DL^{q_1} L^{r_1}[I]$ *and* $d \in DL^{q_2} L^{r_2}[I]$, *its product obeys the Hölder-type bound*

$$\|cd\|_{DL^q_t L^r_x[I]} \lesssim \|c\|_{DL^{q_1} L^{r_1}[I]} \|d\|_{DL^{q_2} L^{r_2}[I]},$$

where $1 \le q_1, q_2, q, r_1, r_2, r \le \infty$, $\frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{q}$ *and* $\frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{r}$.

(3) *Let* $a(t,x,\xi)$ *be an* $\mathrm{End}(\mathfrak{g})$*-valued smooth symbol on* $I \times \mathbb{R}^4_x \times \mathbb{R}^4_\xi$ *whose left quantization* $\mathrm{Op}(a)$ *satisfies the fixed-time bound*

$$\sup_{t \in I} \|\mathrm{Op}(a)(t,x,D)\|_{L^2 \to L^2} \le C_a.$$

Then for any symbol $c \in DL^q L^r$, *we have the space-time bound*

$$\|\mathrm{Op}(ac)(t,x,D)\|_{L^{q_1} L^2[I] \to L^{q_2} L^{r_2}[I]} \lesssim C_a \|c\|_{DL^q L^r[I]},$$

where $1 \le q_1, q_2, q, r_2, r \le \infty$, $\frac{1}{q_1} + \frac{1}{q} = \frac{1}{q_2}$ *and* $\frac{1}{2} + \frac{1}{r} = \frac{1}{r_2}$. *An analogous statement holds in the case of right quantization.*

The proof is essentially the same as the global-in-time versions in [Krieger and Sterbenz 2013, Chapter 10] and [Krieger et al. 2015, Lemma 7.1]; we omit the details.

10C. *Decomposability bounds for A, Ψ and O.* Here we collect some decomposability bounds for A, Ψ and O that we will use in our proof of Theorems 9.6 and 9.9. As before, we state the bounds using continuous Littlewood–Paley projections P_h, but note that the same bounds hold for discrete Littlewood–Paley projections as well. For simplicity of notation, we will usually write $\|G\|_{DL^q L^r} = \|\mathrm{ad}(G)\|_{DL^q L^r}$ for a $\mathfrak{g}$-valued symbol G and $\|O\|_{DL^q L^r} = \|\mathrm{Ad}(O)\|_{DL^q L^r}$ for a $\boldsymbol{G}$-valued symbol O.

For any $\theta > 0$, $h \in \mathbb{R}$ and $* \in \{\mathrm{cone}, \mathrm{null}, \mathrm{out}\}$, recall the definition

$$A^{(\theta)}_{\alpha,h,*,\pm} = P_h \Pi^{\omega,\pm}_{*} \Pi^{\omega,\pm}_{\theta} (\boldsymbol{P} A)_{\alpha}.$$

As before, we will often omit the subscript x for simplicity, and write $A^{(\theta)}_{h,*,\pm} = A^{(\theta)}_{x,h,*,\pm}$ etc.

These symbols obey the following global-in-time decomposability bounds:

Lemma 10.4. *For $q \geq 2$ and $* \in \{\mathrm{cone}, \mathrm{null}, \mathrm{out}\}$, we have*

$$\|A^{(\theta)}_{h,*,\pm} \cdot \omega\|_{DL^q L^\infty} \lesssim 2^{(1-\frac{1}{q})h} \theta^{\frac{5}{2}-\frac{2}{q}} \|A_h\|_{S^1}, \tag{10-10}$$

$$\|A^{(\theta)}_{0,h,*,\pm}\|_{DL^q L^\infty} \lesssim 2^{(1-\frac{1}{q})h} \theta^{\frac{5}{2}-\frac{2}{q}} \|A_{0,h}\|_{Y^1}. \tag{10-11}$$

Furthermore, for $* = \mathrm{cone}$ we have

$$\|\Box A^{(\theta)}_{h,\mathrm{cone},\pm} \cdot \omega\|_{DL^q L^\infty} \lesssim 2^{(3-\frac{1}{p})h} \theta^{\frac{9}{2}-\frac{2}{q}} \|A_h\|_{S^1}, \tag{10-12}$$

$$\|\Delta^{-1}_{\omega^\perp} \Box A^{(\theta)}_{h,\mathrm{cone},\pm} \cdot \omega\|_{DL^q L^\infty} \lesssim 2^{(1-\frac{1}{p})h} \theta^{\frac{5}{2}-\frac{2}{q}} \|A_h\|_{S^1}. \tag{10-13}$$

Proof. The symbols $(\theta \partial_\omega)^n (\Pi^{\omega,\pm}_{*} \Pi^{\omega,\pm}_{\theta})$ are smooth, homogeneous and uniformly bounded, and the corresponding multipliers are disposable for fixed Ω. Then the bounds (10-10) and (10-11) follow by Bernstein's inequality using the Strichartz component of the S^1 norm, and, respectively, the $L^2 \dot{H}^{1/2}$ component of the ∇Y^1 norm.

For the bounds (10-12) and (10-13) we need in addition to consider the size of the symbol of $\Box$, and, respectively, $\Delta^{-1}_{\omega^\perp}$, within the support of $P_h \Pi^{\omega,\pm}_{\mathrm{cone}} \Pi^{\omega,\pm}_{\theta}$. This is $\theta^2 2^{2h}$, respectively $\theta^{-2} 2^{-2h}$. Precisely, we have the representations

$$\Box P_h \Pi^{\omega,\pm}_{\mathrm{cone}} \Pi^{\omega,\pm}_{\theta} = \theta^2 2^{2h} \boldsymbol{O} \Pi^{\omega,\pm}_{\mathrm{cone}} \Pi^{\omega,\pm}_{\theta}, \quad \Delta^{-1}_{\omega^\perp} P_h \Pi^{\omega,\pm}_{*} \Pi^{\omega,\pm}_{\theta} = \theta^{-2} 2^{-2h} \boldsymbol{O} \Pi^{\omega,\pm}_{\mathrm{cone}} \Pi^{\omega,\pm}_{\theta},$$

with $\boldsymbol{O}$ disposable; see, e.g., Lemma 9.5. Then (10-12) and (10-13) immediately follow from (10-10). $\Box$

Next, we consider the phase $\Psi_\pm$, which was defined in (9-42). Given $\theta > 0$ and $h \in \mathbb{R}$, let

$$\Psi^{(\theta)}_{h,\pm} = P_h \Pi^{\omega,\pm}_{\theta} \Psi_\pm.$$

We have the following global-in-time decomposability bounds.

Lemma 10.5. *For $q, r \geq 2$ and $\frac{2}{q} + \frac{3}{r} \leq \frac{3}{2}$, we have*

$$\|(\Psi^{(\theta)}_{h,\pm}, 2^{-h} \nabla \Psi^{(\theta)}_{h,\pm})\|_{DL^q L^r} \lesssim 2^{-(\frac{1}{q}+\frac{4}{r})h} \theta^{\frac{1}{2}-\frac{2}{q}-\frac{3}{r}} \|A_h\|_{S^1}. \tag{10-14}$$

In addition, suppose that $\theta \lesssim 2^a$ for some $a \in -\mathbb{N}$. Then for $q, r \geq 2$, we also have

$$\|Q_{h+2a}(\Psi^{(\theta)}_{h,\pm}, 2^{-h} \nabla \Psi^{(\theta)}_{h,\pm})\|_{DL^q L^r} \lesssim 2^{-(\frac{1}{q}+\frac{4}{r})h} 2^{-\frac{2}{q}a} \theta^{\frac{1}{2}-\frac{3}{r}} \|A_h\|_{S^1}. \tag{10-15}$$

Furthermore,

$$\|\Box \Psi^{(\theta)}_{h,\pm}\|_{DL^2 L^\infty} \lesssim \theta^{\frac{3}{2}} 2^{\frac{3}{2}h} \|A_h\|_{S^1}. \tag{10-16}$$

Proof. Observing that within the support of $P_h \Pi^{\omega,\pm}_{\text{cone}} \Pi^{\omega,\pm}_{\theta}$ the symbol $L^{\mp} \Delta^{-1}_{\omega^\perp}$ has the form $2^{-h}\theta^{-2}\boldsymbol{O}$ with $\boldsymbol{O}$ disposable and depending smoothly on ω on the θ scale, the first bound (10-14) is again a direct consequence of the Strichartz bounds in the S^1 norm for A.

For (10-15) it suffices to prove the case $p = q = 2$ and then use Bernstein's inequality. But in this case it suffices to use the $X^{1,1/2}_\infty$ component of the S^1 norm at fixed modulation.

For the last bound (10-16) it suffices to combine the $L^2 L^\infty$ case of (10-14) with Lemma 9.5. □

We now consider the $\boldsymbol{G}$-valued symbol $O_{<h,\pm}$, which was defined in (9-43). It obeys the following global-in-time decomposability bounds.

Lemma 10.6. *Let c_h be an admissible frequency envelope for A in S^1. Then for any $q > 4$, we have*

$$\|(O_{<h,\pm;x}, O_{<h,\pm;t})\|_{DL^q L^\infty} \lesssim_{\|A\|_{S^1}} 2^{(1-\frac{1}{q})h} c_h. \tag{10-17}$$

When $q = 2$, an analogous bound with a slight loss holds:

$$\|(O_{<h,\pm;x}, O_{<h,\pm;t})\|_{DL^2 L^\infty} \lesssim_{\|A\|_{S^1}} 2^{\frac{1}{2}(1-\delta)h} c_h. \tag{10-18}$$

Proof. These bounds are a consequence of the $\Psi^{(\theta)}_{h,\pm}$ bounds in the previous lemma. The proof is similar to the proof of the similar result in [Krieger and Tataru 2017, Lemma 7.9] and is omitted. We note that the constraint $q > 4$ in the first bound is to prevent losses in the θ summation in (10-14). □

Finally, we consider *interval-localized* decomposability bounds, which will be needed to exploit divisibility (i.e., the hypothesis (9-3)) to gain smallness.

Lemma 10.7. *Let $|I| \geq 2^{-h-\kappa}$, where $h \in \mathbb{R}$ and $\kappa \geq 0$. For $q \geq 2$, we have*

$$\|\Psi^{(\theta)}_h\|_{DL^q L^\infty[I]} \lesssim 2^{C\kappa}\theta^{-C} 2^{-h} \|A_h\|_{L^q L^\infty[I]}, \tag{10-19}$$

$$\|\Delta^{-1}_{\omega^\perp} \Box (\omega \cdot A^{(\theta)}_{h,\text{cone},\pm})\|_{DL^q L^\infty[I]} \lesssim 2^{C\kappa}\theta^{-C} \|A_h\|_{L^q L^\infty[I]}, \tag{10-20}$$

$$\|\omega \cdot A^{(\theta)}_h\|_{DL^q L^\infty[I]} \lesssim 2^{C\kappa}\theta^{-C} \|A_h\|_{L^q L^\infty[I]}, \tag{10-21}$$

$$\|\omega \cdot A^{(\theta)}_{0,h}\|_{DL^q L^\infty[I]} \lesssim 2^{C\kappa}\theta^{-C} \|A_{0,h}\|_{L^q L^\infty[I]}. \tag{10-22}$$

Proof. We will prove (10-19), and leave the similar cases of (10-20), (10-21), (10-22) to the reader.

By scaling, we set $h = 0$. By the definition of the class $DL^q L^\infty[I]$, we have

$$\begin{aligned} \|\Psi^{(\theta)}_0\|_{DL^q L^\infty[I]} &\lesssim \theta^{-2} \bigg(\sum_{n=0}^{40} \sum_{\phi} \sup_{\omega} \|m^{\phi}_{\theta}(\omega) \theta^n \partial^{(n)}_{\xi} \Pi^{\omega}_{\theta} \Pi^{\omega}_{\text{cone}} P_0(\omega \cdot \boldsymbol{P}A)\|^2_{L^q L^\infty[I]} \bigg)^{\frac{1}{2}} \\ &\lesssim \theta^{-C} \sum_{n=0}^{40} \|\theta^n \partial^{(n)}_{\xi} \Pi^{\omega}_{\theta} \Pi^{\omega}_{\text{cone}} P_0(\omega \cdot \boldsymbol{P}A)\|_{L^q L^\infty[I]}. \end{aligned}$$

Fix $n \in [1, 40]$ and $\omega \in \mathbb{S}^3$. From the proof of Lemma 9.3, we see that the projection $\theta^{n'} \partial^{(n')}_{\xi} \Pi^{\omega}_{\theta} \Pi^{\omega}_{\text{cone}} P_0$, when viewed as a Fourier multiplier in (σ, η), has a symbol which is supported in a space-time cube of radius $\lesssim 1$, and its derivatives (up to 40, say) are bounded by θ^{-C} for some large universal constant C. Moreover, we have $|\theta^{n''}\partial_\xi^{(n'')}\omega| \lesssim_{n''} 1$. Denoting by χ_I^0 a generalized cutoff adapted at the unit scale as in (4-22), we have

$$\|\theta^n \partial_\xi^{(n)} \Pi_\theta^\omega \Pi_{\mathrm{cone}}^\omega P_0(\omega\cdot \boldsymbol{P}A)\|_{L^q L^\infty[I]} \lesssim \theta^{-C}\|\chi_I^0 P_0 A\|_{L^q L^\infty}.$$

Recall that A is extended outside I by homogeneous waves. By Proposition 4.10, the last expression is bounded by

$$\lesssim 2^{C\kappa}\theta^{-C}\|P_0 A\|_{L^q L^\infty[I]},$$

which proves (10-19). □

10D. *Collection of symbol bounds.* Before we continue, we introduce the quantity M_σ, which collects various symbol bounds that we have so far.

We fix large enough N and a small universal constant $\delta_\sigma > 0$. Then we let $M_\sigma > 0$ be the minimal constant such that:

- The following pointwise bounds hold for all $0 \le n \le \delta^{-1}$ and $0 \le m \le N$:

$$|\partial_\xi^{(n)}\partial_x^{(m-1)}\nabla\Psi_{\pm,h}^{(\theta)}| \le 2^{mh}\theta^{\frac12-n}M_\sigma,$$

$$|\Psi_{\pm,h}(t,x,\xi)-\Psi_{\pm,h}(s,y,\xi)| \le \min\{2^h\langle t-s,x-y\rangle,1\}M_\sigma,$$

$$|\partial_\xi^{(n)}(\Psi_{\pm,h}(t,x,\xi)-\Psi_{\pm,h}(s,y,\xi))| \le \min\{2^h\langle t-s,x-y\rangle,1\}2^{-(n-\frac12)\delta h}M_\sigma,$$

$$|\partial_\xi^{(n)}(O_{<h,\pm})_{;t,x}(t,x,\xi)| \le 2^{(1-n\delta)h}M_\sigma,$$

$$d(O_{<h,\pm}(t,x,\xi)O_{<h,\pm}^{-1}(s,y,\xi),\mathrm{Id}) \le \log(1+2^h\langle t-s,x-y\rangle)M_\sigma,$$

$$|\partial_\xi^{(n-1)}(O_{<h,\pm}(t,x,\xi)O_{<h,\pm}^{-1}(s,y,\xi))_{;\xi}| \le \min\{2^h\langle t-s,x-y\rangle,1\}^{1-(n-\frac12)\delta}(1+\langle t-s,x-y\rangle)^{(n-\frac12)\delta}M_\sigma.$$

- The following decomposability bounds hold for all $* \in \{\mathrm{cone},\mathrm{null},\mathrm{out}\}$, $q,r\ge 2$ and $\frac2q+\frac3r\le\frac32$:

$$\begin{aligned}
\|A_{h,*,\pm}^{(\theta)}\cdot\omega\|_{DL^qL^\infty} &\le 2^{(1-\frac1q)h}\theta^{\frac52-\frac2q}M_\sigma,\\
\|A_{0,h,*,\pm}^{(\theta)}\|_{DL^qL^\infty} &\le 2^{(1-\frac1q)h}\theta^{\frac52-\frac2q}M_\sigma,\\
\|\Box A_{h,\mathrm{cone},\pm}^{(\theta)}\cdot\omega\|_{DL^qL^\infty} &\le 2^{(3-\frac1p)h}\theta^{\frac92-\frac2q}M_\sigma,\\
\|\Delta_{\omega^\perp}^{-1}\Box A_{h,\mathrm{cone},\pm}^{(\theta)}\cdot\omega\|_{DL^qL^\infty} &\le 2^{(1-\frac1p)h}\theta^{\frac52-\frac2q}M_\sigma,\\
\|(\Psi_{h,\pm}^{(\theta)},2^{-h}\nabla\Psi_{h,\pm}^{(\theta)})\|_{DL^qL^r} &\le 2^{-(\frac1q+\frac4r)h}\theta^{\frac12-\frac2q-\frac3r}M_\sigma,\\
\|Q_{h+2a}(\Psi_{h,\pm}^{(\theta)},2^{-h}\nabla\Psi_{h,\pm}^{(\theta)})\|_{DL^qL^r} &\le 2^{-(\frac1q+\frac4r)h}2^{-\frac2q a}\theta^{\frac12-\frac3r}M_\sigma \quad (\theta\lesssim 2^a\lesssim 1),\\
\|\Box\Psi_{h,\pm}^{(\theta)}\|_{DL^2L^\infty} &\le \theta^{\frac32}2^{\frac32 h}M_\sigma,\\
\|(O_{<h,\pm;x},O_{<h,\pm;t})\|_{DL^qL^\infty} &\le 2^{(1-\frac1q)h}M_\sigma \quad (q\ge 4+\delta_\sigma),\\
\|(O_{<h,\pm;x},O_{<h,\pm;t})\|_{DL^2L^\infty} &\le 2^{\frac12(1-\delta)h}M_\sigma.
\end{aligned}$$

By the preceding results, there exists an M_σ such that

$$M_\sigma \lesssim_M \|A\|_{\ell^\infty S^1} + \|A_0\|_{\ell^\infty Y^1}. \tag{10-23}$$

In particular, note that all of the above symbol bounds are small if $\|A\|_{\ell^\infty S^1}$ and $\|A_0\|_{\ell^\infty Y^1}$ are.

10E. *Oscillatory integral bounds.* Given a smooth function a, let

$$K^a_{<0}(t,x;s,y) = \int \mathrm{Ad}(O_{<h,\pm})_{<0}(t,x,\xi)a(\xi)e^{\pm i(t-s)|\xi|}e^{i\xi\cdot(x-y)}\,\mathrm{Ad}(O^{-1}_{<h,\pm})_{<0}(\xi,y,s)\,\frac{d\xi}{(2\pi)^4}.$$

Lemma 10.8. *For a sufficiently small universal constant $\delta > 0$, the following bounds hold for the kernel $K^a_{<0}(t,x;s,y)$:*

(1) *Assume that a is a smooth bump function on the unit scale. Then*

$$|K^a_{<0}(t,x;s,y)| \lesssim_{M_\sigma} \langle t-s\rangle^{-\frac{3}{2}}\langle |t-s|-|x-y|\rangle^{-100}. \tag{10-24}$$

(2) *Let $a = a_{\mathcal{C}}$ be a smooth bump function on a radially oriented rectangular box $\mathcal{C}$ of size $2^k \times (2^{k+\ell})^3$, where $k, \ell \le 0$. Then*

$$|K^a_{<0}(t,x;s,y)| \lesssim_{M_\sigma} 2^{4k+3\ell}\langle 2^{2(k+\ell)}(t-s)\rangle^{-\frac{3}{2}}\langle 2^k(|t-s|-|x-y|)\rangle^{-100}. \tag{10-25}$$

(3) *Let $a = a_{\mathcal{C}}$ be a smooth bump function on a radially oriented rectangular box $\mathcal{C}$ of size $1 \times (2^\ell)^3$, where $\ell \le 0$. Let $\omega \in \mathbb{S}^3$ be at angle $\simeq 2^\ell$ from $\mathcal{C}$. Then, for $t-s = (x-y)\cdot\omega + O(1)$,*

$$|K^a_{<0}(t,x;s,y)| \lesssim_{M_\sigma} 2^{3\ell}\langle 2^{2\ell}(t-s)\rangle^{-100}\langle 2^\ell(x'-y')\rangle^{-100}, \tag{10-26}$$

where $x' = x-(x\cdot\omega)\omega$ and $y' = y-(y\cdot\omega)\omega$.

This lemma is proved as in [Krieger and Tataru 2017, Section 8.1] by stationary phase, using the symbol bounds in Lemmas 10.1 and 10.2.

10F. *Fixed-time L^2 bounds.* The goal of this subsection is to prove (9-44), (9-46), (9-47) and (9-44)′ for $Z = L^2$. The common key ingredient is the following fixed-time L^2 estimate:

Proposition 10.9. *For $\delta > 0$ sufficiently small, there exists $\delta_{(0)} > 0$ such that the following statement holds. Let $h + 10 \le k \le 0$. Then for every fixed t, we have*

$$\big\|\big(\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_{<k})(x,D)\,\mathrm{Op}(\mathrm{Ad}(O^{-1}_{<h,\pm})_{<k})(D,y)-1\big)P_0\big\|_{L^2\to L^2} \lesssim_{M_\sigma} 2^{\delta_{(0)}h}+2^{-10(k-h)}. \tag{10-27}$$

Lemma 10.10. *There exists $\delta_{(0)} > 0$ such that the following statement holds. Let $h \le 0$ and $a(\xi)$ be a smooth bump function adapted to $\{|\xi| \lesssim 1\}$. Then for every fixed t, we have*

$$\|\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm}))(x,D)a(D)\,\mathrm{Op}(\mathrm{Ad}(O^{-1}_{<h,\pm}))(D,y) - a(D)\|_{L^2\to L^2} \lesssim_{M_\sigma} 2^{\delta_{(0)}h}. \tag{10-28}$$

Proof. For simplicity of notation, we omit $\pm$ in $O_{<h,\pm}$, $O^{-1}_{<h,\pm}$ and $\Psi_{\pm,h}$. Following the hypothesis, we fix $t \in \mathbb{R}$.

The idea is to derive a kernel estimate as in Lemma 10.8, but taking into account the frequency gap. The kernel of the End($\mathfrak{g}$)-valued operator in (10-28) is given by

$$K_{<h}(x,y) = \int (\mathrm{Ad}(O_{<h}(x,\xi)O_{<h}^{-1}(y,\xi)) - 1)a(\xi)e^{i(x-y)\cdot\xi}\,\frac{d\xi}{(2\pi)^4}. \tag{10-29}$$

We obtain two different estimates depending on whether $|x-y| \lesssim 2^{-\delta_{(0)}h}$ or $|x-y| \gtrsim 2^{-\delta_{(0)}h}$.

Case 1: $|x-y| \lesssim 2^{-\delta_{(0)}h}$. In this case, we use the fundamental theorem of calculus and simply bound

$$\begin{aligned}|K_{<h}(x,y)| &\lesssim \iint_{-\infty}^{h} \left|\frac{d}{d\ell}(\mathrm{Ad}(O_{<\ell}(x,\xi)O_{<\ell}^{-1}(y,\xi)))\right| |a(\xi)|\,d\ell\,d\xi \\ &\lesssim \sup_{|\xi|\lesssim 1} \int_{-\infty}^{h} \left|\frac{d}{d\ell}(\mathrm{Ad}(O_{<\ell}(x,\xi)O_{<\ell}^{-1}(y,\xi)))\right| d\ell.\end{aligned}$$

By the algebraic property

$$O[u,v]O^{-1} = [OuO^{-1}, OvO^{-1}], \quad O \in \boldsymbol{G},\ u,v \in \mathfrak{g},$$

we have

$$\mathrm{ad}(u)\,\mathrm{Ad}(O) = \mathrm{Ad}(O)\,\mathrm{ad}(\mathrm{Ad}(O^{-1})u), \quad \mathrm{Ad}(O^{-1})\,\mathrm{ad}(u) = \mathrm{ad}(\mathrm{Ad}(O^{-1})u)\,\mathrm{Ad}(O^{-1}).$$

Therefore,

$$\begin{aligned}&\frac{d}{d\ell}(\mathrm{Ad}(O_{<\ell}(x,\xi)O_{<\ell}^{-1}(y,\xi))) \\ &\qquad = \mathrm{ad}(\Psi_\ell)\,\mathrm{Ad}(O_{<\ell})(x,\xi)\,\mathrm{Ad}(O_{<\ell}^{-1})(y,\xi) - \mathrm{Ad}(O_{<\ell})(x,\xi)\,\mathrm{Ad}(O_{<\ell}^{-1})\,\mathrm{ad}(\Psi_\ell)(y,\xi) \\ &\qquad = \mathrm{Ad}(O_{<\ell})(x,\xi)\,\mathrm{ad}(\mathrm{Ad}(O_{<\ell}^{-1})\Psi_\ell(x,\xi) - \mathrm{Ad}(O_{<\ell}^{-1})\Psi_\ell(y,\xi))\,\mathrm{Ad}(O_{<\ell}^{-1})(y,\xi).\end{aligned}$$

Then using the fact that the norm on End($\mathfrak{g}$) is invariant under $\mathrm{Ad}(O)$ for any $O \in \boldsymbol{G}$, we have

$$\left|\frac{d}{d\ell}(\mathrm{Ad}(O_{<\ell}(x,\xi)O_{<\ell}^{-1}(y,\xi)))\right| = |\mathrm{Ad}(O_{<\ell}^{-1})\Psi_\ell(x,\xi) - \mathrm{Ad}(O_{<\ell}^{-1})\Psi_\ell(y,\xi)|.$$

By the symbol bounds (10-5) and (10-4), we have $|\partial_x(\mathrm{Ad}(O_{<\ell}^{-1})\Psi_\ell)| \lesssim_{M_\sigma} 2^\ell$. Thus, by the mean value theorem,

$$\left|\frac{d}{d\ell}(\mathrm{Ad}(O_{<\ell}(x,\xi)O_{<\ell}^{-1}(y,\xi)))\right| \lesssim_{M_\sigma} 2^\ell 2^{-\delta_{(0)}h}.$$

Integrating in ℓ, we arrive at

$$|K_{<h}(x,y)| \lesssim_{M_\sigma} 2^{(1-\delta_{(0)})h}. \tag{10-30}$$

Case 2: $|x-y| \gtrsim 2^{-\delta_{(0)}h}$. Here, the idea is to repeatedly integrate by parts in ξ. Since

$$\partial_\xi\,\mathrm{Ad}(O_{<h}(x,\xi)O_{<h}^{-1}(y,\xi)) = \mathrm{ad}((O_{<h}(x,\xi)O_{<h}^{-1}(y,\xi))_{;\xi})\,\mathrm{Ad}(O_{<h}(x,\xi)O_{<h}^{-1}(y,\xi)),$$

the symbol bound (10-5) implies

$$|\partial_\xi^{(n)}\,\mathrm{Ad}(O_{<h}(x,\xi)O_{<h}^{-1}(y,\xi))| \lesssim_{n,M_\sigma} 2^{\delta|n-\frac{1}{2}|h}.$$

Therefore, integrating by parts in ξ for N times in (10-29), we obtain

$$|K_{<h}(x,y)| \lesssim_{\delta,N,M_\sigma} \frac{1}{|x-y|^{(1-\delta)N+\frac{1}{2}\delta}} \quad \text{for } |x-y| \gtrsim 2^{-\delta_{(0)}h},\ 0 \le N < \delta^{-1}.$$

Finally, combining Cases 1 and 2, we obtain

$$\sup_x \int |K_{<h}(x,y)|\,dy + \sup_y \int |K_{<h}(x,y)|\,dx \lesssim_{M_\sigma} 2^{(1-5\delta_{(0)})h} \lesssim 2^{\delta_{(0)}h}$$

provided that $\delta_{(0)}$ is small enough. Bound (10-28) now follows. □

Corollary 10.11. *For any $k \in \mathbb{R}$ we have*

$$\|\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm}))(x,D)P_0\|_{L^2\to L^2} \lesssim_{M_\sigma} 1, \tag{10-31}$$

$$\|\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_{<k})(x,D)P_0\|_{L^2\to L^2} \lesssim_{M_\sigma} 1. \tag{10-32}$$

Proof. The first bound follows by a TT^*-argument from Lemma 10.10. Next, note that $\mathrm{Ad}(O_{<h,\pm})_{<k}(x,\xi)$ is simply a smooth average of translates of $\mathrm{Ad}(O_{<h,\pm})(x,\xi)$ in x. Therefore, the second bound follows from the first by translation invariance of L^2. □

Next, we borrow a lemma from [Krieger and Tataru 2017], which handles $\mathrm{Ad}(O_{<h,\pm})_k$ when k is large compared to h.

Lemma 10.12. *Let $t \in \mathbb{R}$, $h \le 0$ and $k \ge h+10$. Then we have*

$$\|\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_k)(t,x,D)P_0\|_{L^2\to L^2} \lesssim_{M_\sigma} 2^{-10(k-h)}. \tag{10-33}$$

Furthermore, for $1 \le q \le p \le \infty$, $h \le 0$ and $k \ge h+10$, we have

$$\|\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_k)(t,x,D)P_0\|_{L^pL^2\to L^qL^2} \lesssim_{M_\sigma} 2^{(\frac{1}{p}-\frac{1}{q})h}2^{-10(k-h)}. \tag{10-34}$$

Same estimates hold for the right quantization $\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_k(D,s,y))$.

Remark 10.13. The specific factor 10 in the gain $2^{-10(k-h)}$ is not of any significance, but it is important to note that this number is much bigger than 1; see the proof of Proposition 10.14 below.

For the proof, we refer to [Krieger and Tataru 2017, Proof of Lemma 8.4] or [Oh and Tataru 2018, Proof of Lemma 9.11].

Proof of Proposition 10.9. Due to the frequency localization of the symbols in (10-27), we can harmlessly insert a multiplier $a(D)$ whose symbol is a smooth bump function $a(\xi)$ adapted to $\{|\xi| \lesssim 1\}$, and then discard P_0 to replace (10-27) by

$$\|\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_{<k})(x,D)a(D)\,\mathrm{Op}(\mathrm{Ad}(O^{-1}_{<h,\pm})_{<k})(D,y)-a(D)\|_{L^2\to L^2} \lesssim_{M_\sigma} 2^{\delta_{(0)}h}+2^{-10(k-h)}.$$

Now it suffices to combine the last two lemmas. □

Proof of (9-44), (9-46), (9-47) and (9-44)′ in the case $Z = L^2$. By a TT^* argument, the bounds (9-44) and (9-44)′ are immediate consequences of (10-27). Also from (10-27) we obtain the estimate (9-47) with a constant $2^{-\delta_{(0)}\kappa}$, which is less than ε if κ is chosen large enough depending only on M_0.

Finally, for (9-46) we compute

$$\partial_t(\mathrm{Ad}(O))_{<0} = (\mathrm{ad}(O_{;t})\,\mathrm{Ad}(O))_{<0};$$

therefore it suffices to combine the decomposability bound (10-17) for $O_{;t}$ with $q = \infty$ with (10-31). The former bound yields a $2^{-\kappa}$ factor which again yields ε smallness if κ is large enough. □

10G. *Space-time L^2L^2 bounds.* Next, we establish (9-44), (9-46), (9-47) and (9-44)′ when $Z = N$ or N^*. As we will see below, (9-44), (9-46) and (9-44)′ follow from the arguments in [Krieger and Tataru 2017]. In the bulk of this subsection, we focus on the task of establishing (9-47).

To state the key estimates, it is convenient to set up some notation. We introduce the compound $\boldsymbol{G}$-valued symbol

$$\boldsymbol{O}_{<h,\pm}(t,x,s,y,\xi) = O_{<h,\pm}(t,x,\xi)O^{-1}_{<h,\pm}(s,y,\xi).$$

The quantization of $\mathrm{Ad}(\boldsymbol{O}_{<h,\pm})$, which is an $\mathrm{End}(\mathfrak{g})$-valued compound symbol, takes the form

$$\mathrm{Op}(\mathrm{Ad}(\boldsymbol{O}_{<h,\pm}))(t,x,D,y,s) = \mathrm{Op}(\mathrm{Ad}(O_{<h,\pm}))(t,x,D)\,\mathrm{Op}(\mathrm{Ad}(O^{-1}_{<h,\pm}))(D,y,s).$$

Given a compound $\mathrm{End}(\mathfrak{g})$-valued symbol $a(t,x,s,y,\xi)$, we define the double space-time frequency projection

$$(a)_{\ll k}(t,x,s,y,\xi) = S^{t,x}_{<k}S^{s,y}_{<k}a(t,x,s,y,\xi).$$

Therefore, according to our conventions,

$$\mathrm{Ad}(\boldsymbol{O}_{<h,\pm})_{\ll k}(t,x,s,y,\xi) = \mathrm{Ad}(O_{<h,\pm})_{<k}(t,x,\xi)\,\mathrm{Ad}(O^{-1}_{<h,\pm})_{<k}(s,y,\xi).$$

Proposition 10.14. *For $\delta > 0$ sufficiently small, there exists $\delta_{(1)}$ such that the following bound holds for any $h < -20$:*

$$\big\|\big(\mathrm{Op}(\mathrm{Ad}(\boldsymbol{O}_{<h,\pm})_{\ll 0})(t,x,D,t,y) - 1\big)P_0\big\|_{N^*\to X^{0,1/2}_\infty} \lesssim_{M_\sigma} 2^{\delta_{(1)}h}. \tag{10-35}$$

Before we begin the proof, we state a lemma for passing to a double space-time frequency localization of $\mathrm{Ad}(\boldsymbol{O}_{<h,\pm})$, which is used several times in our argument below.

Lemma 10.15. *For $2 \le q \le \infty$ and $h + 10 \le k \le 0$, we have*

$$\big\|\big(\mathrm{Op}(\mathrm{Ad}(\boldsymbol{O}_{<h,\pm})_{\ll 0}) - \mathrm{Op}(\mathrm{Ad}(\boldsymbol{O}_{<h,\pm})_{\ll k})\big)P_0\big\|_{L^pL^2\to L^qL^2} \lesssim_{M_\sigma} 2^{(\frac{1}{p}-\frac{1}{q})h}2^{10(h-k)}. \tag{10-36}$$

This lemma is a straightforward consequence of Lemma 10.12; we omit the proof.

Proof of (10-35). We follow [Oh and Tataru 2018, Proof of Proposition 9.13]. For simplicity, we omit $\pm$ in $O_{<h,\pm}$, $\boldsymbol{O}_{<h,\pm}$ etc.

Step 1: high-modulation input. For any $j \in \mathbb{Z}$ and $j' \ge j - 5$, we claim that

$$\|Q_j(\mathrm{Op}(\mathrm{Ad}(\boldsymbol{O}_{<h})_{\ll 0}) - 1)P_0Q_{j'}\|_{N^*\to X^{0,1/2}_\infty} \lesssim_{M_\sigma} 2^{\delta_{(0)}h}2^{\frac{1}{2}(j-j')}. \tag{10-37}$$

Step 2: low modulation input, $\frac{1}{2}h \le j$. Here, we take care of the easy case $\frac{1}{2}h \le j$. Under this assumption, we claim that

$$\| Q_j (\mathrm{Op}(\mathrm{Ad}(\boldsymbol{O}_{<h})_{\ll 0}) - 1) P_0 Q_{<j-5} \|_{N^* \to X_\infty^{0,1/2}} \lesssim_{M_\sigma} 2^{4h}. \tag{10-38}$$

Note that

$$Q_j (\mathrm{Op}(\mathrm{Ad}(\boldsymbol{O}_{<h})_{\ll j-5}) - 1) P_0 Q_{<j-5} = 0.$$

Thus, using the $L^\infty L^2$ portion of N^*, it suffices to prove

$$\| Q_j (\mathrm{Op}(\mathrm{Ad}(\boldsymbol{O}_{<h})_{\ll 0} - \mathrm{Ad}(\boldsymbol{O}_{<h})_{\ll j-5}) P_0 Q_{<j-5} \|_{N^* \to X_\infty^{0,1/2}} \lesssim_{M_\sigma} 2^{4h}.$$

Since Q_j and $Q_{<j-5}$ are disposable in $L^2 L^2$ and $L^\infty L^2$, respectively, this estimate follows from Lemma 10.15.

Step 3: low modulation input, $j < \frac{1}{2}h$, main decomposition. The goal of Steps 3–6 is to establish

$$\| Q_j (\mathrm{Op}(\mathrm{Ad}(\boldsymbol{O}_{<h})_{\ll 0}) - \mathrm{Ad}(\boldsymbol{O}_{<j+\tilde{\delta}h})_{\ll 0}) P_0 Q_{<j-5} \|_{N^* \to X_\infty^{0,1/2}} \lesssim_{M_\sigma} 2^{\delta_{(0)} h}, \tag{10-39}$$

provided that $j + \tilde{\delta}h \le h$.

At the level of $\mathrm{End}(\mathfrak{g})$-valued compound symbols, we expand

$$\mathrm{Ad}(\boldsymbol{O}_{<h}) - \mathrm{Ad}(\boldsymbol{O}_{<j+\tilde{\delta}h}) = \mathcal{L} + \mathcal{Q} + \mathcal{C},$$

where

$$\begin{aligned}
\mathcal{L} &= \int_{j+\tilde{\delta}h \le \ell \le h} \mathcal{L}_{\ell,<j+\tilde{\delta}h} \, d\ell, \\
\mathcal{Q} &= \int_{j+\tilde{\delta}h \le \ell' \le \ell \le h} \mathcal{Q}_{\ell,\ell',<j+\tilde{\delta}h} \, d\ell' \, d\ell, \\
\mathcal{C} &= \int_{j+\tilde{\delta}h \le \ell'' \le \ell' \le \ell \le h} \mathcal{C}_{\ell,\ell',\ell'',<\ell''} \, d\ell'' \, d\ell' \, d\ell,
\end{aligned}$$

and the integrands $\mathcal{L}_{\ell,<k}$, $\mathcal{Q}_{\ell,\ell',<k}$ and $\mathcal{C}_{\ell,\ell',\ell'',<k}$ are defined recursively as

$$\begin{aligned}
\mathcal{L}_{\ell,<k}(t,x,s,y,\xi) &= \mathrm{ad}(\Psi_\ell)(t,x,\xi)\,\mathrm{Ad}(\boldsymbol{O}_{<k})(t,x,s,y,\xi) - \mathrm{Ad}(\boldsymbol{O}_{<k})(t,x,s,y,\xi)\,\mathrm{ad}(\Psi_\ell)(s,y,\xi), \\
\mathcal{Q}_{\ell,\ell',<k}(t,x,s,y,\xi) &= \mathrm{ad}(\Psi_\ell)(t,x,\xi)\mathcal{L}_{\ell',<k}(t,x,s,y,\xi) - \mathcal{L}_{\ell',<k}(t,x,s,y,\xi)\,\mathrm{ad}(\Psi_\ell)(s,y,\xi), \\
\mathcal{C}_{\ell,\ell',\ell'',<k}(t,x,s,y,\xi) &= \mathrm{ad}(\Psi_\ell)(t,x,\xi)\mathcal{Q}_{\ell',\ell'',<k}(t,x,s,y,\xi) - \mathcal{Q}_{\ell',\ell'',<k}(t,x,s,y,\xi)\,\mathrm{ad}(\Psi_\ell)(s,y,\xi).
\end{aligned}$$

The three terms $\mathcal{L}_{\ell,<k}$, $\mathcal{Q}_{\ell,\ell',<k}$ and $\mathcal{C}_{\ell,\ell',\ell'',<k}$ are successively considered in the next three steps.

Step 4: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{L}$. Our goal here is to prove

$$\| Q_j \mathcal{L}_{\ll 0} P_0 Q_{<j-5} \|_{N^* \to X_\infty^{0,1/2}} \lesssim_{M_\sigma} 2^{\delta_{(0)} h}. \tag{10-40}$$

We introduce

$$\begin{aligned}
\mathcal{L}_{\ell,<k,\ll k'} &= \mathrm{ad}(\Psi_\ell)(t,x,\xi)\,\mathrm{Ad}(\boldsymbol{O}_{<k})_{\ll k'}(t,x,s,y,\xi), - \mathrm{Ad}(\boldsymbol{O}_{<k})_{\ll k'}(t,x,s,y,\xi)\,\mathrm{ad}(\Psi_\ell)(s,y,\xi) \\
\mathcal{L}_{\ell,<-\infty} &= \mathrm{ad}(\Psi_\ell)(t,x,\xi) - \mathrm{ad}(\Psi_\ell)(s,y,\xi)
\end{aligned}$$

and take the decomposition

$$\mathcal{L} = \int_{j+\tilde{\delta}h\le\ell\le h} (\mathcal{L}_{\ell,<j+\tilde{\delta}h} - \mathcal{L}_{\ell,<j+\tilde{\delta}h,\ll j-5})\,d\ell + \int_{j-10\tilde{\delta}h\le\ell\le h} \mathcal{L}_{\ell,<j+\tilde{\delta}h,\ll j-5}\,d\ell$$
$$+ \int_{j+\tilde{\delta}h\le\ell\le j-10\tilde{\delta}h} (\mathcal{L}_{\ell,<j+\tilde{\delta}h,\ll j-5} - \mathcal{L}_{\ell,<-\infty})\,d\ell + \int_{j+\tilde{\delta}h\le\ell\le j-10\tilde{\delta}h} \mathcal{L}_{\ell,<-\infty}\,d\ell$$
$$=: \mathcal{L}_{(1)} + \mathcal{L}_{(2)} + \mathcal{L}_{(3)} + \mathcal{L}_{(4)}.$$

Step 4.1: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{L}_{(1)}$. For this term we can add a double frequency localization $\ll C$ on $\mathcal{L}_{\ell,<j+\tilde{\delta}h}$ and then harmlessly discard the double $\ll 0$ localization in (10-40). Then it suffices to prove that for $\ell > j + \delta m$ we have

$$\| Q_j \operatorname{Op}(\mathcal{L}_{\ell,<j+\tilde{\delta}h,\ll C} - \mathcal{L}_{\ell,<j+\tilde{\delta}h,\ll j-5}) P_0 Q_{<j-5}\|_{L^\infty L^2\to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{6}[\ell-(j+\tilde{\delta}h)]} 2^{(10+\frac{1}{2})\tilde{\delta}h},$$

and then integrate with respect to ℓ. But this is a consequence of the decomposability bound (10-14) with $q = 6$ and $r = \infty$, together with the bound (10-34) with $p = 6$ and $q = 2$.

Step 4.2: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{L}_{(2)}$. Here as well as in the next two cases the $\ll 0$ localization in ℓ has no effect and is discarded. The two terms in $\mathcal{L}_{\ell,<j+\tilde{\delta}h,\ll j-5}$ are similar; we restrict our attention to the first one. Consider now the operator

$$Q_j \operatorname{Op}(\operatorname{ad}(\Psi_\ell) \operatorname{Ad}(\boldsymbol{O}_{<j+\tilde{\delta}h})_{\ll j-5}) Q_{<j-5} = \sum_\theta Q_j \operatorname{Op}(\operatorname{ad}(\Psi_\ell^{(\theta)}) \operatorname{Ad}(\boldsymbol{O}_{<j+\tilde{\delta}h})_{\ll j-5}) Q_{<j-5}.$$

The important observation here is that, because of the geometry of the cone, the frequency localizations for both $\operatorname{Ad}(\boldsymbol{O}_{<j+tdh})_{\ll j-5})$ and $\Psi_\ell^{(\theta)}$ force a large angle $\theta > 2^{(j-\ell)/2}$, or else the above operator vanishes.

Given this bound for θ, we can now use the decomposability bound (10-14) with $q = 2$ and $r = \infty$ combined with (10-34) with $p = \infty$ and $q = \infty$ to obtain

$$\| \operatorname{Op}(\operatorname{ad}(\Psi_\ell^{(\theta)}) \operatorname{Ad}(\boldsymbol{O}_{<j+\tilde{\delta}h})_{\ll j-5}) P_0\|_{L^\infty L^2\to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j} 2^{\frac{1}{2}(j-\ell)} \theta^{-\frac{1}{2}},$$

which after θ summation in the range $\theta > 2^{\frac{1}{2}(j-\ell)}$ yields

$$\| Q_j \operatorname{Op}(\mathcal{L}_{(2)}) P_0 Q_{<j-5}\|_{L^\infty L^2\to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j} 2^{\frac{5}{2}\tilde{\delta}h},$$

which suffices.

Step 4.3: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{L}_{(3)}$. Here we have the same angle constraint as above but this levels off for $\ell < j$, namely $\theta > 2^{-(\ell-j)_+/2}$. However, we can now replace (10-32) with (10-27) to obtain

$$\| \operatorname{Op}(\operatorname{ad}(\Psi_\ell^{(\theta)})(\operatorname{Ad}(\boldsymbol{O}_{<j+\tilde{\delta}h})_{\ll j-5} - I)) P_0\|_{L^\infty L^2\to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j} 2^{-\frac{1}{2}(\ell-j)} \theta^{-\frac{1}{2}} (2^{\delta_{(0)}(j+\tilde{\delta}h)} + 2^{10\tilde{\delta}h}),$$

which after θ and ℓ summation yields

$$\| Q_j \operatorname{Op}(\mathcal{L}_{(3)}) P_0 Q_{<j-5}\|_{L^\infty L^2\to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j} (2^{(\delta_{(0)}-\frac{1}{4}\tilde{\delta})h} + 2^{9\tilde{\delta}h}).$$

This suffices provided that $\tilde{\delta}$ is small enough $\tilde{\delta} < \delta_{(0)}$.

Step 4.4: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{L}_{(4)}$. Here we have the same range $j - \tilde{\delta}h < \ell < j + 10\tilde{\delta}h$ for ℓ. We also have the same constraint on the angle $\theta > 2^{-(\ell-j)_+/2}$ but this is no longer relevant in this case, as we will gain in frequency, and this can override any angular losses.

This time we are able to take advantage of the difference structure for Ψ. Precisely, it suffices to show that for a, a localized at frequency 1, we have

$$\| \operatorname{Op}(\operatorname{ad}(\Psi_\ell^{(\theta)}))(t,x,D)a(D) - a(D)\operatorname{Op}(\operatorname{ad}(\Psi_\ell^{(\theta)}))(t,x,D)\|_{L^\infty L^2 \to L^q L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{q}\ell} 2^{\ell} \theta^{-C}. \quad (10\text{-}41)$$

But this was already proved in [Oh and Tataru 2018, (9.40)].

Step 5: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{Q}$. We proceed in the same manner as in the case of $\mathcal{L}$. Defining the symbols

$$\mathcal{Q}_{\ell,\ell',<k,\ll k'} = \operatorname{ad}(\Psi_\ell)(t,x,\xi)\mathcal{L}_{\ell',<k,\ll k'}(t,x,s,y,\xi) - \mathcal{L}_{\ell',<k,\ll k'}(t,x,s,y,\xi)\operatorname{ad}(\Psi_\ell)(s,y,\xi),$$
$$\mathcal{Q}_{\ell,\ell',<-\infty} = \operatorname{ad}(\Psi_\ell)(t,x,\xi)\mathcal{L}_{\ell',<-\infty}(t,x,s,y,\xi) - \mathcal{L}_{\ell',<-\infty}(t,x,s,y,\xi)\operatorname{ad}(\Psi_\ell)(s,y,\xi),$$

we decompose $\mathcal{Q}$ as

$$\begin{aligned}\mathcal{Q} = &\int_{j+\tilde{\delta}h \le \ell' \le \ell \le h} (\mathcal{Q}_{\ell,\ell',<j+\tilde{\delta}h} - \mathcal{Q}_{\ell,\ell',<j+\tilde{\delta}h,\ll j-10})\, d\ell'\, d\ell \\ &+ \int_{\substack{j+\tilde{\delta}h \le \ell' \le \ell \le h \\ j-10\tilde{\delta}h \le \ell}} \mathcal{Q}_{\ell,\ell',<j+\tilde{\delta}h,\ll j-10}\, d\ell'\, d\ell \\ &+ \int_{j+\tilde{\delta}h \le \ell' \le \ell \le j-10\tilde{\delta}h} (\mathcal{Q}_{\ell,\ell',<j+\tilde{\delta}h,\ll j-10} - \mathcal{Q}_{\ell,\ell',<-\infty})\, d\ell'\, d\ell \\ &+ \int_{j+\tilde{\delta}h \le \ell' \le \ell \le j-10\tilde{\delta}h} \mathcal{Q}_{\ell,\ell',<-\infty}\, d\ell'\, d\ell \\ =:&\ \mathcal{Q}_{(1)} + \mathcal{Q}_{(2)} + \mathcal{Q}_{(3)} + \mathcal{Q}_{(4)}\end{aligned}$$

Then we consider each term separately.

Step 5.1: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{Q}_{(1)}$. Proceeding as in Step 4.1, we have

$$\mathcal{Q}_{\ll 1} = \int_{j+\tilde{\delta}h \le \ell' \le \ell \le h} (\mathcal{Q}_{\ell,\ell',<j+\tilde{\delta}h,\ll C} - \mathcal{Q}_{\ell,\ell',<j+\tilde{\delta}h,\ll j-5})_{\ll 0}\, d\ell'\, d\ell$$

and we can again harmlessly discard the outer $\ll 0$. Applying the decomposability bound (10-14) with $q = 6$ for Ψ_ℓ and with $q = \infty$ for $\Psi_{\ell'}$ and $r = \infty$, together with the bound (10-34) with $p = \infty$ and $q = 3$, we obtain

$$\|\mathcal{Q}_{\ell,\ell',<j+\tilde{\delta}h,\ll C} - \mathcal{Q}_{\ell,\ell',<j+\tilde{\delta}h,\ll j-5}\|_{L^\infty L^2 \to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{6}[\ell-(j+\tilde{\delta}h)]} 2^{(10+\frac{1}{2})\tilde{\delta}h}.$$

Summing up with respect to ℓ and ℓ' we obtain

$$\|\operatorname{Op}(\mathcal{Q}_{(1)})P_0\|_{L^\infty L^2 \to L^2} \lesssim_{M_\sigma} 2^{10\tilde{\delta}h},$$

which suffices.

Step 5.2: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{Q}_{(2)}$. Here and also for $\mathcal{Q}_{(3)}$ and $\mathcal{Q}_{(4)}$ we can remove the outer frequency localization $\ll 0$, which does nothing. The expression $\mathcal{Q}_{(2)}$ contains four terms depending on whether Ψ_ℓ and $\Psi_{\ell'}$ act on the left or on the right. We consider one of them, for which we need to bound the operator

$$Q_j \operatorname{Op}(\operatorname{ad}(\Psi_\ell)\operatorname{Ad}(\mathcal{O}_{<j+\tilde{\delta}h})_{\ll j-5}\operatorname{ad}(\Psi_{\ell'}))Q_{<j-5}P_0.$$

We decompose with respect to angles into

$$\sum_{\theta,\theta'} Q_j \operatorname{Op}(\operatorname{ad}(\Psi_\ell^{(\theta)})\operatorname{Ad}(\mathcal{O}_{<j+\tilde{\delta}h})_{\ll j-5}\operatorname{ad}(\Psi_{\ell'}^{(\theta')}))Q_{<j-5}P_0$$

and consider the nontrivial scenarios. This is as in Step 5.2 but now we have two angles, which must satisfy nonexclusively

$$\text{either } \theta > 2^{\frac{1}{2}(j-\ell)}, \qquad \text{or } \theta' > 2^{\frac{1}{2}(j-\ell')}.$$

We can now use the decomposability bound (10-14) with $q = 3$ and $r = \infty$ for the large[11] angle and $q = 6$ and $r = \infty$ for the other angle combined with (10-34) with $p = \infty$ and $q = \infty$ to obtain either

$$\|\operatorname{Op}(\operatorname{ad}(\Psi_\ell^{(\theta)})\operatorname{Ad}(\mathcal{O}_{<j+\tilde{\delta}h})_{\ll j-5}\operatorname{ad}(\Psi_{\ell'}^{(\theta')}))P_0\|_{L^\infty L^2\to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j}2^{\frac{1}{3}(j-\ell)}\theta^{-\frac{1}{6}}2^{\frac{1}{6}(j-\ell')}\theta'^{\frac{1}{6}}$$

or the same bound with the pairs (l,θ) and (l',θ') reversed. Summing with respect to ℓ, ℓ', and also with respect to θ, θ' subject to the constraints above, we obtain

$$\|Q_j \operatorname{Op}(\mathcal{Q}_{(2)})P_0 Q_{<j-5}\|_{L^\infty L^2\to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j}2^{\frac{5}{3}\tilde{\delta}h},$$

which suffices.

Step 5.3: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{Q}_{(3)}$. We repeat the angle localization analysis in the previous step, but as in Step 4.3, we again replace (10-32) with (10-27). The outcome is similar to the one in Step 4.3; details are omitted.

Step 5.4: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{Q}_{(4)}$. Again we apply the same angle localization analysis as in the previous two steps. However, as in Step 4.4, we also need to exploit the difference between one of the two Ψ's and its adjoint. Consider one such term, e.g.,

$$\operatorname{ad}(\Psi_\ell^{(\theta)})(t,x,\xi)[\operatorname{ad}(\Psi_{\ell'}^{(\theta')})(t,x,\xi) - \operatorname{ad}(\Psi_{\ell'}^{(\theta')})(\xi,y,s)].$$

For this it suffices to apply the disposability bound (10-14) for $\Psi_\ell^{(\theta)}$ combined with (10-41). The choice of the exponents is no longer important. We obtain

$$\|\operatorname{Op}(\mathcal{Q}_{(4)})P_0\|_{L^\infty L^2\to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j}2^{(1-C\tilde{\delta})j}.$$

Step 6: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{C}$. This repeats the analysis for $\mathcal{L}$ and $\mathcal{Q}$, but we no longer need to keep track of angular separation. Setting

$$\mathcal{C}_{\ell,\ell',\ell'',<k,\ll k'} = \operatorname{ad}(\Psi_\ell)(t,x,\xi)\mathcal{Q}_{\ell',\ell'',<k,\ll k'}(t,x,s,y,\xi) - \mathcal{Q}_{\ell',\ell'',<k,\ll k'}(t,x,s,y,\xi)\operatorname{ad}(\Psi_\ell)(s,y,\xi),$$
$$\mathcal{C}_{\ell,\ell',\ell'',<-\infty} = \operatorname{ad}(\Psi_\ell)(t,x,\xi)\mathcal{Q}_{\ell',\ell'',<-\infty}(t,x,s,y,\xi) - \mathcal{Q}_{\ell',\ell'',<-\infty}(t,x,s,y,\xi)\operatorname{ad}(\Psi_\ell)(s,y,\xi),$$

[11] That is, which satisfies the bound on the previous line.

we decompose $\mathcal{C}$ as

$$\begin{aligned}\mathcal{C} = &\int_{j+\tilde{\delta}h \le \ell'' \le \ell' \le \ell \le h} (\mathcal{C}_{\ell,\ell',\ell'',<\ell''} - \mathcal{C}_{\ell,\ell',\ell'',<\ell'',\ll -5})\, d\ell''\, d\ell'\, d\ell \\ &+ \int_{\substack{j+\tilde{\delta}h \le \ell'' \le \ell' \le \ell \le h \\ j-10\tilde{\delta}h \le \ell}} \mathcal{C}_{\ell,\ell',\ell'',<\ell'',\ll -5}\, d\ell''\, d\ell'\, d\ell \\ &+ \int_{j+\tilde{\delta}h \le \ell'' \le \ell' \le \ell \le j-10\tilde{\delta}h} (\mathcal{C}_{\ell,\ell',\ell'',<\ell'',\ll j-5} - \mathcal{C}_{\ell,\ell',\ell'',<-\infty})\, d\ell''\, d\ell'\, d\ell \\ &+ \int_{j+\tilde{\delta}h \le \ell'' \le \ell' \le \ell \le j-10\tilde{\delta}h} \mathcal{C}_{\ell,\ell',\ell'',<-\infty}\, d\ell''\, d\ell'\, d\ell \\ =:&\, \mathcal{C}_{(1)} + \mathcal{C}_{(2)} + \mathcal{C}_{(3)} + \mathcal{C}_{(4)}\end{aligned}$$

and consider each of the terms separately.

Step 6.1: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{C}_{(1)}$. The same argument as in Steps 4.1 and 5.1 yields the bound

$$\begin{aligned}&\| \operatorname{Op}(\operatorname{ad}(\Psi_\ell)\operatorname{ad}(\Psi_{\ell'})\operatorname{ad}(\Psi_{\ell''})(\operatorname{Ad}((\boldsymbol{O}_{<\ell''}) - \operatorname{Ad}((\boldsymbol{O}_{<\ell''})_{\ll -5}))_{\ll 0} \|_{L^\infty L^2 \to L^2} \\ &\qquad \lesssim_{M_\sigma} 2^{-\frac{1}{2}j} 2^{\frac{1}{6}(j+\tilde{\delta}h-\ell)} 2^{\frac{1}{6}(j+\tilde{\delta}h-\ell')} 2^{\frac{1}{6}(j+\tilde{\delta}h-\ell')} 2^{10\ell''} 2^{\frac{1}{2}\tilde{\delta}h},\end{aligned}$$

as well as for any of the other choices of left/right quantizations for the Ψ's. Integration over $j+\tilde{\delta}h < \ell'' < \ell' < \ell < \frac{m}{2}$ is now harmless.

Step 6.2: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{C}_{(2)}$. Applying the decomposability bound (10-14) with $q = 6$ for each of the three Ψ's in the $\mathcal{C}_2$ integrand, as well as the L^2 bound for $\operatorname{Op}(\operatorname{Ad}((\boldsymbol{O}_{<\ell''})_{\ll -5})$ yields the bound

$$\| \operatorname{Op}(\operatorname{ad}(\Psi_\ell)\operatorname{ad}(\Psi_{\ell'})\operatorname{ad}(\Psi_{\ell''})\operatorname{Ad}((\boldsymbol{O}_{<j+\tilde{\delta}h}))_{\ll -5} \|_{L^\infty L^2 \to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j} 2^{\frac{1}{6}(j-\ell)} 2^{\frac{1}{6}(j-\ell')} 2^{\frac{1}{6}(j-\ell')},$$

which suffices after integration in $\ell > j - 10\tilde{\delta}h$ and $\ell', \ell'' > j + \tilde{\delta}h$.

Step 6.3: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{C}_{(3)}$. This is the same argument as in the previous step, but using (10-27) instead of (10-32).

Step 6.4: low modulation input, $j < \frac{1}{2}h$, contribution of $\mathcal{C}_{(4)}$. Here we are concerned with symbols of the form

$$\operatorname{ad}(\Psi_\ell)(t,x,\xi)\operatorname{ad}(\Psi_{\ell'})(t,x,\xi)[\operatorname{ad}(\Psi_{\ell''}(t,x,\xi) - \operatorname{ad}(\Psi_{\ell''}(\xi,y,s)],$$

where one or both of $\operatorname{ad}(\Psi_\ell)$ and $\operatorname{ad}(\Psi_{\ell'})$ may be switched to the right and in the right quantization. Here we use again the decomposability bound (10-14) with $q = 6$ for Ψ_ℓ and $\operatorname{ad}(\Psi_{\ell'})$, and (10-41) for the $\Psi_{\ell''}$ difference.

Step 7: low modulation input, $j < \frac{1}{2}h$, low frequency $\boldsymbol{O}$. To complete the proof of the estimate (10-35) it remains to show that

$$\| Q_j \operatorname{Op}(\operatorname{Ad}(\boldsymbol{O}_{<j+\tilde{\delta}h})_{\ll 0}(t,x,D,y,s) - 1) P_0 Q_{<j-5} \|_{N^* \to X^{0,1/2}_\infty} \lesssim_{M_\sigma} 2^{\delta_{(1)}h}. \tag{10-42}$$

If $j+\tilde{\delta}h \le h$, this is combined with the bound (10-39), which is the main outcome of Steps 3–6. Else, this is used by itself, simply observing that we can harmlessly replace $j+\tilde{\delta}h$ by h.

The above bound is identical to

$$\|Q_j \operatorname{Op}(\operatorname{Ad}(\boldsymbol{O}_{<j+\tilde{\delta}h})_{\ll 0} - \operatorname{Ad}(\boldsymbol{O}_{<j+\tilde{\delta}h})_{\ll j-5})(t,x,D,y,s) P_0 Q_{<j-5}\|_{N^* \to X^{0,1/2}_\infty} \lesssim_{M_\sigma} 2^{\delta_{(1)} h},$$

which in turn would follow from

$$\|\operatorname{Op}(\operatorname{Ad}(\boldsymbol{O}_{<j+\tilde{\delta}h})_{\ll 0} - \operatorname{Ad}(\boldsymbol{O}_{<j+\tilde{\delta}h})_{\ll j-5})(t,x,D,y,s)) P_0\|_{L^\infty L^2 \to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j} 2^{\delta_{(1)} h}.$$

But this is a direct consequence of the bound (10-34). □

Proof of (9-47) *in the case $Z = N$ or N^*.* For the estimate (9-47) with $Z = N^*$ we combine the $L^\infty L^2$ bound given by (10-27) with (10-35). If on the other hand $Z = N$, then the same bound follows by duality. □

It remains to prove (9-44), (9-46) and (9-44)′ when $Z = N$ or N^*. For this purpose, we recall the following result from [Krieger and Tataru 2017]:

Lemma 10.16. *For $\ell \le k' \pm O(1)$, we have*

$$\|Q_\ell \operatorname{Op}(\operatorname{Ad}(O_{<h,\pm})_{k'})(t,x,D) Q_{<0} P_0\|_{N^* \to X^{0,1/2}_\infty} \lesssim_{M_\sigma} 2^{\delta_1(\ell - k')}, \tag{10-43}$$

$$\|Q_\ell \operatorname{Op}(\operatorname{Ad}(O^{-1}_{<h,\pm})_{k'})(D,y,s) Q_{<0} P_0\|_{N^* \to X^{0,1/2}_\infty} \lesssim_{M_\sigma} 2^{\delta_1(\ell - k')}. \tag{10-44}$$

In particular, summing over all (ℓ, k') with $\ell \le k$ and $k \le k' + O(1)$, we have

$$\|Q_{<k}(\operatorname{Op}(\operatorname{Ad}(O_{<h,\pm})_{<0}) - \operatorname{Op}(\operatorname{Ad}(O_{<h,\pm})_{<k-C}))(t,x,D) Q_{<0} P_0\|_{N^* \to X^{0,1/2}_\infty} \lesssim_{M_\sigma} 1, \tag{10-45}$$

$$\|Q_{<k}(\operatorname{Op}(\operatorname{Ad}(O^{-1}_{<h,\pm})_{<0}) - \operatorname{Op}(\operatorname{Ad}(O^{-1}_{<h,\pm})_{<k-C}))(D,y,s) Q_{<0} P_0\|_{N^* \to X^{0,1/2}_\infty} \lesssim_{M_\sigma} 1. \tag{10-46}$$

Proof. The proof of this lemma is similar to that of Proposition 10.14, but simpler in the sense the frequency gap need not be exploited. It can be proved with exactly the same arguments as in [Krieger and Tataru 2017, Proof of Proposition 8.5] (there, $M_\sigma \lesssim \varepsilon$). Because of this, we will merely indicate here how to modify the preceding proof of (10-35) to obtain (10-43). We leave the details, as well as the entire case of (10-44), to the reader.

As before, we omit $\pm$ in the symbols. We replace $\operatorname{Ad}(\boldsymbol{O}_{<h})_{\ll k}(t,x,s,y,\xi) - 1$ by $\operatorname{Ad}(O_{<h})_{<k}(t,x,\xi)$ throughout the proof of (10-35). The main decomposition (Step 4) now takes the form

$$\begin{aligned}
\operatorname{Ad}(O_{<h})(t,x,\xi) - \operatorname{Ad}(O_{<j+\tilde{\delta}h}) &= \mathcal{L}' + \mathcal{Q}' + \mathcal{C}' \\
&= \int_{j+\tilde{\delta}h \le \ell \le h} \mathcal{L}'_{\ell,<j+\tilde{\delta}h}\, d\ell + \int_{j+\tilde{\delta}h \le \ell' \le \ell \le h} \mathcal{Q}'_{\ell,\ell',<j+\tilde{\delta}h}\, d\ell'\, d\ell \\
&\quad + \int_{j+\tilde{\delta}h \le \ell'' \le \ell' \le \ell \le h} \mathcal{C}'_{\ell,\ell',\ell'',<j+\tilde{\delta}h}\, d\ell''\, d\ell'\, d\ell,
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{L}'_{\ell,<k}(t,x,\xi) &= \operatorname{ad}(\Psi_\ell) \operatorname{Ad}(O_{<k})(t,x,\xi), \\
\mathcal{Q}'_{\ell,<k}(t,x,\xi) &= \operatorname{ad}(\Psi_\ell) \mathcal{L}'_{\ell',<k}(t,x,\xi) = \operatorname{ad}(\Psi_\ell) \operatorname{ad}(\Psi_{\ell'}) \operatorname{Ad}(O_{<k})(t,x,\xi), \\
\mathcal{C}'_{\ell,<k}(t,x,\xi) &= \operatorname{ad}(\Psi_\ell) \mathcal{Q}'_{\ell',\ell'',<k}(t,x,\xi) = \operatorname{ad}(\Psi_\ell) \operatorname{ad}(\Psi_{\ell'}) \operatorname{ad}(\Psi_{\ell''}) \operatorname{Ad}(O_{<k})(t,x,\xi).
\end{aligned}$$

For the expansion of $\mathcal{L}$, $\mathcal{Q}$ and $\mathcal{C}$ in Steps 5, 6 and 7, we replace $\mathcal{L}_{\ell,<k,\ll k'}$, $\mathcal{L}_{\ell,<-\infty}$, $\mathcal{Q}_{\ell,\ell',<k,\ll k'}$, $\mathcal{Q}_{\ell,\ell',<-\infty}$, $\mathcal{C}_{\ell,\ell',\ell'',<k,\ll k'}$ and $\mathcal{C}_{\ell,\ell',\ell'',<-\infty}$ by, respectively,

$$\begin{aligned}
\mathcal{L}'_{\ell,<k,<k'} &= \mathrm{ad}(\Psi_\ell)\,\mathrm{Ad}(O_{<k})_{<k'}(t,x,\xi),\\
\mathcal{L}'_{\ell,<-\infty} &= \mathrm{ad}(\Psi_\ell)(t,x,\xi),\\
\mathcal{Q}'_{\ell,\ell',<k,<k'} &= \mathrm{ad}(\Psi_\ell)\mathcal{L}'_{\ell',<k,<k'}(t,x,\xi) = \mathrm{ad}(\Psi_\ell)\,\mathrm{ad}(\Psi_{\ell'})\,\mathrm{Ad}(O_{<k})_{<k'}(t,x,\xi),\\
\mathcal{Q}'_{\ell,\ell',<-\infty} &= \mathrm{ad}(\Psi_\ell)\mathcal{L}'_{\ell',<-\infty}(t,x,\xi) = \mathrm{ad}(\Psi_\ell)\,\mathrm{ad}(\Psi_{\ell'})(t,x,\xi),\\
\mathcal{C}'_{\ell,\ell',\ell''<k,<k'} &= \mathrm{ad}(\Psi_\ell)\mathcal{Q}'_{\ell',\ell'',<k,<k'}(t,x,\xi) = \mathrm{ad}(\Psi_\ell)\,\mathrm{ad}(\Psi_{\ell'})\,\mathrm{ad}(\Psi_{\ell''})\,\mathrm{Ad}(O_{<k})_{<k'}(t,x,\xi),\\
\mathcal{C}'_{\ell,\ell',\ell'',<-\infty} &= \mathrm{ad}(\Psi_\ell)\mathcal{Q}'_{\ell',\ell'',<-\infty}(t,x,\xi) = \mathrm{ad}(\Psi_\ell)\,\mathrm{ad}(\Psi_{\ell'})\,\mathrm{ad}(\Psi_{\ell''})(t,x,\xi).
\end{aligned}$$

Accordingly, we replace the use of (10-27) and (10-36) by (10-32) and (10-34), respectively, which results in loss of the smallness factor $2^{\delta_{(1)}h}$ in (10-43) compared to (10-35). □

Proof of (9-44), (9-46) *and* (9-44)′ *in the case $Z = N$ or N^*.* It suffices to consider the $Z = N^*$; then the case $Z = N$ follows by duality. The $L^\infty L^2$ bound follows from the $Z = L^2$ case, so for (9-44) and (9-44)′ it remains to establish that

$$\|Q_j\,\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_{<0})P_0\|_{N^*\to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j}.$$

By Lemma 10.16 this reduces to

$$\|Q_j\,\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_{<j-5})P_0\|_{N^*\to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j}.$$

Now due to the frequency localization for $\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_{<j-5}$ we can insert a (slight enlargement of) Q_j on the right, in which case we can simply use again the $Z = L^2$ case.

Similarly, in the case of (9-44)′ it suffices to show that

$$\|Q_j[\partial_t, \mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_{<0})]Q_{<j}P_0\|_{N^*\to L^2} \lesssim_{M_\sigma} 2^{-\frac{1}{2}j}2^h.$$

We split into two cases. If $j \le \frac{3}{4}h$ then we write

$$\partial_t\,\mathrm{Ad}(O_{<h,\pm}) = \mathrm{ad}(O_{<h,\pm;t})\,\mathrm{Ad}(O_{<h,\pm})_{<0}),$$

and then we can easily combine the decomposability bound (10-18) with the L^2 boundedness of $\mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_{<0})$. Else we have

$$Q_j[\partial_t, \mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_{<0})]Q_{<j}P_0 = Q_j[\partial_t, \mathrm{Op}(\mathrm{Ad}(O_{<h,\pm})_{[j-5,0]})]Q_{<j}P_0.$$

Now we discard Q_j, $Q_{<j-5}$ and ∂_t and use directly (10-34) with $p = \infty$ and $q = 2$. □

10H. ***Dispersive estimates.*** Finally, we sketch the proofs of (9-45) and (9-45)′. As in [Krieger and Tataru 2017], we exactly follow the argument in [Krieger et al. 2015, Section 11]. In the case of (9-45), we replace the use of the oscillatory integral estimates (108), (110) and (111) in [loc. cit.] by (10-24), (10-25) and (10-26), and the fixed-time L^2 bound (114) in [loc. cit.] by (10-32), (118) in [loc. cit.] by (10-45) etc. In case of (9-45)′, observe that all the constants in these bounds are *universal* under the smallness assumption (9-48) for a suitable choice of $\delta_o(M)$, as we may take $M_\sigma \lesssim 1$.

There is one exception to the above strategy, namely the square function bound

$$\| \mathrm{Op}(\mathrm{Ad}(O_{\pm})_{<0}(t,x,D)\|_{S_0^{\sharp}\to L_x^{10/3}L_t^2} \lesssim_{M_\sigma} 1. \tag{10-47}$$

This is due to the fact that the square function norm was not part of the S_0 norm in [Krieger et al. 2015; Krieger and Tataru 2017], and was added only here. The same approach as in [Krieger and Tataru 2017] allows us, via a TT^*-type argument, to reduce the problem to an estimate of the form

$$\left\| \int \chi_{-l}(t-s)\boldsymbol{S}(t,s)B(s)\,ds \right\|_{L_x^{10/3}L_t^2} \lesssim_{M_\sigma} \|B\|_{L_x^{10/7}L_t^2},$$

where

$$\boldsymbol{S}(t,s) = \mathrm{Op}(\mathrm{Ad}(O_{\pm})_{<0}(t,x,D)e^{\pm i(t-s)|D|}\,\mathrm{Op}(\mathrm{Ad}(O_{\pm})_{<0}(D,s,y)$$

and the bump function χ_{-l} corresponds to the modulation scale 2^l in $S_0^{\sharp}$. It is easily seen that the bump function is disposable and can be harmlessly discarded. Hence in order to prove (10-47) it remains to show that

$$\left\| \int \boldsymbol{S}(t,s)B(s)\,ds \right\|_{L_x^{10/3}L_t^2} \lesssim_{M_\sigma} \|B\|_{L_x^{10/7}L_t^2}. \tag{10-48}$$

To prove this we use Stein's analytic interpolation theorem. We consider the analytic family of operators

$$T_z B(t) = e^{z^2}\int (t-s)^z \boldsymbol{S}(t,s)B(s)\,ds$$

for z in the strip

$$-1 \le \mathrm{Im} z \le \tfrac{3}{2}.$$

Then it suffices to establish the uniform bounds

$$\|T_z\|_{L^2\to L^2} \lesssim_{M_\sigma} 1, \quad \mathrm{Re} z = -1, \tag{10-49}$$

$$\|T_z\|_{L_x^1 L_t^2 \to L_x^\infty L_t^2} \lesssim_{M_\sigma} 1, \quad \mathrm{Re} z = \tfrac{3}{2}. \tag{10-50}$$

For (10-49) we can use the bound (10-31) to discard the L^2 bounded operators

$$\mathrm{Op}(\mathrm{Ad}(O_{\pm})_{<0}(t,x,D)e^{\pm it|D|}, \qquad e^{\mp is|D|}\,\mathrm{Op}(\mathrm{Ad}(O_{\pm})_{<0}(D,s,y).$$

Then we are left with the time convolutions with the kernels $e^{z^2}t^z$. But these are easily seen to be multipliers with uniformly bounded symbols.

For (10-50), on the other hand, we consider the kernel $K_z(t,x,s,y)$ of T_z. This is given by

$$K_z(t,x,s,y) = e^{z^2}(t-s)^z K^a_{<0}(t,x,s,y)$$

with a a smooth bump function on the unit scale. Hence by (10-24) we have the kernel bound

$$|K_z(t,x,s,y)| \lesssim_{M_\sigma} \langle |t-s| - |x-y| \rangle^{-100}, \quad \mathrm{Re} z = \tfrac{3}{2}.$$

Fixing x and y we have the obvious bound

$$\|K_z(\cdot,x,\cdot,y)\|_{L^2\to L^2} \lesssim_{M_\sigma} 1.$$

Then (10-50) easily follows.

11. Renormalization error bounds

Without loss of generality, we fix the sign $\pm = +$. In this section, unless we specify otherwise, $\mathrm{Op}(\cdot)$ denotes the left quantization. For the sake of simplicity, we also adopt the convention of simply writing A_x for $\boldsymbol{P}_x A$.

11A. *Preliminaries.* We collect here some technical tools for proving the renormalization error bound.

We begin with a tool that allows us to split $\mathrm{Op}(ab)$ into $\mathrm{Op}(a)\,\mathrm{Op}(b)$. The idea of the proof is based on the heuristic identity $\mathrm{Op}(ab) - \mathrm{Op}(a)\,\mathrm{Op}(b) \approx \mathrm{Op}(-i\,\partial_\xi a \cdot \partial_x b)$ for left-quantized pseudodifferential operators; see [Krieger et al. 2015, Lemma 7.2] and [Krieger and Tataru 2017, Lemma 7.2].

Lemma 11.1 (composition via pseudodifferential calculus). *Let $a(t,x,\xi)$ and $b(t,x,\xi)$ be $\mathrm{End}(\mathfrak{g})$-valued symbols on $I_t \times \mathbb{R}^4_x \times \mathbb{R}^4_\xi$ with bounded derivatives, such that $a(t,x,\xi)$ is homogeneous of degree 0 in ξ and $b(t,x,\xi) = P^x_{<h_\theta - 10} b(t,x,\xi)$ for some $0 < \theta < 1$ and $2^{h_\theta} = \theta$. Then we have*

$$\begin{aligned}\|(\mathrm{Op}(a)\,\mathrm{Op}(b) - \mathrm{Op}(ab))P_0\|_{L^q L^2[I] \to L^r L^2[I]} \\ \lesssim \|\theta\partial_\xi a\|_{D_\theta L^{p_2} L^\infty[I]} \|\mathrm{Op}(\theta^{-1}\partial_x b)P_0\|_{L^q L^2[I] \to L^{p_1} L^2[I]}, \qquad (11\text{-}1)\end{aligned}$$

where $r^{-1} = p_1^{-1} + p_2^{-1}$.

Proof. For simplicity, in this proof we only present formal computation, which can be justified using the qualitative assumptions on a and b.

Let us fix $t \in I$. Thanks to the frequency-localization condition $b(x,\xi) = P^x_{<h_\theta-10} b(x,\xi)$, we may write

$$(\mathrm{Op}(a)\,\mathrm{Op}(b) - \mathrm{Op}(ab))P_0 = \sum_\phi \mathrm{Op}(a^\phi_\theta)\,\mathrm{Op}(b^\phi_\theta) - \mathrm{Op}(a^\phi_\theta b^\phi_\theta),$$

where

$$a^\phi_\theta(x,\xi) = a(x,\xi)(m^\phi_\theta)^2(\xi)\tilde{m}^2_0(\xi), \quad b^\phi_\theta(x,\xi) = b(x,\xi)\tilde{m}^\phi_\theta(\xi) m_0(\xi).$$

Here ϕ runs over caps of radius $\simeq \theta$ on $\mathbb{S}^3$ with uniformly finite overlaps, $(m^\phi_\theta)^2(\xi) = (m^\phi_\theta)^2(\xi/|\xi|)$ are the associated smooth partition of unity on $\mathbb{S}^3$ and $m_0(\xi)$ is the symbol for P_0. The functions $\tilde{m}^\phi_\theta(\xi) = \tilde{m}^\phi_\theta(\xi/|\xi|)$ and $\tilde{m}^2_0(\xi)$ are smooth cutoffs to the supports of m^ϕ_θ and m_0, respectively, which can be inserted thanks to the frequency-localization condition $b(x,\xi) = P^x_{<h_\theta-10} b(x,\xi)$.

For each ϕ, we claim that

$$\begin{aligned}\|\mathrm{Op}(a^\phi_\theta)\,\mathrm{Op}(b^\phi_\theta) - \mathrm{Op}(a^\phi_\theta b^\phi_\theta)\|_{L^2 \to L^2} \\ \lesssim \left(\sum_{n=1}^{20} \sup_\omega m^\phi_\theta(\omega) \|\theta^n \partial^{(n)}_\xi a(\cdot,\omega)\|_{L^\infty}\right) \|\mathrm{Op}(\theta^{-1}\partial_x b^\phi_\theta)\|_{L^2 \to L^2}. \qquad (11\text{-}2)\end{aligned}$$

Assuming the claim, the proof can be completed as follows. Let us restore the dependence of the symbols on t. By the definition of $D_\theta L^q L^r$, we have

$$\left\| \left(\sum_\phi \left(\sum_{n=1}^{20} \sup_\omega m^\phi_\theta(\omega) \|\theta^n \partial^{(n)}_\xi a(t,\cdot,\omega)\|_{L^\infty} \right)^2 \right)^{\frac{1}{2}} \right\|_{L^{p_2}_t[I]} \lesssim \|\theta\partial_\xi a\|_{D_\theta L^{p_2} L^\infty[I]}.$$

On the other hand, by L^2-almost orthogonality of $\tilde{m}^{\phi}_{\theta}(\xi)$ and Hölder in t, we have

$$\Big\| \Big(\sum_{\phi} \|\mathrm{Op}(\theta^{-1}\partial_x b^{\phi}_{\theta})\|^2_{L^2\to L^2} \Big)^{\frac{1}{2}} \Big\|_{L^{p_0}_t[I]} \lesssim \|\mathrm{Op}(\theta^{-1}\partial_x b) P_0\|_{L^q L^2 \to L^{p_1} L^2[I]},$$

where $r^{-1} + p_0^{-1} = p_1^{-1}$. Therefore, by Cauchy–Schwarz in ϕ and Hölder in t, (11-1) follows.

We now turn to the proof of (11-2). For simplicity of notation, we use the shorthand $a = a^{\phi}_{\theta}$ and $b = b^{\phi}_{\theta}$ for now. Then the kernel of $\mathrm{Op}(a)\,\mathrm{Op}(b) - \mathrm{Op}(ab)$ can be computed as follows:

$$\begin{aligned} K(x,y) &= \int e^{i(x-z)\cdot\xi} e^{i(z-y)\cdot\eta} (a(x,\xi) - a(x,\eta)) b(z,\eta)\, dz \frac{d\xi}{(2\pi)^4} \frac{d\eta}{(2\pi)^4} \\ &= \int_0^1 \!\! \int e^{i(x-z)\cdot\xi} e^{i(z-y)\cdot\eta} (\xi - \eta)\cdot(\partial_\xi a)(x, s\xi + (1-s)\eta) b(z,\eta)\, dz \frac{d\xi}{(2\pi)^4} \frac{d\eta}{(2\pi)^4}\, ds \\ &= -i \int_0^1 \!\! \int e^{i(x-z)\cdot\xi} e^{i(z-y)\cdot\eta} (\partial_\xi a)(x, s\xi + (1-s)\eta)(\partial_x b)(z,\eta)\, dz \frac{d\xi}{(2\pi)^4} \frac{d\eta}{(2\pi)^4}\, ds. \end{aligned}$$

Expanding

$$\partial_\xi a(x,\cdot) = \int e^{-i(\cdot)\cdot\Xi} (\partial_\xi a)^{\vee}(x,\Xi)\, d\Xi$$

and making the change of variables $\tilde{z} = z - (1-s)\Xi$, we further compute

$$\begin{aligned} K(x,y) &= -i \int_0^1 \!\! \int e^{i(x - s\Xi - z)\cdot\xi} e^{i(z-(1-s)\Xi - y)\cdot\eta} (\partial_\xi a)^{\vee}(x,\Xi)(\partial_x b)(z,\eta)\, d\Xi\, dz \frac{d\xi}{(2\pi)^4} \frac{d\eta}{(2\pi)^4}\, ds \\ &= -i \int_0^1 \!\! \int e^{i(x-\Xi-\tilde{z})\cdot\xi} e^{i(\tilde{z}-y)\cdot\eta} (\partial_\xi a)^{\vee}(x,\Xi)(\partial_x b)(\tilde{z} + (1-s)\Xi, \eta)\, d\Xi\, d\tilde{z} \frac{d\xi}{(2\pi)^4} \frac{d\eta}{(2\pi)^4}\, ds \\ &= -i \int_0^1 \!\! \int (\partial_\xi a)^{\vee}(x,\Xi) \Big(\int e^{i(x - s\Xi - y)\cdot\eta} (\partial_x b)(x - s\Xi, \eta) \frac{d\eta}{(2\pi)^4} \Big) d\Xi\, ds. \end{aligned}$$

On the last line, note that the η-integral inside the parentheses is precisely the kernel of $\mathrm{Op}(\partial_x b)(x - s\Xi, D)$. By translation invariance, we have

$$\theta^{-1} \|(\partial_x b)(x - s\Xi, D)\|_{L^2\to L^2} = \|(\theta^{-1}\partial_x b)(x, D) P_0\|_{L^2\to L^2}.$$

On the other hand, returning to the full notation $a^{\phi}_{\theta} = a$ and rotating the axes so that $\phi = (1,0,0,0)$, note that $a^{\phi}_{\theta}(x,\cdot)$ is supported on a rectangle of dimension $\simeq 1\times\theta\times\theta\times\theta$, and smooth on the corresponding scale. Integrating by parts in ξ to obtain rapid decay in Ξ (of the form $\langle \Xi^1\rangle^{-N} \langle \theta\Xi'\rangle^{-N}$, where $\Xi' = (\Xi^2, \Xi^3, \Xi^4)$), we may estimate

$$\begin{aligned} \theta \int \|(\partial_\xi a^{\phi}_{\theta})^{\vee}(\cdot,\Xi)\|_{L^\infty}\, d\Xi &\leq \int \Big\| \int e^{i\Xi\cdot\xi} \theta \partial_\xi a(\cdot,\xi) (m^{\phi}_{\theta})^2(\xi) \tilde{m}_0^2(\xi) \frac{d\xi}{(2\pi)^4} \Big\|_{L^\infty} d\Xi \\ &\lesssim \theta^{-3} \sum_{n=1}^{20} \int \|\theta^n \partial^{(n)}_{\xi} a(\cdot,\xi)\|_{L^\infty} m^{\phi}_{\theta}(\xi) \tilde{m}_0(\xi)\, d\xi. \end{aligned}$$

Passing to the polar coordinates $\xi = \lambda\omega$ (where $\lambda = |\xi|$), integrating out λ and using Hölder in ω (which cancels the factor θ^{-3}), we arrive at

$$\theta \int \|(\partial_\xi a^{\phi}_{\theta})^{\vee}(\cdot, \Xi)\|_{L^\infty}\, d\Xi \lesssim \sum_{n=1}^{20} \sup_{\omega} m^{\phi}_{\theta}(\omega)\|\theta^n \partial_\xi^{(n)} a(\cdot,\omega)\|_{L^\infty},$$

which proves (11-2). ☐

Remark 11.2. As is evident from the proof, we in fact have the simpler bound

$$\|(\mathrm{Op}(a)\,\mathrm{Op}(b) - \mathrm{Op}(ab))P_0\|_{L^q L^2[I]\to L^r L^2[I]} \lesssim \|a\|_{D_\theta L^{p_2} L^\infty[I]} \|\mathrm{Op}(\theta^{-1}\partial_x b)P_0\|_{L^q L^2[I]\to L^{p_1} L^2[I]}. \tag{11-1$'$}$$

In other words, control of the $D_\theta L^{p_2} L^\infty$-norm already encodes the fact that a is smooth in ξ on the scale θ.

In practice, Lemma 11.1 can be only be applied when we know that the symbol on the right (b in Lemma 11.1) is smooth in x on the scale θ^{-1}. Fortunately, when $b = \mathrm{Ad}(O)$, the remainder can be controlled using decomposability bounds for Ψ. We therefore have the following useful composition lemma.

Lemma 11.3 (composition lemma). *Let $G = G(t,x,\xi)$ be a smooth $\mathfrak{g}$-valued symbol on $I \times \mathbb{R}^4 \times \mathbb{R}^4$, which is homogeneous of degree 0 in ξ and admits a decomposition of the form $G = \sum_{\theta\in 2^{-\mathbb{N}}} G^{(\theta)}$, where*

$$\|G^{(\theta)}\|_{D_\theta L^2 L^\infty[I]} \le \theta^\alpha B$$

for some $B > 0$ and $\alpha > \frac{1}{2} + \delta$. Then for every $\ell \le 0$ we have

$$\|\mathrm{Op}(\mathrm{ad}(G)\,\mathrm{Ad}(O_{<\ell}))P_0 - \mathrm{Op}(\mathrm{ad}(G))\,\mathrm{Op}(\mathrm{Ad}(O_{<\ell}))P_0\|_{N^*[I]\to N[I]} \lesssim_M B. \tag{11-3}$$

Proof. Let us assume that $\ell > h_\theta - 20$, as the alternative case is easier.

We decompose the expression on the left-hand side of (11-3) into $\sum_{\theta\in 2^{-\mathbb{N}}} D^{(\theta)}$, where

$$D^{(\theta)} = \mathrm{Op}(\mathrm{ad}(G^{(\theta)})\,\mathrm{Ad}(O_{<\ell}))P_0 - \mathrm{Op}(\mathrm{ad}(G^{(\theta)}))\,\mathrm{Op}(\mathrm{Ad}(O_{<\ell}))P_0.$$

In order to reduce to the case when Lemma 11.1 is applicable, we introduce $h_\theta = \log_2 \theta$ and further decompose $D^{(\theta)}$ as

$$\begin{aligned} D^{(\theta)} = &\int_{h_\theta-20}^{\ell} \mathrm{Op}(\mathrm{ad}(G^{(\theta)})\,\mathrm{ad}(\Psi_h)\,\mathrm{Ad}(O_{<h}))P_0\,dh - \int_{h_\theta-20}^{\ell} \mathrm{Op}(\mathrm{ad}(G^{(\theta)}))\,\mathrm{Op}(\mathrm{ad}(\Psi_h)\,\mathrm{Ad}(O_{<h}))P_0\,dh \\ &+\mathrm{Op}(\mathrm{ad}(G^{(\theta)})\,\mathrm{Ad}(O_{<h_\theta-20})_{\ge h_\theta-10})P_0 - \mathrm{Op}(\mathrm{ad}(G^{(\theta)}))\,\mathrm{Op}(\mathrm{Ad}(O_{<h_\theta-20})_{h_\theta-10})P_0 \\ &+\mathrm{Op}(\mathrm{ad}(G^{(\theta)})\,\mathrm{Ad}(O_{<h_\theta-20})_{<h_\theta-10})P_0 - \mathrm{Op}(\mathrm{ad}(G^{(\theta)}))\,\mathrm{Op}(\mathrm{Ad}(O_{<h_\theta-20})_{<h_\theta-10})P_0. \end{aligned}$$

We claim that

$$\|D^{(\theta)}\|_{L^\infty L^2[I]\to L^1 L^2[I]} \lesssim \theta^{\alpha-\frac{1}{2}} B. \tag{11-4}$$

Assuming (11-4), the proof can be completed by simply summing up in $\theta \in 2^{-\mathbb{N}}$, which is possible since $\alpha > \frac{1}{2} + \delta$.

For the first term in the above splitting of $D^{(\theta)}$, we have

$$\int_{h_\theta-20}^{\ell} \|\mathrm{Op}(\mathrm{ad}(G^{(\theta)})\,\mathrm{ad}(\Psi_h)\,\mathrm{Ad}(O_{<h}))P_0\|_{L^\infty L^2[I]\to L^1L^2[I]}\,dh$$
$$\lesssim_M \int_{h_\theta-20}^{\ell} \|G^{(\theta)}\|_{D_\theta L^2L^\infty[I]}\|\Psi_h\|_{DL^2L^\infty[I]}\,dh$$
$$\lesssim_M \int_{h_\theta-20}^{\ell} \theta^\alpha 2^{(-\frac{1}{2}-\delta)h}B \lesssim_M \theta^{\alpha-\frac{1}{2}-\delta}B.$$

The second term can be handled similarly. For the third term, we use the DL^2L^∞ bound for $G^{(\theta)}$ and apply Lemma 10.12 to $\mathrm{Ad}(O_{<h_\theta-20})_{\geq h_\theta-10})$, which leads to the acceptable bounds

$$\|\mathrm{Op}(\mathrm{ad}(G^{(\theta)})\,\mathrm{Ad}(O_{<h_\theta-20})_{\geq h_\theta-10})P_0\|_{L^\infty L^2[I]\to L^1L^2[I]} \lesssim_M \theta^\alpha B,$$
$$\|\mathrm{Op}(\mathrm{ad}(G^{(\theta)}))\,\mathrm{Op}(\mathrm{Ad}(O_{<h_\theta-20})_{\geq h_\theta-10})P_0\|_{L^\infty L^2[I]\to L^1L^2[I]} \lesssim_M \theta^\alpha B.$$

Finally, for the last term we use Lemma 11.1 (in fact, (11-1)′). □

11B. ***Decomposition of the error.*** Let

$$E = \Box_A^{p,\kappa}\,\mathrm{Op}(\mathrm{Ad}(O)_{<0}) - \mathrm{Op}(\mathrm{Ad}(O)_{<0})\Box.$$

We may take the decomposition

$$E = E_1 + \cdots + E_6,$$

where

$$E_1 = 2i\,\mathrm{Op}\big((\mathrm{ad}(\omega\cdot A_{x,<-\kappa} + A_{0,<-\kappa} + L_+^\omega\Psi)\,\mathrm{Ad}(O))_{<0}\big)|D_x|,$$
$$E_2 = 2i\,\mathrm{Op}\big((\mathrm{ad}(\omega\cdot O_{;x} + O_{;t} - L_+^\omega\Psi)\,\mathrm{Ad}(O))_{<0}\big)|D_x|,$$
$$E_3 = 2\,\mathrm{Op}\big(\mathrm{ad}(A_{\alpha,<-\kappa})(\mathrm{ad}(O^{;\alpha})\,\mathrm{Ad}(O))_{<0}\big) + \mathrm{Op}\big((\mathrm{ad}(O_{;\alpha})\,\mathrm{ad}(O^{;\alpha})\,\mathrm{Ad}(O))_{<0}\big),$$
$$E_4 = \mathrm{Op}((\mathrm{ad}(\partial^\alpha O_{;\alpha})\,\mathrm{Ad}(O))_{<0}),$$
$$E_5 = -2i\,\mathrm{Op}(\mathrm{ad}(A_{0,<-\kappa})\,\mathrm{Ad}(O)_{<0})(D_t+|D_x|) - 2i\,\mathrm{Op}((\mathrm{ad}(O_{<-\kappa;t})\,\mathrm{Ad}(O))_{<0})(D_t+|D_x|),$$
$$E_6 = -2i\,\mathrm{Op}\big([S_{<0}, \mathrm{ad}(\omega\cdot A_{x,<-\kappa} + A_{0,<-\kappa})]\,\mathrm{Ad}(O)\big)|D_x|.$$

In the remainder of this section, we estimate each error term in order.

11C. ***Estimate for E_1.*** Here, our goal is to prove

$$\|E_1P_0\|_{S_0^\sharp[I]\to N[I]} \leq \varepsilon, \tag{11-5}$$

with κ_1 large enough and δ_p sufficiently small.

11C1. *Preliminary reduction.* For this term, we may simply work with $I = \mathbb{R}$ by extending the input by homogeneous waves outside I. The desired smallness comes from κ and bounds for $\Box A_x$ and ΔA_0 on I, which controls the size of the symbol of E_1 through our extension of A_α as in Section 9B.

We first dispose of the symbol regularization $(\cdot)_{<0}$ by translation invariance, and also throw away $|D_x|$ using P_0. Using (9-42) and the identity

$$L_+^{\omega} L_-^{\omega} \Delta_{\omega^\perp}^{1} = -\Delta_{\omega^\perp}^{-1} \Box + 1,$$

(11-5) reduces to showing

$$\left\| \int_{-\infty}^{-\kappa} \mathrm{Op}(\mathrm{ad}(G_h)\,\mathrm{Ad}(O)) P_0 \, dh \right\|_{S_0^\sharp \to N} \ll \varepsilon,$$

where

$$G_h = \omega \cdot A_{x,h} - \omega \cdot A_{x,h,\mathrm{cone}}^{(\geq|\eta|^\delta)} + \Delta_{\omega^\perp}^{-1} \Box (\omega \cdot A_{x,h,\mathrm{cone}}^{(\geq|\eta|^\delta)}) + A_{0,h}.$$

Note that each angular component $G_h^{(\theta)} = \Pi_\theta^{\omega,+} G_h$ obeys

$$\|G_h^{(\theta)}\|_{DL^2L^\infty} \lesssim 2^{\frac{1}{2}h} \theta^{\frac{3}{2}} (\|A_{x,h}\|_{S^1} + \|A_{0,h}\|_{Y^1}).$$

Therefore, by Lemma 11.3, we have

$$\left\| \int_{-\infty}^{-\kappa} \big(\mathrm{Op}(\mathrm{ad}(G_h)\,\mathrm{Ad}(O)) - \mathrm{Op}(\mathrm{ad}(G_h))\,\mathrm{Op}(\mathrm{Ad}(O))\big) P_0 \, dh \right\|_{N^* \to N} \lesssim_M 2^{-\frac{1}{2}\kappa},$$

which is acceptable. By Lemma 10.12 applied to $\mathrm{Op}(\mathrm{Ad}(O)_{\geq 0})$, we also have

$$\begin{aligned} \left\| \int_{-\infty}^{-\kappa} \mathrm{Op}(\mathrm{ad}(G_h))\,\mathrm{Op}(\mathrm{Ad}(O)_{\geq 0}) P_0 \, dh \right\|_{N^* \to N} &\lesssim_M \int_{-\infty}^{-\kappa} 2^{\frac{1}{2}h} \|\mathrm{Op}(\mathrm{Ad}(O)_{\geq 0}) P_0\|_{L^\infty L^2 \to L^2 L^2} \, dh \\ &\lesssim_M 2^{-\frac{1}{2}\kappa}. \end{aligned}$$

Thus it suffices to show that

$$\left\| \int_{-\infty}^{-\kappa} \mathrm{Op}(\mathrm{ad}(G_h))\,\mathrm{Op}(\mathrm{Ad}(O)_{<0}) P_0 \, dh \right\|_{S_0^\sharp \to N} \ll \varepsilon.$$

By (9-45), we have $\mathrm{Op}(\mathrm{Ad}(O)_{<0}) P_0 : S_0^\sharp \to S_0$. Thus, in order to prove (11-5), we are left to establish

$$\left\| \int_{-\infty}^{-\kappa} \mathrm{Op}(\mathrm{ad}(G_h)) P_0 \, dh \right\|_{S_0 \to N} \ll \varepsilon, \tag{11-6}$$

where we abuse the notation a bit and denote by P_0 a frequency projection to a slightly enlarged region of the form $\{|\xi| \simeq 1\}$.

At this point it is convenient to observe that the contribution of $\widetilde{R}_0$ to A_0 in (9-27) is easy to estimate in $L^1 L^\infty$ and can be harmlessly discarded. Thus from here on we assume that

$$\widetilde{R}_0 = 0. \tag{11-7}$$

In order to proceed, we write

$$G_h = G_{h,\mathrm{cone}} + G_{h,\mathrm{null}} + G_{h,\mathrm{out}},$$

where

$$G_{h,\text{cone}} = \omega \cdot A^{(<|\eta|^{\delta})}_{x,h,\text{cone}} + \Delta^{-1}_{\omega^{\perp}} \Box(\omega \cdot A^{(\geq|\eta|^{\delta})}_{x,h,\text{cone}}) + A_{0,h,\text{cone}},$$
$$G_{h,\text{null}} = \omega \cdot A_{x,h,\text{null}} + A_{0,h,\text{null}},$$
$$G_{h,\text{out}} = \omega \cdot A_{x,h,\text{out}} + A_{0,h,\text{out}}.$$

11C2. *Estimate for $G_{h,\text{cone}}$.* We claim that

$$\left\| \int_{-\infty}^{-\kappa} \text{Op}(\text{ad}(G_{h,\text{cone}})) P_0 \, dh \right\|_{N^* \to N} \ll \varepsilon. \tag{11-8}$$

Let $G^{(\theta)}_{h,\text{cone}} = \Pi^{\omega,\pm}_{\theta} G_{h,\text{cone}}$ and consider the expression $\text{Op}(\text{ad}(G^{(\theta)}_{h,\text{cone}}))P_0$. By the Fourier support property of $G^{(\theta)}_{h,\text{cone}}$ (more precisely, the mismatch between its modulation $\lesssim 2^h\theta^2$ and the angle θ), it is impossible that both the input and the output have modulation $\ll 2^h\theta^2$. Using the L^2L^2 norm for the input or the output (whichever that has modulation $\gtrsim 2^h\theta^2$), we may estimate

$$\begin{aligned}
&\|\text{Op}(G_{h,\text{cone}})P_0\|_{N^*\to N} \\
&\qquad \lesssim \sum_{\theta<1} 2^{-\frac{1}{2}h}\theta^{-1}\|G^{(\theta)}_{h,\text{cone}}\|_{DL^2L^\infty} \\
&\qquad \lesssim 2^{\frac{\delta}{2}h}\|A_{x,h}\|_{S^1} + \sum_{\theta<1} 2^{-\frac{1}{2}h}\theta^{-\frac{1}{2}}\|Q_{<h+2\log_2\theta+C}\Box A_x\|_{L^2L^2} + \sum_{\theta<1} 2^{-\frac{1}{2}h}\theta^{\frac{1}{2}}\|\Delta A_{0,h}\|_{L^2L^2}.
\end{aligned}$$

We now treat each term separately.

Case 1: contribution of small angle interaction. The term $2^{(\delta/2)h}\|A_{x,h}\|_{S^1}$ is acceptable since it is integrable in $-\infty < h < -\kappa$, and we gain a small factor $2^{-(\delta/2)\kappa}$ as a result.

Case 2: contribution of $\Box A_x$. For the second term, we split the θ-summation into $\theta < 2^{-\kappa}$ and $\theta \geq 2^{-\kappa}$. In the former case, note that

$$\|Q_{<h+2\log_2\theta+C}\Box A_x\|_{L^2L^2} \lesssim \theta^{2b_1}\|\Box A_{x,h}\|_{X^{-1/2+b_1,-b_1}}.$$

Since $b_1 > \frac{1}{4}$, we may estimate

$$\sum_{\theta<2^{-\kappa}} 2^{-\frac{1}{2}h}\theta^{-\frac{1}{2}}\|Q_{<h+2\log_2\theta+C}\Box A_x\|_{L^2L^2} \lesssim 2^{-(2b_1-\frac{1}{2})\kappa}\|\Box A_{x,h}\|_{X^{-1/2+b_1,-b_1}}.$$

The last line is acceptable, since it is integrable in $-\infty < h < -\kappa$, and it is small thanks to $2^{-(2b-1/2)\kappa}$. In the case $\theta \geq 2^{-\kappa}$, we estimate

$$\sum_{\theta\geq 2^{-\kappa}} 2^{-\frac{1}{2}h}\theta^{-\frac{1}{2}}\|Q_{<h+2\log_2\theta+C}\Box A_x\|_{L^2L^2} \lesssim 2^{\frac{1}{2}\kappa}\|\Box A_{x,h}\|_{L^2\dot{H}^{-1/2}}.$$

After integration in h, this is acceptable thanks to (9-22).

Case 3: contribution of A_0. In this case, we simply sum up in $\theta < 1$ and observe that

$$\sum_{\theta<1} 2^{-\frac{1}{2}h}\theta^{\frac{1}{2}}\|\Delta A_{0,h}\|_{L^2L^2} \lesssim \|\Delta A_{0,h}\|_{L^2\dot{H}^{-1/2}}.$$

After integration in h, this term is then acceptable by (9-29).

11C3. *Estimate for $G_{h,\mathrm{out}}$.* We claim that

$$\Big\| \int_{-\infty}^{-\kappa} \mathrm{Op}(\mathrm{ad}(G_{h,\mathrm{out}})) P_0 \, dh \Big\|_{N^* \to N} \ll \varepsilon. \tag{11-9}$$

As in the case of $G_{h,\mathrm{cone}}$, the idea is again to make use of the mismatch between modulation of $G_{h,\mathrm{out}}$ and the angle θ. Let $G^{(\theta)}_{h,\mathrm{out}} = \Pi^{\omega,\pm}_{\theta} G_{h,\mathrm{out}}$, and consider the expression $\mathrm{Op}(\mathrm{ad}(G^{(\theta)}_{h,\mathrm{out}})) P_0$. By definition, $G^{(\theta)}_{h,\mathrm{out}}$ has modulation $\gtrsim 2^h \theta^2$. Thus, we take the decomposition $G^{(\theta)}_{h,\mathrm{out}} = \sum_{a:2^a \gtrsim \theta} Q_{h+2a} G^{(\theta)}_{h,\mathrm{out}}$. By the Fourier support property of the symbol $Q_{h+2a} G^{(\theta)}_{h,\mathrm{out}}$ (more precisely, the mismatch between the angle θ and the modulation 2^{h+2a}), it is impossible that both the input and the output have modulation $\ll 2^{h+2a}$. Using the $L^2 L^2$ norm for the input or the output, we have

$$\begin{aligned}
&\|\mathrm{Op}(\mathrm{ad}(G_{h,\mathrm{out}})) P_0\|_{N^* \to N} \\
&\quad \lesssim \sum_{a} \sum_{\theta < \min\{C 2^a, 1\}} 2^{-\frac{1}{2}(h+2a)} \| Q_{h+2a} G^{(\theta)}_{h,\mathrm{out}} \|_{D L^2 L^\infty} \\
&\quad \lesssim \sum_{a} \sum_{\theta < \min\{C 2^a, 1\}} (2^{-\frac{1}{2}(h+2a)} 2^{2h} \theta^{\frac{5}{2}} \| Q_{h+2a} A_{x,h} \|_{L^2 L^2} + 2^{-\frac{1}{2}(h+2a)} 2^{2h} \theta^{\frac{3}{2}} \| A_{0,h} \|_{L^2 L^2}) \\
&\quad \lesssim \sum_{a} (2^{\frac{5}{2} a -} 2^{-3a} 2^{-\frac{1}{2} h} \| Q_{h+2a} \Box A_{x,h} \|_{L^2 L^2} + 2^{\frac{3}{2} a -} 2^{-a} 2^{-\frac{1}{2} h} \| \Delta A_{0,h} \|_{L^2 L^2}).
\end{aligned}$$

We split the a-summation into $a < -\kappa$ and $a > -\kappa$. In the former case, the sum is bounded by

$$2^{-(2b_1 - \frac{1}{2})\kappa} \| \Box A_{x,h} \|_{X^{b_1 - 1/2, -b_1}} + 2^{-\frac{1}{2}\kappa} \| \Delta A_{0,h} \|_{L^2 \dot{H}^{-1/2}},$$

which is integrable in h and small thanks to $2^{-(2b_1 - 1/2)\kappa}$; therefore it is acceptable. When $a > -\kappa$, the sum is bounded by

$$2^{\frac{1}{2}\kappa} \| \Box A_{x,h} \|_{L^2 \dot{H}^{1/2}} + \| \Delta A_{0,h} \|_{L^2 \dot{H}^{-1/2}}.$$

After integrating in h, this term is therefore acceptable by (9-22) and (9-29).

11C4. *Estimate for $G_{h,\mathrm{null}}$.* We claim that

$$\Big\| \int_{-\infty}^{-\kappa} \mathrm{Op}(\mathrm{ad}(G_{h,\mathrm{null}})) P_0 \, dh \Big\|_{S_0 \to N} \ll \varepsilon. \tag{11-10}$$

Let $G^{(\theta)}_{h,\mathrm{null}} = \Pi^{\omega,\pm}_{\theta} G_{h,\mathrm{null}}$. Note that $G^{(\theta)}_{h,\mathrm{null}}$ has modulation $\simeq 2^h \theta^2$. Hence if either the input or the output have modulation $\geq 2^{-C} 2^h \theta^2$, the same argument as in the case of $G_{h,\mathrm{cone}}$ applies. Writing $\theta = 2^\ell$, it remains to prove

$$\Big\| \sum_{\ell \in -\mathbb{N}} \int_{-\infty}^{-\kappa} Q_{<h+2\ell-C} \, \mathrm{Op}(\mathrm{ad}(\omega \cdot A^{(2^\ell)}_{x,h,\mathrm{null}} + A^{(2^\ell)}_{0,h,\mathrm{null}})) P_0 Q_{<h+2\ell-C} \, dh \Big\|_{S_0 \to N} \ll \varepsilon. \tag{11-11}$$

Our next simplification is to observe that we can harmlessly replace the symbols $A^{(2^\ell)}_{x,h,\mathrm{null}}$ and $A^{(2^\ell)}_{0,h,\mathrm{null}}$ with the functions $Q_{h+2\ell} A_{x,h}$ and $Q_{h+2\ell} A_{x,h}$. This is because the difference of the two is localized still at modulation $2^{h+2\ell}$, but also at distance $2^{h+2\ell}$ from the null plane $\{\sigma + \omega \cdot \eta = 0\}$. This would force either the input or the output modulation in (11-11) to be $\geq 2^{-C}2^{h+2\ell}$, and again the same argument as in the case of $G_{h,\mathrm{cone}}$ applies. Thus with $j = h + 2\ell$ we have reduced the problem to estimating

$$\Big\| \sum_{j<h} \int_{-\infty}^{-\kappa} Q_{<j-C} \,\mathrm{ad}(Q_j A_{\alpha,h}) \partial^\alpha P_0 Q_{<j-C} \, dh \Big\|_{S_0 \to N} \ll \varepsilon, \tag{11-12}$$

$$\Big\| \sum_{j<h} \int_{-\infty}^{-\kappa} Q_{<j-C} \,\mathrm{ad}(Q_j A_{0,h})(D_0 + |D_x|) P_0 Q_{<j-C} \, dh \Big\|_{S_0 \to N} \ll \varepsilon. \tag{11-13}$$

The second bound is straightforward since $(D_0 + |D_x|)P_0 Q_{<0} : S_0 \to L^2$ and $A_0 \in L^2 \dot{H}^{3/2}$.

Thus it remains to consider (11-12). From here on, we assume that A is determined by the expressions (9-27) and (9-30) in terms of $\tilde{A}$. By (11-7) we have already set $\widetilde{R}_0 = 0$. It is equally easy to see that we can set $\widetilde{R}_x = 0$. Indeed, by (4-6) and (8-30) we have

$$\| Q_{<j-C} \,\mathrm{ad}(\Box^{-1} P_h R_\ell) \partial^\ell P_0 Q_{<j-C} \|_{S_0 \to N} \lesssim 2^{\delta_1 (j-h)} \| \Box^{-1} P_h R_\ell \|_{Z^1} \lesssim 2^{\delta_1 (j-h)} \| P_h R_\ell \|_{L^1 L^2},$$

where $R = \chi_I \boldsymbol{P} \widetilde{R}$. Now the summability in $j < h$ and the smallness is assured due to (9-26).

Once we have dispensed with the error terms, we are left with $A_{t,x}$ given by

$$A_0 = \Delta^{-1} \boldsymbol{O}(\chi_I \tilde{A}^\ell, \partial_t \tilde{A}_\ell), \tag{11-14}$$

$$A = \Box^{-1} \boldsymbol{P}(\boldsymbol{O}(\chi_I \tilde{A}^\ell, \partial_x \tilde{A}_\ell) + \boldsymbol{O}'(\boldsymbol{P}_\ell \tilde{A}, \chi_I \partial^\ell \tilde{A}) - \boldsymbol{O}'(\tilde{A}_0, \chi_I \partial_t \tilde{A}) + \boldsymbol{O}'(\widetilde{G}_\ell, \chi_I \partial^\ell \tilde{A})). \tag{11-15}$$

We consider the contributions of each of these terms in (11-12).

<u>Step 1</u>: the contribution of $A_0 = \Delta^{-1} \boldsymbol{O}(\chi_I \tilde{A}^\ell, \partial_t \tilde{A}_\ell)$ and $A_x = \Box^{-1} \boldsymbol{P}\boldsymbol{O}(\chi_I \tilde{A}^\ell, \partial_x \tilde{A}_\ell)$. This is the main component, which we have to treat in a trilinear fashion. In particular we have to ensure that we gain smallness. For this we use a trilinear Littlewood–Paley decomposition to set

$$A = \sum_{k,k_1,k_2} A(k,k_1,k_2) = \sum_{k,k_1,k_2} \mathcal{H} A(k,k_1,k_2) + \sum (1 - \mathcal{H}^*) A(k,k_1,k_2),$$

where

$$\mathcal{H} A(k,k_1,k_2) := \mathcal{H} P_k \boldsymbol{P} A(P_{k_1} \chi_I \tilde{A}^\ell, P_{k_2} \partial_t \tilde{A}_\ell),$$
$$(1-\mathcal{H}) A(k,k_1,k_2) := (1-\mathcal{H}) P_k \boldsymbol{P} A(P_{k_1} \chi_I \tilde{A}^\ell, P_{k_2} \partial_t \tilde{A}_\ell).$$

For the terms in the first sum we use the trilinear estimate (8-43), which gives

$$\| Q_{<j-C} \,\mathrm{ad}(Q_j \mathcal{H} A_\alpha(k,k_1,k_2)) \partial^\alpha P_0 Q_{<j-C} \|_{S_0 \to L^1 L^2} \lesssim 2^{-\delta_1 |k_{\max} - k_{\min}|} 2^{\delta_1 (j-k)} \| P_{k_1} \tilde{A} \|_{S^1} \| P_{k_2} \tilde{A} \|_{S^1}.$$

For the A_x terms in the second sum we first use (8-21) and (8-33), (8-34) to obtain

$$\| (1-\mathcal{H}) A_x(k,k_1,k_2) \|_{Z^1} \lesssim 2^{-\delta_1 |k_{\max} - k_{\min}|} \| P_{k_1} \tilde{A} \|_{S^1} \| P_{k_2} \tilde{A} \|_{S^1}$$

and then use (8-30) to conclude that

$$\| Q_{<j-C} \,\mathrm{ad}(Q_j (1-\mathcal{H}) A_\ell(k,k_1,k_2)) \partial^\ell P_0 Q_{<j-C} \|_{S_0 \to N} \lesssim 2^{-\delta_1 |k_{\min} - k_{\max}|} 2^{\delta_1 (j-k)} \| P_{k_1} \tilde{A} \|_{S^1} \| P_{k_2} \tilde{A} \|_{S^1}.$$

Similarly, for the A_0 terms in the second sum we use (8-35) and then (8-31) to obtain

$$\| Q_{<j-C} \,\mathrm{ad}(Q_j (1-\mathcal{H}) A_0(k,k_1,k_2)) \partial^0 P_0 Q_{<j-C} \|_{S_0 \to N} \lesssim 2^{-\delta_1 |k_{\min} - k_{\max}|} 2^{\delta_1 (j-k)} \| P_{k_1} \tilde{A} \|_{S^1} \| P_{k_2} \tilde{A} \|_{S^1}.$$

Adding the last three bounds, we obtain

$$\|Q_{<j-C}\,\mathrm{ad}(Q_j A_\alpha(k,k_1,k_2))\partial^\alpha P_0 Q_{<j-C}\|_{S_0\to N} \lesssim 2^{-\delta_1|k_{\max}-k_{\min}|}2^{\delta_1(j-k)}\|P_{k_1}\tilde{A}\|_{S^1}\|P_{k_2}\tilde{A}\|_{S^1}.$$

This gives both summability in k, k_1, k_2 and smallness provided we exclude the range of indices $j, k_1, k_2 \in [k-\kappa', k+\kappa']$ with $\kappa' \gg 1$.

On the other hand, in the range excluded above, the operator $P_k Q_j$ is disposable, while both $\Box$ and Δ are elliptic, i.e., of size 2^{2k}. Then we can estimate

$$\|Q_j A(k,k_1,k_2)\|_{L^1L^\infty} \lesssim 2^{C\kappa'}\|P_{k_1}\tilde{A}\|_{DS^1}\|P_{k_2}\tilde{A}\|_{DS^1};$$

therefore we gain smallness from the divisible norm; see (9-5).

Step 2: the contribution of $A_x = \Box^{-1}\boldsymbol{PO}'(\boldsymbol{P}_\ell\tilde{A}, \chi_I\partial^\ell\tilde{A})$. This is a milder contribution, which we can deal with in a bilinear fashion. Taking again the decomposition

$$A_x = \sum_{k,k_1,k_2} A(k,k_1,k_2),$$

we use (8-38) to obtain

$$\|A_x(k,k_1,k_2)\|_{Z^1} \lesssim 2^{-\delta_1|k_{\max}-k_{\min}|}\|P_{k_1}\tilde{A}\|_{\underline{S}^1}\|P_{k_1}\tilde{A}\|_{S^1}.$$

Then by (8-30) it follows that

$$\|Q_{<j-C}\,\mathrm{ad}(Q_j\mathcal{H}A_x(k,k_1,k_2))\partial^\alpha P_0 Q_{<j-C}\|_{S_0\to L^1L^2} \lesssim 2^{-\delta_1|k_{\max}-k_{\min}|}2^{\delta_1(j-k)}\|P_{k_1}\tilde{A}\|_{\underline{S}^1}\|P_{k_1}\tilde{A}\|_{S^1}. \quad (11\text{-}16)$$

Again this is suitable outside the range $j, k_1, k_2 \in [k-\kappa', k+\kappa']$ with $\kappa' \gg 1$, whereas in this range we can use divisible norms as in the previous step.

Step 3: the contribution of $\boldsymbol{PO}'(\tilde{A}_0, \chi_I\partial_t\tilde{A}) + \boldsymbol{PO}'(\widetilde{G}_\ell, \chi_I\partial^\ell\tilde{A})$. These two terms are similar, as we have the same bounds available for $\tilde{A}_0$ and $\widetilde{G}_l$. We will discuss $\tilde{A}_0$. Setting

$$A_x = \Box^{-1}\boldsymbol{PO}'(\tilde{A}_0, \chi_I\partial_t\tilde{A}), \quad A_0 = 0,$$

we decompose A as before,

$$A_x = \sum A_x(k,k_1,k_2).$$

We can estimate the terms in the sum using (8-41) to get

$$\|A_x(k,k_1,k_2)\|_{Z^1} \lesssim 2^{-\delta_1|k_{\max}-k_{\min}|}\|P_{k_1}\tilde{A}_0\|_{Y^1}\|P_{k_1}\tilde{A}\|_{S^1}.$$

Then (11-16) follows again from (8-30), and we conclude as in Step 2.

11D. *Estimate for E_2.* Our next goal is to estimate the error term E_2, which arises from the multilinear error between $O_{;\alpha}$ and $\partial_\alpha\Psi$. For this purpose, we rely crucially on interval localization of decomposable norms (Lemma 10.7).

11D1. *Expansion of $O_{;\alpha}$.* We will prove that

$$\|E_2 P_0\|_{N^*[I]\to N[I]} \le \varepsilon \tag{11-17}$$

provided that κ_1 is large enough, and δ_p is sufficiently small.

As usual, we may dispose of the symbol regularization $(\cdot)_{<0}$ by translation invariance. Also disposing of $|D_x|$ using P_0, it suffices to prove

$$\big\|\mathrm{Op}\big(\mathrm{ad}(\omega\cdot(O_{;x}-\partial_x\Psi)+(O_{;t}-\partial_t\Psi))\,\mathrm{Ad}(O)\big)P_0\big\|_{N^*[I]\to N[I]} \ll \varepsilon. \tag{11-18}$$

Recall that $\partial_h O_{<h;\alpha} = \Psi_{h,\alpha} + [\Psi_h, O_{<h;\alpha}]$. Therefore,

$$\partial_h(\mathrm{ad}(O_{<h;\alpha})\,\mathrm{Ad}(O_{<h})) = \mathrm{ad}(\partial_\alpha\Psi_h)\,\mathrm{Ad}(O_{<h}) + \mathrm{ad}(\Psi_h)\,\mathrm{Ad}(O_{<h;\alpha})\,\mathrm{Ad}(O_{<h}).$$

Repeatedly applying the fundamental theorem of calculus and this equation, we obtain the expansion

$$\begin{aligned}
&\mathrm{ad}(O_{;\alpha})\,\mathrm{Ad}(O)\\
&\quad= \int_{-\infty}^{-\kappa} \mathrm{ad}(\partial_\alpha\Psi_{h_1})\,\mathrm{Ad}(O_{<h_1})\,dh_1 &&(11\text{-}19)\\
&\qquad+ \int_{-\infty}^{-\kappa}\int_{-\infty}^{h_1} \mathrm{ad}(\Psi_{h_1})\,\mathrm{ad}(\partial_\alpha\Psi_{h_2})\,\mathrm{Ad}(O_{<h_2})\,dh_2\,dh_1 &&(11\text{-}20)\\
&\qquad+\cdots\\
&\qquad\quad+ \int_{-\infty}^{-\kappa}\int_{-\infty}^{h_1}\cdots\int_{-\infty}^{h_5} \mathrm{ad}(\Psi_{h_1})\,\mathrm{ad}(\Psi_{h_2})\cdots\mathrm{ad}(\partial_\alpha\Psi_{h_6})\,\mathrm{Ad}(O_{<h_6})\,dh_6\cdots dh_2\,dh_1. &&(11\text{-}21)
\end{aligned}$$

On the other hand,

$$\partial_h(\mathrm{ad}(\partial_\alpha\Psi_{<h})\,\mathrm{Ad}(O_{<h})) = \mathrm{ad}(\partial_\alpha\Psi_h)\,\mathrm{Ad}(O_{<h}) + \mathrm{ad}(\partial_\alpha\Psi_{<h})\,\mathrm{ad}(\Psi_h)\,\mathrm{Ad}(O_{<h}),$$

so we have

$$\begin{aligned}
\mathrm{ad}(\partial_\alpha\Psi)\,\mathrm{Ad}(O) &= \int_{-\infty}^{-\kappa} \mathrm{ad}(\partial_\alpha\Psi_{h_1})\,\mathrm{Ad}(O_{<h_1})\,dh_1 &&(11\text{-}22)\\
&\qquad+ \int_{-\infty}^{-\kappa}\int_{-\infty}^{h_1} \mathrm{ad}(\partial_\alpha\Psi_{h_2})\,\mathrm{ad}(\Psi_{h_1})\,\mathrm{Ad}(O_{<h_1})\,dh_2\,dh_1. &&(11\text{-}23)
\end{aligned}$$

Observe that (11-19) and (11-22) coincide. Thus, we only need to consider the contribution of (11-20)–(11-21) and (11-23) in (11-18).

11D2. *Estimate for quadratic expressions.* We begin with the contribution of the quadratic terms in Ψ, namely (11-20) and (11-23), which are most delicate. We claim that

$$\left\|\int_{-\infty}^{-\kappa}\int_{-\infty}^{h_1} \mathrm{Op}(\mathrm{ad}(\Psi_{h_1})\,\mathrm{ad}(L_+^\omega\Psi_{h_2})\,\mathrm{Ad}(O_{<h_2}))P_0\,dh_2\,dh_1\right\|_{N^*[I]\to N[I]} \le \varepsilon, \tag{11-24}$$

$$\left\|\int_{-\infty}^{-\kappa}\int_{-\infty}^{h_1} \mathrm{Op}(\mathrm{ad}(L_+^\omega\Psi_{h_2})\,\mathrm{ad}(\Psi_{h_1})\,\mathrm{Ad}(O_{<h_1}))P_0\,dh_2\,dh_1\right\|_{N^*[I]\to N[I]} \le \varepsilon, \tag{11-25}$$

provided that κ_1 is large enough and δ_p is sufficiently small. In what follows, we will focus on establishing (11-24), as the proof for the other claim is analogous.

By (9-42) and the identity $L^\omega_+ L^\omega_- \Delta^1_{\omega^\perp} = -\Delta^{-1}_{\omega^\perp}\Box + 1$, (11-24) would follow once we establish

$$\left\| \int_{-\infty}^{-\kappa} \int_{-\infty}^{h_1} \mathrm{Op}\big(\mathrm{ad}(\Psi_{h_1})\,\mathrm{ad}(\omega \cdot A^{\mathrm{main}}_{h_2})\,\mathrm{Ad}(O_{<h_2})\big) P_0 \, dh_2 \, dh_1 \right\|_{N^*[I]\to N[I]} \ll \varepsilon, \tag{11-26}$$

$$\left\| \int_{-\infty}^{-\kappa} \int_{-\infty}^{h_1} \mathrm{Op}\big(\mathrm{ad}(\Psi_{h_1})\,\mathrm{ad}(\Delta^{-1}_{\omega^\perp}\Box(\omega \cdot A^{\mathrm{main}}_{h_2}))\,\mathrm{Ad}(O_{<h_2})\big) P_0 \, dh_2 \, dh_1 \right\|_{N^*[I]\to N[I]} \ll \varepsilon. \tag{11-27}$$

In Lemmas 10.4 and 10.7, note that $\omega \cdot A^{\mathrm{main},(\theta)}_h (= \omega \cdot A^{(\theta)}_{x,h,\mathrm{cone},+})$ and $\Delta^{-1}_{\omega^\perp}\Box(\omega \cdot A^{\mathrm{main},(\theta)}_h)$ obey the same bounds. Therefore, (11-26) and (11-27) are proved in exactly the same way. In what follows, we only consider (11-26).

Our first task is to remove $\mathrm{Ad}(O_{<h_2})$. For $\theta \in 2^{-\mathbb{N}}$, define

$$G^{(\theta)} = \mathrm{ad}(\Psi^{(\theta)}_{h_1})\,\mathrm{ad}(\omega \cdot A^{\mathrm{main},(<\theta)}_{h_2}) + \mathrm{ad}(\Psi^{(\le\theta)}_{h_1})\,\mathrm{ad}(\omega \cdot A^{\mathrm{main},(\theta)}_{h_2}).$$

so that

$$G := \mathrm{ad}(\Psi_{h_1})\,\mathrm{ad}(\omega \cdot A^{\mathrm{main}}_{h_2}) = \sum_{\theta \in 2^{-\mathbb{N}}} G^{(\theta)}.$$

Note that

$$\|G^{(\theta)}\|_{DL^2L^\infty} \lesssim_M 2^{\frac{1}{2}h_1} 2^{\frac{1}{2}(h_2-h_1)} \theta^{\frac{3}{2}},$$

by Lemma 10.4 and Lemma 10.5. Applying Lemma 11.3, then integrating $-\infty < h_2 < h_1 < -\kappa$, it follows that

$$\left\| \int_{-\infty}^{-\kappa} \int_{-\infty}^{h_1} \big(\mathrm{Op}(\mathrm{ad}(G)\,\mathrm{Ad}(O_{<h_2})) - \mathrm{Op}(\mathrm{ad}(G))\,\mathrm{Op}(\mathrm{Ad}(O_{<h_2}))\big) P_0 \, dh_2 \, dh_1 \right\|_{N^*[I]\to N[I]} \lesssim 2^{-\frac{1}{2}\kappa},$$

which is acceptable. On the other hand, using the DL^2L^∞ bound for G and Lemma 10.12, we have

$$\begin{aligned} &\left\| \int_{-\infty}^{-\kappa} \int_{-\infty}^{h_1} \mathrm{Op}(\mathrm{ad}(G))\,\mathrm{Op}(\mathrm{Ad}(O_{<h_2})_{\ge 0}) P_0 \, dh_2 \, dh_1 \right\|_{N^*[I]\to N[I]} \\ &\qquad \lesssim_M \int_{-\infty}^{-\kappa} \int_{-\infty}^{h_1} 2^{\frac{1}{2}h_1} 2^{\frac{1}{2}(h_2-h_1)} \|\mathrm{Op}(\mathrm{Ad}(O_{<h_2})_{\ge 0}) P_0\|_{L^\infty L^2[I]\to L^2L^2[I]} \, dh_2 \, dh_1 \\ &\qquad \lesssim_M 2^{-\frac{1}{2}\kappa}, \end{aligned}$$

so we may replace $\mathrm{Op}(\mathrm{Ad}(O_{<h_2}))$ by $\mathrm{Op}(\mathrm{Ad}(O_{<h_2}))_{<0}$. Finally, by (9-44) we have

$$\mathrm{Op}(\mathrm{Ad}(O_{<h_2})_{<0}) P_0 : N^*[I] \to N^*[I],$$

so we are left to prove

$$\left\| \int_{-\infty}^{0} \int_{-\infty}^{h_1} \mathrm{Op}(\mathrm{ad}(\Psi_{h_1})\,\mathrm{ad}(\omega \cdot A^{\mathrm{main}}_{h_2})) \, dh_2 \, dh_1 \right\|_{N^*[I]\to N[I]} \ll \varepsilon. \tag{11-28}$$

In order to place ourselves in a context where we can apply Lemma 10.7, we begin by dispensing with the case of short intervals

$$|I| \leq 2^{-h_2 - C\kappa}.$$

For very short intervals $|I| \leq 2^{-h_1 - C\kappa}$ we have the bound

$$\Big\| \int_{-\infty}^{0} \int_{-\infty}^{h_1} \mathrm{Op}(\mathrm{ad}(\Psi_{h_1})\,\mathrm{ad}(\omega \cdot A_{h_2}^{\mathrm{main}}))\, dh_2\, dh_1 \Big\|_{L^\infty L^2 \to L^1 L^2} \lesssim_M 2^{h_2} |I|,$$

which is a consequence of fixed-time decomposability bounds, namely (10-10) with $q = \infty$ and (10-14) with $q = \infty$ and $r = \infty$, combined with Hölder's inequality in time. This suffices for the integration with respect to h_1 and h_2 in this range.

For merely short intervals $2^{-h_1 - C\kappa} \leq |I| \leq 2^{-h_2 - C\kappa}$ we are allowed to use space-time decomposabilty bounds but only for Ψ_{h_1}. In this case we apply (10-10) with $q = \infty$ and (10-14) with $q = 6$ and $r = \infty$, combined with Hölder's inequality in time, to obtain

$$\Big\| \int_{-\infty}^{0} \int_{-\infty}^{h_1} \mathrm{Op}(\mathrm{ad}(\Psi_{h_1})\,\mathrm{ad}(\omega \cdot A_{h_2}^{\mathrm{main}}))\, dh_2\, dh_1 \Big\|_{L^\infty L^2 \to L^1 L^2} \lesssim_M 2^{-\frac{1}{6}h_1} 2^{h_2} |I|^{\frac{5}{6}}.$$

This again suffices for the integration with respect to h_1 and h_2 in this range.

For large intervals, on the other hand, we will use Lemma 10.7. We begin by decomposing $\Psi_{h_1} = \sum_{\theta_1} \Psi_{h_1}^{(\theta_1)}$ and $A_{h_2}^{\mathrm{main}} = \sum_{\theta_2} A_{h_2}^{\mathrm{main},(\theta_2)}$. First, we consider the case $2^{h_1}\theta_1^2 \geq 2^{-2\kappa} 2^{h_2} \theta_2^2$. For fixed h_1, h_2 and θ_2, we use interval localized decomposability calculus to estimate

$$\begin{aligned} \sum_{\theta_1 \geq 2^{-\kappa} 2^{(1/2)(h_2 - h_1)} \theta_2} & \| \mathrm{Op}(\mathrm{ad}(\Psi_{h_1}^{(\theta_1)})\,\mathrm{ad}(\omega \cdot A_{h_2}^{\mathrm{main},(\theta_2)})) \|_{L^\infty L^2[I] \to L^1 L^2[I]} \\ &\lesssim \sum_{\theta_1 \geq 2^{-\kappa} 2^{(1/2)(h_2 - h_1)} \theta_2} \| \Psi_{h_1}^{(\theta_1)} \|_{DL^2 L^\infty[I]} \| \omega \cdot A_{h_2}^{\mathrm{main},(\theta_2)} \|_{DL^2 L^\infty[I]} \\ &\lesssim 2^{\kappa} 2^{\frac{1}{4}(h_2 - h_1)} \theta_2 \| A_{h_1} \|_{S^1} (2^{-\frac{1}{2}h_2} \theta_2^{-\frac{3}{2}} \| \omega \cdot A_{h_2}^{\mathrm{main},(\theta_2)} \|_{DL^2 L^\infty[I]}). \end{aligned}$$

Summing up in $\theta_2 < 2^{-2\kappa}$, we see that

$$\begin{aligned} \sum_{\theta_2 < 2^{-2\kappa}} \sum_{\theta_1 \geq 2^{-\kappa} 2^{(1/2)(h_2 - h_1)} \theta_2} & \| \mathrm{Op}(\mathrm{ad}(\Psi_{h_1}^{(\theta_1)})\,\mathrm{ad}(\omega \cdot A_{h_2}^{\mathrm{main},(\theta_2)})\,\mathrm{Ad}(O_{<h_2})|\xi|) \|_{L^\infty L^2[I] \to L^1 L^2[I]} \\ &\lesssim 2^{-\kappa} 2^{\frac{1}{4}(h_2 - h_1)} \| A_{h_1} \|_{S^1} \| A_{h_2} \|_{S^1}, \end{aligned}$$

which is acceptable. On the other hand, in the large angle case $\theta_2 \geq 2^{-2\kappa}$, we use Lemma 10.7 to bound

$$2^{-\frac{1}{2}h_2} \theta_2^{-\frac{3}{2}} \| \omega \cdot A_{h_2}^{\mathrm{main},(\theta_2)} \|_{DL^2 L^\infty[I]} \lesssim 2^{C\kappa} \| A_{h_2} \|_{DS^1[I]}.$$

When $2^{h_1}\theta_1^2 < 2^{-2\kappa} 2^{h_2} \theta_2^2$, we extend the input to $\mathbb{R} \times \mathbb{R}^4$ by zero outside I and use modulation localization. Here we do not apply Lemma 10.7, but rather gain smallness from $-\kappa$. In this case, observe that it is impossible for the input, the output and $\Psi_{h_1}^{(\theta_1)}$ to all have modulation $\ll 2^{h_2}\theta_2^2 =: j_2$. Therefore, we split into three cases:

Case 1: high-modulation input. We estimate

$$\begin{aligned}\sum_{\theta_2}\sum_{\theta_1<2^{-\kappa/2}2^{(1/2)(h_2-h_1)}\theta_2} &\|\mathrm{Op}(\mathrm{ad}(\Psi_{h_1}^{(\theta_1)})\,\mathrm{ad}(\omega\cdot A_{h_2}^{\mathrm{main},(\theta_2)}))Q_{\geq j_2-C}\|_{X_0^{1/2,\infty}\to L^1L^2}\\ &\lesssim \sum_{\theta_2}\sum_{\theta_1<2^{-\kappa/2}2^{(1/2)(h_2-h_1)}\theta_2} 2^{-\frac{1}{2}h_2}\theta_2^{-1}\|\Psi_{h_1}^{(\theta_1)}\|_{DL^6L^\infty}\|\omega\cdot A_{h_2}^{\mathrm{main},(\theta_2)}\|_{DL^3L^\infty}\\ &\lesssim \sum_{\theta_2}\sum_{\theta_1<2^{-\kappa/2}2^{(1/2)(h_2-h_1)}\theta_2} 2^{\frac{1}{6}(h_2-h_1)}\theta_1^{\frac{1}{6}}\theta_2^{\frac{5}{6}}\|A_{x,h_1}\|_{S^1}\|A_{x,h_2}\|_{S^1}\\ &\lesssim 2^{-\frac{1}{6}\kappa}2^{\frac{1}{4}(h_2-h_1)}\|A_{x,h_1}\|_{S^1}\|A_{x,h_2}\|_{S^1},\end{aligned}$$

which is acceptable.

Case 2: high-modulation output. When the output has modulation $\geq 2^{j_2-C}$, then we have exactly the same bound for $L^\infty L^2\to X_0^{-1/2,1}$ (we use boundedness of $Q_{<j_2-C}$ on $L^\infty L^2$).

Case 3: high modulation for Ψ_{h_1}. By boundedness of $Q_{<j_2-C}$ on $L^\infty L^2$ and L^1L^2, it suffices to have the estimate

$$\begin{aligned}\sum_{\theta_2}\sum_{\theta_1<2^{-\kappa/2}2^{(1/2)(h_2-h_1)}\theta_2} &\|\mathrm{Op}(\mathrm{ad}(Q_{\geq j_2-C}\Psi_{h_1}^{(\theta_1)})\,\mathrm{ad}(\omega\cdot A_{h_2}^{\mathrm{main},(\theta_2)}))\|_{L^\infty L^2\to L^1L^2}\\ &\lesssim \sum_{\theta_2}\sum_{\theta_1<2^{-\kappa/2}2^{(1/2)(h_2-h_1)}\theta_2} \|Q_{\geq j_2-C}\Psi_{h_1}^{(\theta_1)}\|_{DL^2L^\infty}\|\omega\cdot A_{h_2}^{\mathrm{main},(\theta_2)}\|_{DL^2L^\infty}\\ &\lesssim \sum_{\theta_2}\sum_{\theta_1<2^{-\kappa/2}2^{(1/2)(h_2-h_1)}\theta_2} \theta_1^{\frac{1}{2}}\theta_2^{\frac{1}{2}}\|A_{x,h_1}\|_{S^1}\|A_{x,h_2}\|_{S^1}\\ &\lesssim 2^{-\frac{1}{2}\kappa}2^{\frac{1}{4}(h_2-h_1)}\|A_{x,h_1}\|_{S^1}\|A_{x,h_2}\|_{S^1}.\end{aligned}$$

Here, we have use (10-15) for $\sum_{j\geq j_2-C}Q_j\Psi_{h_1}^{(\theta_1)}$.

11D3. *Estimate for higher-order expressions.* The contribution of the cubic, quartic and quintic terms in Ψ in the expansion of $O_{;\alpha}$ are treated in a similar manner as in the quadratic case; therefore, we omit the proof. The only remaining case is the contribution of (11-21). For this term, we claim that

$$\left\|\int_{-\infty}^{-\kappa}\int_{-\infty}^{h_1}\cdots\int_{-\infty}^{h_5}\mathrm{Op}(\mathrm{ad}(\Psi_{h_1})\cdots\mathrm{ad}(\Psi_{h_5})\,\mathrm{ad}(O_{<h_6;\alpha})\,\mathrm{Ad}(O_{<h_6}))\,dh_6\cdots dh_2\,dh_1\right\|_{N^*[I]\to N[I]}\leq\varepsilon$$

for κ_1 large enough and δ_p in (9-3) adequately small.

As in the case of the quadratic part, we start with very short intervals and move up the line. If $|I|<2^{-h_1-C\kappa}$ then we only apply fixed-time decomposability estimates, namely (10-14) with $q=\infty$ and $r=\infty$ and (10-17) also with $q=\infty$, together with Hölder in time, to obtain

$$\big\|\mathrm{Op}\big(\mathrm{ad}(\Psi_{h_1})\cdots\mathrm{ad}(\Psi_{h_5})\,\mathrm{ad}(O_{<h_6;\alpha})\,\mathrm{Ad}(O_{<h_6})\big)\big\|_{L^\infty L^2\to L^1L^2}\lesssim_M 2^{h_6}|I|,$$

which suffices for the h integration.

If $2^{-h_1-C\kappa} \leq |I| < 2^{-h_2-C\kappa}$ then we switch to (10-14) with $q=6$ and $r=\infty$ for Ψ_{h_1}, to obtain

$$\big\|\mathrm{Op}\big(\mathrm{ad}(\Psi_{h_1})\cdots\mathrm{ad}(\Psi_{h_5})\,\mathrm{ad}(O_{<h_6;\alpha})\,\mathrm{Ad}(O_{<h_6})\big)\big\|_{L^\infty L^2\to L^1L^2} \lesssim_M 2^{-\frac{1}{6}h_1}2^{h_6}|I|^{\frac{5}{6}},$$

which again suffices for the h integration.

Repeating this procedure for increasingly large I we eventually arrive at the last case $|I| > 2^{-h_6-C\kappa}$. There by Lemma 10.3 and boundedness of $\mathrm{Ad}(O_{<h_6})$ on L^2, we have

$$\begin{aligned}\big\|\mathrm{Op}\big(\mathrm{ad}(\Psi_{h_1})\cdots\mathrm{ad}(\Psi_{h_5})\,\mathrm{ad}(O_{<h_6;\alpha})\,\mathrm{Ad}(O_{<h_2})\big)\big\|_{L^\infty L^2[I]\to L^1L^2[I]}&\\ \lesssim \|\Psi_{h_1}\|_{DL^6L^\infty[I]}\cdots\|\Psi_{h_5}\|_{DL^6L^\infty[I]}\|O_{<h_6;\alpha}\|_{DL^6L^\infty[I]}&.\end{aligned}$$

Using Lemma 10.5 for $\Psi_h^{(\theta)}$ with $\theta < 2^{-\kappa}$ and Lemma 10.7 for the rest, we have

$$\|\Psi_h\|_{DL^6L^\infty[I]} \leq 2^{-\frac{1}{6}h}(2^{-\kappa}\|A_{x,h}\|_{S^1[I]} + C2^{C\kappa}\|A_{x,h}\|_{DS^1[I]}).$$

This bound provides us with the desired smallness. By the previous estimate and (10-17), the h-integrals converge as well, which proves our claim.

11E. *Estimates for E_3, …, E_6.* We finally handle the error terms E_3, …, E_6, for which we gain smallness from the frequency gap κ.

11E1. *The estimate for E_3.* It suffices to show that

$$\|E_3P_0\|_{L^\infty L^2\to L^1L^2} \lesssim_M 2^{-\frac{1}{2}\kappa}.$$

But this is a consequence of the L^2 boundedness for $\mathrm{Op}(\mathrm{Ad}(O))$, combined with the L^2L^∞ decomposability estimates for A_α and $O_{;\alpha}$ in Lemmas 10.4 and 10.6.

11E2. *The estimate for E_4.* We expand with respect to h,

$$\mathrm{ad}(\partial^\alpha O_{;\alpha})\,\mathrm{Ad}(O) = \int_{-\infty}^{-\kappa} \partial^\alpha(\mathrm{ad}(O_{<h;\alpha})\,\mathrm{ad}(\Psi_h))\,\mathrm{Ad}(O_{<h})\,\mathrm{ad}(\Box\Psi_h)\,\mathrm{Ad}(O_{<h})\,dh.$$

For the first term we simply use two L^2L^∞ decomposability estimates as in the case of E_3. For the second term, in view of the bound (10-16), we can apply Lemma 11.3 to discard the $\mathrm{Ad}(O_{<h})$ factor. Then it suffices to show that

$$\left\|\int_{-\infty}^{-\kappa}\mathrm{Op}(\mathrm{ad}(\Box\Psi_h))P_0\,dh\right\|_{S_0\to N} \lesssim_M 2^h.$$

After expanding Ψ_h in θ, we note that, due to the frequency localization of $\Psi_h^{(\theta)}$, either the input or the output has modulation $\gtrsim 2^h\theta^2$. We assume the former, as the other case is similar. Then we only need to prove the bound

$$\left\|\int_{-\infty}^{-\kappa}\mathrm{Op}(\mathrm{ad}(\Box\Psi_h^{(\theta)}))P_0\,dh\right\|_{L^2\to L^1L^2} \lesssim_M \theta 2^{\frac{3}{2}h},$$

which is an immediate consequence of the decomposability bound (10-16) for $\Box\Psi_h^{(\theta)}$.

11E3. *The estimate for E_5.* It suffices to show that

$$\|E_3 P_0\|_{S_0^\sharp \to L^1 L^2} \lesssim_M 2^{-\frac{1}{2}\kappa}.$$

Since $(D_t + |D_x|)P_0 : S_0^\sharp \to L^2$, this follows from the L^2 boundedness for $\mathrm{Op}(\mathrm{Ad}(O))$, combined with the $L^2 L^\infty$ decomposability estimates for A_α in Lemma 10.4.

11E4. *The estimate for E_6.* In view of the $L^2 L^\infty$ decomposability estimates for A_α in Lemmas 10.4 and 11.3, we can discard the $\mathrm{Ad}(O)$ factor. In addition, as in Proposition 4.30, we can express the commutator $[S_0, A_h]$ in the form

$$[S_0, A_h] f = 2^h \mathcal{O}(A_h, f).$$

Then we have reduced our problem to proving

$$\left\| \int_{-\infty}^{-\kappa} 2^h \,\mathrm{Op}(\mathrm{ad}(\omega \cdot \nabla A_{x,h})) P_0 \, dh \right\|_{S_0 \to N} \ll \varepsilon,$$

$$\left\| \int_{-\infty}^{-\kappa} 2^h \,\mathrm{Op}(\mathrm{ad}(A_{0,h})) P_0 \, dh \right\|_{S_0 \to N} \ll \varepsilon.$$

But then these follow, with the $2^{-\delta_1 \kappa}$ gain, from (8-21) and (8-23), thanks to the extra derivative (i.e., the 2^h factor).

Acknowledgments

Part of the work was carried out during the semester program "New Challenges in PDE" held at MSRI in Fall 2015. Oh was supported by the Miller Research Fellowship from the Miller Institute, UC Berkeley and the TJ Park Science Fellowship from the POSCO TJ Park Foundation. Tataru was partially supported by the NSF grant DMS-1266182 as well as by the Simons Investigator grant from the Simons Foundation.

References

[Gavrus 2019] C. Gavrus, "Global well-posedness for the massive Maxwell–Klein–Gordon equation with small critical Sobolev data", *Ann. PDE* **5**:1 (2019), art. id. 10. MR Zbl

[Gavrus and Oh 2016] C. Gavrus and S.-J. Oh, "Global well-posedness of high dimensional Maxwell–Dirac for small critical data", preprint, 2016. arXiv

[Kenig and Merle 2006] C. E. Kenig and F. Merle, "Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear Schrödinger equation in the radial case", *Invent. Math.* **166**:3 (2006), 645–675. MR Zbl

[Kenig and Merle 2008] C. E. Kenig and F. Merle, "Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation", *Acta Math.* **201**:2 (2008), 147–212. MR Zbl

[Klainerman and Machedon 1994] S. Klainerman and M. Machedon, "On the Maxwell–Klein–Gordon equation with finite energy", *Duke Math. J.* **74**:1 (1994), 19–44. MR Zbl

[Klainerman and Machedon 1995] S. Klainerman and M. Machedon, "Finite energy solutions of the Yang–Mills equations in $\mathbb{R}^{3+1}$", *Ann. of Math.* (2) **142**:1 (1995), 39–119. MR Zbl

[Krieger and Lührmann 2015] J. Krieger and J. Lührmann, "Concentration compactness for the critical Maxwell–Klein–Gordon equation", *Ann. PDE* **1**:1 (2015), art. id. 5. MR Zbl

[Krieger and Schlag 2012] J. Krieger and W. Schlag, *Concentration compactness for critical wave maps*, European Mathematical Society, Zürich, 2012. MR Zbl

[Krieger and Sterbenz 2013] J. Krieger and J. Sterbenz, *Global regularity for the Yang–Mills equations on high dimensional Minkowski space*, Mem. Amer. Math. Soc. **1047**, Amer. Math. Soc., Providence, RI, 2013. MR Zbl

[Krieger and Tataru 2017] J. Krieger and D. Tataru, "Global well-posedness for the Yang–Mills equation in 4+1 dimensions: small energy", *Ann. of Math.* (2) **185**:3 (2017), 831–893. MR Zbl

[Krieger et al. 2015] J. Krieger, J. Sterbenz, and D. Tataru, "Global well-posedness for the Maxwell–Klein–Gordon equation in 4+1 dimensions: small energy", *Duke Math. J.* **164**:6 (2015), 973–1040. MR Zbl

[Machedon and Sterbenz 2004] M. Machedon and J. Sterbenz, "Almost optimal local well-posedness for the (3+1)-dimensional Maxwell–Klein–Gordon equations", *J. Amer. Math. Soc.* **17**:2 (2004), 297–359. MR Zbl

[Oh 2014] S.-J. Oh, "Gauge choice for the Yang–Mills equations using the Yang–Mills heat flow and local well-posedness in H^1", *J. Hyperbolic Differ. Equ.* **11**:1 (2014), 1–108. MR Zbl

[Oh 2015] S.-J. Oh, "Finite energy global well-posedness of the Yang–Mills equations on $\mathbb{R}^{1+3}$: an approach using the Yang–Mills heat flow", *Duke Math. J.* **164**:9 (2015), 1669–1732. MR Zbl

[Oh and Tataru 2016a] S.-J. Oh and D. Tataru, "Global well-posedness and scattering of the (4+1)-dimensional Maxwell–Klein–Gordon equation", *Invent. Math.* **205**:3 (2016), 781–877. MR Zbl

[Oh and Tataru 2016b] S.-J. Oh and D. Tataru, "Local well-posedness of the (4+1)-dimensional Maxwell–Klein–Gordon equation at energy regularity", *Ann. PDE* **2**:1 (2016), art. id. 2. MR Zbl

[Oh and Tataru 2017a] S.-J. Oh and D. Tataru, "The threshold conjecture for the energy critical hyperbolic Yang–Mills equation", preprint, 2017. arXiv

[Oh and Tataru 2017b] S.-J. Oh and D. Tataru, "The Yang–Mills heat flow and the caloric gauge", preprint, 2017. arXiv

[Oh and Tataru 2018] S.-J. Oh and D. Tataru, "Energy dispersed solutions for the (4+1)-dimensional Maxwell–Klein–Gordon equation", *Amer. J. Math.* **140**:1 (2018), 1–82. MR Zbl

[Oh and Tataru 2019a] S.-J. Oh and D. Tataru, "The hyperbolic Yang–Mills equation for connections in an arbitrary topological class", *Comm. Math. Phys.* **365**:2 (2019), 685–739. MR Zbl

[Oh and Tataru 2019b] S.-J. Oh and D. Tataru, "The threshold theorem for the (4+1)-dimensional Yang–Mills equation: an overview of the proof", *Bull. Amer. Math. Soc.* (*N.S.*) **56**:2 (2019), 171–210. MR Zbl

[Rodnianski and Tao 2004] I. Rodnianski and T. Tao, "Global regularity for the Maxwell–Klein–Gordon equation with small critical Sobolev norm in high dimensions", *Comm. Math. Phys.* **251**:2 (2004), 377–426. MR Zbl

[Sterbenz and Tataru 2010a] J. Sterbenz and D. Tataru, "Energy dispersed large data wave maps in 2+1 dimensions", *Comm. Math. Phys.* **298**:1 (2010), 139–230. MR Zbl

[Sterbenz and Tataru 2010b] J. Sterbenz and D. Tataru, "Regularity of wave-maps in dimension 2+1", *Comm. Math. Phys.* **298**:1 (2010), 231–264. MR Zbl

[Tao 2001] T. Tao, "Global regularity of wave maps, II: Small energy in two dimensions", *Comm. Math. Phys.* **224**:2 (2001), 443–544. MR Zbl

[Tao 2004] T. Tao, "Geometric renormalization of large energy wave maps", exposé XI in *Journées "Équations aux Dérivées Partielles"*, École Polytech., Palaiseau, 2004. MR Zbl

[Tao 2008a] T. Tao, "Global regularity of wave maps, III: Large energy from $\mathbb{R}^{1+2}$ to hyperbolic spaces", preprint, 2008. arXiv

[Tao 2008b] T. Tao, "Global regularity of wave maps, IV: Absence of stationary or self-similar solutions in the energy class", preprint, 2008. arXiv

[Tao 2008c] T. Tao, "Global regularity of wave maps, V: Large data local wellposedness and perturbation theory in the energy class", preprint, 2008. arXiv

[Tao 2009a] T. Tao, "Global regularity of wave maps, VI: Abstract theory of minimal-energy blowup solutions", preprint, 2009. arXiv

[Tao 2009b] T. Tao, "Global regularity of wave maps, VII: Control of delocalised or dispersed solutions", preprint, 2009. arXiv

[Tataru 2001] D. Tataru, "On global existence and scattering for the wave maps equation", *Amer. J. Math.* **123**:1 (2001), 37–77. MR Zbl

Received 23 Nov 2018. Revised 29 Jan 2020. Accepted 8 Mar 2020.

SUNG-JIN OH: sjoh@math.berkeley.edu
Department of Mathematics, University of California, Berkeley, CA, United States
School of Mathematics, Korea Institute for Advanced Study, Seoul, South Korea

DANIEL TATARU: tataru@math.berkeley.edu
Department of Mathematics, University of California, Berkeley, CA, United States

PURE and APPLIED
ANALYSIS
Vol. 2, No. 2, 2020
https://doi.org/10.2140/paa.2020.2.385

SCATTERING RESONANCES ON TRUNCATED CONES

DEAN BASKIN AND MENGXUAN YANG

We consider the problem of finding the resonances of the Laplacian on truncated Riemannian cones. In a similar fashion to Cheeger and Taylor, we construct the resolvent and scattering matrix for the Laplacian on cones and truncated cones. Following Stefanov, we show that the resonances on the truncated cone are distributed asymptotically as $Ar^n + o(r^n)$, where A is an explicit coefficient. We also conclude that the Laplacian on a nontruncated cone has no resonances.

1. Introduction

In this note, we consider the resonances on truncated Riemannian cones and establish a Weyl-type formula for their distribution. To fix notation, we let (Y, h) be a compact $(n-1)$-dimensional Riemannian manifold (with or without boundary) and let $C(Y)$ denote the cone over Y. In other words, $C(Y)$ is diffeomorphic to the product $(0, \infty)_r \times Y$ and is equipped with the incomplete Riemannian metric $g = dr^2 + r^2 h$. We refer the reader to the foundational works [Cheeger and Taylor 1982a; 1982b] for more details on the geometric set-up. We also introduce the *truncated* Riemannian cone $C_a(Y)$ formed by introducing a boundary at $r = a$; i.e., $C_a(Y)$ is diffeomorphic to $[a, \infty)_r \times Y$ and equipped with the same metric.

The (negative-definite) Laplacian on $C(Y)$ (or $C_a(Y)$ with a choice of boundary conditions) has the form

$$\partial_r^2 + \frac{n-1}{r}\partial_r + \frac{1}{r^2}\Delta_h,$$

where Δ_h denotes the Laplacian of (Y, h). Its resolvent $R(\lambda)$ is given by

$$R(\lambda) = (\Delta + \lambda^2)^{-1}.$$

We consider the *cutoff resolvent* $\chi R(\lambda)\chi$, where χ is a (fixed) smooth compactly supported function on $C(Y)$ (or $C_a(Y)$). One consequence of the resolvent formula of Theorem 2.1 is that the cutoff resolvent extends meromorphically to the logarithmic cover of $\mathbb{C} \setminus \{0\}$.

More precisely, we identify elements λ of the logarithmic cover of $\mathbb{C} \setminus \{0\}$ by a magnitude $|\lambda|$ and a phase $\arg\lambda \in \mathbb{R}$. We identify the "physical half-plane" as those λ with $\arg\lambda \in (0, \pi)$. These λ correspond to the resolvent set $\mathbb{C} \setminus [0, \infty)$ via the map $\lambda \mapsto |\lambda|^2 e^{2i \arg\lambda}$. The cutoff resolvent then extends to be meromorphic as a function of λ on this logarithmic cover.

MSC2010: 33C10, 35L05, 58J50.
Keywords: resonances, cones.

The poles of the cutoff resolvent consist of possibly finitely many L^2-eigenvalues lying in the upper half-plane (which do not appear with Dirichlet boundary conditions) and poles lying on other sheets of the cover. The latter poles are called the *resonances* of Δ.

The main theorem of this paper counts the most physically relevant resonances for the truncated cone. In particular, we count those resonances λ nearest to the physical half-plane, i.e., those with $\arg\lambda \in \left(-\frac{\pi}{2}, 0\right)$ and $\arg\lambda \in \left(\pi, \frac{3\pi}{2}\right)$. The resonances on other "sheets" of the cover remain more mysterious and are related to the zeros of Hankel functions near the real axis. We consider the resonance counting function on these sheets, defined by

$$N(r) = \#\{\lambda : \lambda \text{ is a resonance and } |\lambda| \le r\}.$$

The following theorem provides an asymptotic formula for $N(r)$.

Theorem 1.1. *Suppose either that the set of periodic geodesics of (Y, h) has Liouville measure zero or that $Y = \mathbb{S}^{n-1}$ equipped with a constant rescaling of the standard metric. Consider the truncated cone $C_1(Y)$ equipped with the Dirichlet Laplacian and let $N(r)$ denote its resonance counting function on the neighboring sheets as above. We then have, as $r \to \infty$,*

$$N(r) = A_n \operatorname{Vol}(Y, h) r^n + o(r^n),$$

where A_n is an explicit constant (defined below in (7)) and $\operatorname{Vol}(Y, h)$ denotes the volume of the Riemannian manifold (Y, h).

The constant $A_n \operatorname{Vol}(Y, h)$ in Theorem 1.1 is the same constant as computed in [Stefanov 2006] for the resonance counting function on the domain exterior to a ball in $\mathbb{R}^n$. When $Y = \mathbb{S}^{n-1}$ is equipped with its standard metric, the truncated cone $C_1(Y)$ can be thought of as the exterior of the unit ball in Euclidean space. Theorem 1.1 recovers Stefanov's result. (When $Y = \mathbb{S}^{n-1}$, n odd, is equipped with its standard metric, the cutoff resolvent in fact continues to the complex plane; this can be seen in the resolvent formulae below.)

We also state the following theorem, which is known to the community but does not seem to be in the literature.

Theorem 1.2. *If (Y, h) is a compact Riemannian manifold (with or without boundary) then the cone $C(Y)$ has no resonances.*

In fact, Theorem 2.1 below shows that λ is a resonance of the truncated cone $C_1(Y)$ if and only if λ/a is a resonance of the truncated cone $C_a(Y)$. Sending a to 0 then pushes all resonances out to infinity and provides evidence for Theorem 1.2.

The proof of Theorem 1.1 has two main steps. We first separate variables and obtain an explicit resolvent formula in Theorem 2.1 to characterize the resonances as zeros of a Hankel function. In Section 3 we consider the asymptotic distribution of the zeros of each Hankel function appearing in the resolvent formula. The hypothesis on the link (Y, h) is used to control the error terms when synthesizing the result. Theorem 1.2 is an immediate corollary of the resolvent formula in Theorem 2.1.

The proof of Theorem 1.1 follows an argument of [Stefanov 2006] very closely. Stefanov established a Weyl-type law for the distribution of resonances for the exterior of a ball in odd-dimensional Euclidean space. The main contribution of this paper is the observation that, after some natural modifications, the core of Stefanov's argument applies to the setting of cones. Borthwick [2010; 2012] and Borthwick and Philipp [2014] showed that a similar approach works in the asymptotically hyperbolic setting.

We further remark that we have specialized to the Dirichlet Laplacian in Theorem 1.1 only for simplicity. For Neumann or Robin boundary conditions, the resolvent formula of Theorem 2.1 has an analogous expression. The resonance counting problem then involves counting zeros of $H_\nu^{(2)'} + C\nu H_\nu^{(2)}$, which can be handled with similar arguments.

2. Resolvent construction

In this section we write down an explicit formula (via separation of variables) for the resolvent and then show that the cut-off resolvent has a meromorphic continuation to the logarithmic cover Λ of the complex plane. The construction is essentially contained in the [Cheeger and Taylor 1982a; 1982b], but the resolvent is not explicitly written there.

Suppose ϕ_j form an orthonormal family of eigenfunctions for $-\Delta_h$ with corresponding eigenvalues μ_j^2. We decompose $L^2(C(Y))$ into a direct sum in terms of the eigenspaces of $-\Delta_h$, i.e.,

$$L^2(C_a(Y); \mathbb{C}) = \bigoplus_{j=0}^{\infty} L^2((a,\infty); E_j), \qquad f(r,y) = \sum_{j=0}^{\infty} f_j(r)\phi_j(y),$$

where the first space is defined with respect to the volume form induced by the metric and the latter spaces can be identified (via the identification $f(r)\phi_j(y) \mapsto f(r)$) with the space $L^2((a,\infty); \mathbb{C})$ equipped with the volume form $r^{n-1}\,dr$.

For $\arg\lambda \in (0,\pi)$, the resolvent $R(\lambda)$ splits as a direct sum of operators $R_j(\lambda)$ acting on $L^2((a,\infty), E_j)$, with measure $r^{n-1}\,dr$.

$$R(\lambda)\left(\sum_{j=1}^{\infty} f_j(r)\phi_j(y)\right) = \bigoplus_{j=1}^{\infty} (R_j(\lambda) f_j)\phi_j(y).$$

In this section, we prove the following explicit formula for the j-th piece of the resolvent. For the cone $C(Y)$ (i.e., for $a = 0$), we use the Friedrichs extension of the Laplacian to guarantee self-adjointness (though in high enough dimension the Laplacian is essentially self-adjoint):

Theorem 2.1. *The piece of the resolvent corresponding to the j-th eigenvalue has the following explicit expression on the truncated cone $C_a(Y)$ or the cone $C(Y)$ $(a = 0)$:*

$$(R_j(\lambda)f)(r) = \int_a^\infty K_{a,j}(r,\tilde{r}) f(\tilde{r})\tilde{r}^{n-1}\,d\tilde{r},$$

where $K_{a,j}(r,\tilde{r})$ is given by

$$K_{a,j}(r,\tilde{r}) = \frac{\pi}{2i}(\tilde{r}r)^{-(n-2)/2} \begin{cases} H_{\nu_j}^{(1)}(\lambda\tilde{r})J_{\nu_j}(\lambda r) - (J_{\nu_j}(\lambda a)/H_{\nu_j}^{(1)}(\lambda a))H_{\nu_j}^{(1)}(\lambda\tilde{r})H_{\nu_j}^{(1)}(\lambda r), & r < \tilde{r}, \\ J_{\nu_j}(\lambda\tilde{r})H_{\nu_j}^{(1)}(\lambda r) - (J_{\nu_j}(\lambda a)/H_{\nu_j}^{(1)}(\lambda a))H_{\nu_j}^{(1)}(\lambda\tilde{r})H_{\nu_j}^{(1)}(\lambda r), & r > \tilde{r}. \end{cases}$$

Here J_ν are the standard Bessel functions of the first kind and $H_\nu^{(1)}$ are the Hankel functions of the first kind. The second term in both expressions should be interpreted as 0 *when* $a = 0$.

Proof. After separating variables, we may assume that $f = f_j(r)\phi_j(y)$. We construct the resolvent for $\Im\lambda > 0$ and then meromorphically continue the expression.

Writing $u = u_j(r)\phi_j(y)$, the equation $(\Delta+\lambda^2)u = f$ induces the following differential equation for u_j:

$$\partial_r^2 u_j + \frac{n-1}{r}\partial_r u_j - \frac{\mu_j^2}{r^2}u_j + \lambda^2 u_j = f_j. \tag{1}$$

We solve this equation by showing it is equivalent to a Bessel equation.

Changing variables to $\rho = \lambda r$ and writing $\tilde{u}(\rho) = u(\rho/\lambda)$ yields

$$\partial_\rho^2 \tilde{u} + \frac{n-1}{\rho}\partial_\rho \tilde{u} + \left(1 - \frac{\mu_j^2}{\rho^2}\right)\tilde{u} = \frac{1}{\lambda^2}\tilde{f}(\rho).$$

Writing $v = \rho^{(n-2)/2}\tilde{u}$, we obtain a Bessel equation for v:

$$v'' + \frac{1}{\rho}v' + \left(1 - \frac{\nu_j^2}{\rho^2}\right)v = g(\rho), \tag{2}$$

where

$$\nu_j^2 = \mu_j^2 + ((n-2)/2)^2 \quad \text{and} \quad g(\rho) = \frac{\rho^{(n-2)/2}}{\lambda^2}\tilde{f}(\rho).$$

We now proceed by the standard ODE technique of variation of parameters. One basis for the space of solutions of the homogeneous version of this Bessel equation is $\{J_{\nu_j}(\rho), H_{\nu_j}^{(1)}(\rho)\}$, where J_ν is the Bessel function of the first kind and $H_\nu^{(1)}$ is the Hankel function of the first kind. We thus may use the following basis for the space of solutions of the homogeneous equation:

$$w_1(r) = r^{-(n-2)/2}J_{\nu_j}(\lambda r), \quad w_2(r) = r^{-(n-2)/2}H_{\nu_j}^{(1)}(\lambda r). \tag{3}$$

For $\Im\lambda > 0$, $R_j(\lambda) f_j$ must lie in $L^2((a,\infty), r^{n-1}\,dr)$. If f_j is compactly supported, this means that $u_j = R_j(\lambda) f_j$ must be a multiple of $r^{-(n-2)/2}H_{\nu_j}^{(1)}(\lambda r)$ near infinity. When $a > 0$, u_j must satisfy the boundary condition at $r = a$. When $a = 0$, the choice of the Friedrichs extension requires that both u_j and u_j' lie in the weighted L^2 space near 0 and so u_j must be a multiple of $r^{-(n-2)/2}J_{\nu_j}(\lambda r)$ near $r = 0$ as any nonzero multiple of w_2 will not have this property.

We may thus write

$$u_j(r) = \left(\int_r^\infty \frac{w_2(\tilde{r}) f_j(\tilde{r})}{W(w_1,w_2)(\tilde{r})}\,d\tilde{r}\right)w_1(r) + \left(C + \int_a^r \frac{w_1(\tilde{r}) f_j(\tilde{r})}{W(w_1,w_2)(\tilde{r})}\,d\tilde{r}\right)w_2(r),$$

where C is a yet-to-be-determined constant, the functions w_1 and w_2 are as in (3), and $W(w_1, w_2)$ is their Wronskian. The Wronskian W can be easily computed in terms of the Wronskian of the Bessel and Hankel functions and is

$$W(w_1,w_2)(r) = r^{-(n-1)}\frac{2i}{\pi}.$$

We now turn our attention to the boundary condition. For $a = 0$, the requirement that the solution and its derivative live in L^2 forces $C = 0$, yielding the result. For $a \neq 0$, we require that $u_j(a) = 0$; i.e.,

$$\left(\frac{\pi}{2i}\int_a^\infty H^{(1)}_{\nu_j}(\lambda\tilde{r})\tilde{r}^{n/2} f(\tilde{r})\,d\tilde{r}\right) a^{-(n-2)/2} J_{\nu_j}(\lambda a) + C a^{-(n-2)/2} H^{(1)}_{\nu_j}(\lambda a) = 0,$$

and so we must have

$$C = -\frac{\pi}{2i}\frac{J_{\nu_j}(\lambda a)}{H^{(1)}_{\nu_j}(\lambda a)}\int_a^\infty H^{(1)}_{\nu_j}(\lambda\tilde{r})\tilde{r}^{n/2} f(x)\,dx,$$

finishing the proof. □

We now claim that $\chi R(\lambda)\chi$ has a meromorphic continuation:

Lemma 2.2. *Given a fixed $\chi \in C_c^\infty(\mathbb{R}_+ \times Y)$, $\chi R(\lambda)\chi$ meromorphically continues from*

$$\{\lambda \in \mathbb{C} : \Im\lambda > 0\}$$

to the logarithmic cover Λ of the complex plane.

Proof. We first prove the statement for the full cone; the statement for the truncated cone will follow by an appeal to the analytic Fredholm theorem.

Fix $\chi \in C_c^\infty((0, \infty))$ and regard $\chi(r)$ as a compactly supported smooth function on $C(Y)$. We let $R(\lambda)$ denote the resolvent on the nontruncated cone (i.e., $a = 0$) and $K(\lambda; r, y, \tilde{r}, \tilde{y})$ denote its integral kernel. In order to show that $\chi R(\lambda)\chi$ meromorphically continues, it suffices to show that for any $f, g \in L^2(C(Y))$, the function

$$\lambda \mapsto \langle \chi R(\lambda)\chi f, g\rangle$$

meromorphically continues to Λ.

Fix two such functions $f, g \in L^2(C(Y))$ and let $f_j(r)$ and $g_j(r)$ denote their coefficients in the expansion in terms of eigenfunctions of Δ_h, i.e.,

$$f(r, y) = \sum_{j=0}^\infty f_j(r)\phi_j(y).$$

We observe that because f and g are square-integrable, the sum and the integral commute; i.e.,

$$\|f\|^2_{L^2(C(Y))} = \int_0^\infty \sum_{j=0}^\infty |f_j(r)|^2 r^{n-1}\,dr = \sum_{j=0}^\infty \int_0^\infty |f_j(r)|^2 r^{n-1}\,dr.$$

From Theorem 2.1, we may write

$$\begin{aligned}\langle \chi R(\lambda)\chi f, g\rangle = \sum_{j=0}^\infty \Bigg(&\int_0^\infty\int_0^r (\tilde{r}r)^{-(n-2)/2}\chi(r)\chi(\tilde{r}) f_j(\tilde{r}) g_j(r) J_{\nu_j}(\lambda\tilde{r}) H^{(1)}_{\nu_j}(\lambda r)\tilde{r}^{n-1} r^{n-1}\,d\tilde{r}\,dr \\ &+ \int_0^\infty\int_r^\infty (\tilde{r}r)^{-(n-2)/2}\chi(r)\chi(\tilde{r}) f_j(\tilde{r}) g_j(r) J_{\nu_j}(\lambda r) H^{(1)}_{\nu_j}(\lambda\tilde{r})\tilde{r}^{n-1} r^{n-1}\,d\tilde{r}\,dr\Bigg), \quad (4)\end{aligned}$$

where J_ν and $H_\nu^{(1)}$ are as above. Because each term in (4) meromorphically continues to the Riemann surface Λ, it suffices to show that the partial sums of the series converge locally (in λ) uniformly (in j).

By the asymptotic expansions of Bessel functions for large order, we know [DLMF 2018, 10.19] that, locally in $\lambda \in \Lambda$, and for $r \in \operatorname{supp} \chi$,

$$J_\nu(\lambda r) = \frac{1}{\sqrt{2\pi\nu}}\left(\frac{e\lambda r}{2\nu}\right)^{\nu} + o\left(\frac{1}{\sqrt{\nu}}\left(\frac{e\lambda r}{2\nu}\right)^{\nu}\right),$$
$$H_\nu^{(1)}(\lambda r) = \frac{1}{i}\sqrt{\frac{2}{\pi\nu}}\left(\frac{e\lambda r}{2\nu}\right)^{-\nu} + o\left(\frac{1}{\sqrt{\nu}}\left(\frac{e\lambda r}{2\nu}\right)^{-\nu}\right),$$

as $\nu \to \infty$ through the positive reals. In particular, for j large enough, each term in (4) can be bounded by

$$C\int_0^\infty\int_0^r \frac{1}{\pi\nu_j}\chi(r)\chi(\tilde{r}) f_j(\tilde{r}) g_j(r)\left[\left(\frac{\tilde{r}}{r}\right)^{\nu_j}(1+o(1))\right](\tilde{r}r)^{n/2}\,d\tilde{r}\,dr$$
$$+ C\int_0^\infty\int_r^\infty \frac{1}{\pi\nu_j}\chi(r)\chi(\tilde{r}) f_j(\tilde{r}) g_j(r)\left[\left(\frac{r}{\tilde{r}}\right)^{\nu_j}(1+o(1))\right](\tilde{r}r)^{n/2}\,d\tilde{r}\,dr.$$

Observe that in the first integral, $\tilde{r}/r$ is bounded by 1, while $r/\tilde{r}$ is bounded by 1 in the second.

Because χ is compactly supported, we may therefore bound each term (for j large enough) by

$$\frac{C_\chi}{\nu_j}\|f_j\|_{L^2}\|g_j\|_{L^2}.$$

This sequence is absolutely summable, so the partial sums of the series in (4) converge locally uniformly. This establishes that the cut-off resolvent on the full cone ($a = 0$) meromorphically extends to the logarithmic cover Λ of the complex plane.

We now proceed to the case of the truncated cone ($a > 0$). We proceed by an appeal to the analytic Fredholm theorem.

Fix $\chi_0, \chi_\infty \in C^\infty((a, \infty))$ so that $\chi_0(r)$ is supported near $r = a$, $\chi_\infty(r)$ is identically zero near $r = a$, and $\chi_0 + \chi_\infty = 1$. We let $R_\infty(\lambda)$ denote the resolvent on the nontruncated cone and $R_0(\lambda)$ denote the resolvent on a compact manifold with boundary into which the support of χ_0 embeds isometrically. We define the parametrix

$$Q(\lambda) = \tilde{\chi}_0 R_0(\lambda)\chi_0 + \tilde{\chi}_\infty R_\infty(\lambda)\chi_\infty,$$

where $\tilde{\chi}$ have similar support properties and are identically 1 on the support of their counterparts. Applying $\Delta + \lambda^2$ yields a remainder of the form $I + \sum[\Delta, \tilde{\chi}_i]R_i(\lambda)\chi_i$. Both terms are compact and the operator is invertible for large $\Im\lambda$ by Neumann series, so applying $R_a(\lambda)$ to both sides and inverting the remainder shows that it has a meromorphic continuation. □

3. Proof of Theorem 1.1

By the formula for the resolvent in Theorem 2.1, the resonances of $R_a(\lambda)$ correspond to those λ for which $H_{\nu_j}^{(1)}(\lambda a) = 0$ for some j. For simplicity we will discuss only the case $a = 1$ as the other cases can be found by rescaling. As mentioned in the introduction, we consider only those resonances nearest to the upper half-plane, i.e., those with

$$-\tfrac{\pi}{2} < \arg\lambda < 0 \quad \text{or} \quad \pi < \arg\lambda < \tfrac{3\pi}{2}. \tag{5}$$

Because ν_j is real, we may relate the zeros of $H^{(1)}_{\nu_j}(\lambda)$ in the region given by (5) to zeros of $H^{(2)}_{\nu_j}(\lambda)$ in the quadrant $0 < \arg\lambda < \frac{\pi}{2}$ via analytic continuation formulae. Indeed, it is well known [DLMF 2018, 10.11.5, 10.11.9] that

$$\begin{aligned} H^{(1)}_{\nu}(ze^{\pi}) &= -e^{-\nu\pi\imath} H^{(2)}_{\nu}(z), \\ H^{(1)}_{\nu}(\bar{z}) &= \overline{H^{(2)}_{\nu}(z)}. \end{aligned} \tag{6}$$

The first of these equations identifies zeros of $H^{(1)}_{\nu}$ in $\pi < \arg\lambda < \frac{3\pi}{2}$ to zeros of $H^{(2)}_{\nu}$ in the first quadrant; the second equation does the same for zeros of $H^{(1)}_{\nu}$ with $-\frac{\pi}{2} < \arg\lambda < 0$. In particular, each zero of $H^{(2)}_{\nu}$ with $0 \le \arg\lambda \le \frac{\pi}{2}$ corresponds to exactly two resonances.

For large enough ν, the zeros of the Hankel function $H^{(2)}_{\nu}$ in the first quadrant lie near the boundary of (a scaling of) an "eye-like" domain $K \subset \mathbb{C}$. The domain K is symmetric about the real axis and is bounded by the following curve and its conjugate:

$$z = \pm(t\coth t - t^2)^{1/2} + i(t^2 - t\tanh t)^{1/2}, \quad 0 \le t \le t_0,$$

where t_0 is the positive root of $t = \coth t$. We refer to the piece of the boundary of K lying in the upper half-plane by ∂K_+.

The constant A_n above is given by

$$A_n = \frac{2(n-1)\,\mathrm{Vol}(B_{n-1})}{n(2\pi)^n} \int_{\partial K_+} \frac{|1-z^2|^{1/2}}{|z|^{n+1}}\, d|z|, \tag{7}$$

where B_{n-1} is the $(n-1)$-dimensional unit ball. Observe that, up to a factor of the volume of the unit sphere (which is replaced by the volume of Y in the theorem statement), the constant A_n is the same constant computed in [Stefanov 2006].

We use below two different parametrizations of the piece of ∂K_+ lying the in the quadrant $0 \le \arg z \le \frac{\pi}{2}$. The first parametrization is by the argument of z, i.e., by the map

$$\left[0, \tfrac{\pi}{2}\right] \to \partial K_+, \quad \theta = \arg z \mapsto z = z(\theta).$$

For the second parametrization, we introduce the function ρ, defined by

$$\rho(z) = \tfrac{2}{3}\zeta^{3/2} = \log\frac{1+\sqrt{1-z^2}}{z} - \sqrt{1-z^2}, \quad |\arg z| < \pi, \tag{8}$$

where (following [Stefanov 2006, Section 4; Olver 1974, Chapter 10]) the branches of the functions above are chosen so that ζ is real when z is. Another characterization is that the principal branches are chosen when $0 < z < 1$ and continuity is demanded elsewhere.

The boundary ∂K is the vanishing set of $\Re\rho$. This yields a parametrization of the part of ∂K_+ lying in $0 \le \arg z \le \frac{\pi}{2}$:

$$\left[0, \tfrac{\pi}{2}\right] \to \partial K_+, \quad t \mapsto \rho^{-1}(-it) = z.$$

The transition between the two parametrizations is given by

$$\frac{dt}{d\theta} = \frac{dt}{dz}\frac{dz}{d\theta} = (i\rho'(z))(iz) = \sqrt{1-z^2}.$$

The function ζ defined in (8) is the solution of the ODE

$$\left(\frac{d\zeta}{dz}\right)^2 = \frac{1-z^2}{\zeta z^2}$$

that is infinitely differentiable on the positive real axis (including at $z = 1$). As is implicit in (8), it can be analytically continued to the complex plane with a branch cut along the negative real axis.

Because the resonances correspond to zeros of $H^{(2)}_{\nu_j}$, we must also consider the asymptotic distribution of the ν_j. In what follows, we consider only the case when the periodic geodesics of (Y, h) have measure zero.[1] The eigenvalues μ_j^2 of Δ_h obey Weyl's law:

$$\begin{aligned} N_h(\mu) &= \#\{\mu_j : \mu_j \le \mu \text{ with multiplicity}\} \\ &= \frac{\operatorname{Vol} B_{n-1}}{(2\pi)^{n-1}} \operatorname{Vol}(Y, h)\mu^{n-1} + R(\mu). \end{aligned}$$

Here $\operatorname{Vol}(B_{n-1})$ denotes the volume of the unit ball in $\mathbb{R}^{n-1}$ and $\operatorname{Vol}(Y, h)$ is the volume of Y equipped with the metric h. In general, $R(\mu) = O(\mu^{n-2})$, but if we now impose the dynamical hypothesis (that the set of periodic geodesics of (Y, h) has Liouville measure zero), then a theorem of [Duistermaat and Guillemin 1975] (in the boundaryless case) and [Ivrii 1980; 1982] (in the boundary case) shows that

$$R(\lambda) = o(\mu^{n-2}).$$

The nonperiodicity assumption then allows us to count eigenvalues on intervals of length 1:

$$\begin{aligned} N_h(\mu, \mu+1) &= \#\{\mu_j : \mu \le \mu_j \le \mu+1 \text{ with multiplicity}\} \\ &= (n-1)\frac{\operatorname{Vol}(B_{n-1})}{(2\pi)^{n-1}} \operatorname{Vol}(Y, h)\mu^{n-2} + o(\mu^{n-2}). \end{aligned}$$

As $\nu_j^2 = \mu_j^2 + (n-2)^2/4$, the same counting formula holds for ν_j; i.e.,

$$\begin{aligned} N_\nu(\rho, \rho+1) &= \#\{\nu_j : \rho \le \nu_j \le \rho+1 \text{ with multiplicity}\} \\ &= (n-1)\frac{\operatorname{Vol}(B_{n-1})}{(2\pi)^{n-1}} \operatorname{Vol}(Y, h)\rho^{n-2} + o(\rho^{n-2}). \end{aligned} \tag{9}$$

We now turn our attention to the zeros of the Hankel function $H^{(2)}_\nu(z)$ with $\arg z \in \left[0, \frac{\pi}{2}\right]$. An argument from [Watson 1944, pages 511–513] is easily adapted to give a precise count of the number of zeros of

[1] When (Y, h) is a sphere, the analysis is simplified slightly. In that case, one replaces the use of the Weyl formula with explicit formulae for the eigenvalues μ_j^2 and their multiplicities.

$H_\nu^{(2)}$ in this sector. Indeed, that argument shows that the number of zeros is given by the closest integer to $\frac{1}{2}\nu - \frac{1}{4}$ (when $\nu - \frac{1}{2}$ is an integer, there is a zero on the imaginary axis and so rounds up).

As $\nu \to \infty$ through positive real values, we have an asymptotic expansion [DLMF 2018, 10.20.6] relating the Hankel function to the Airy function

$$H_\nu^{(2)}(\nu z) \sim 2e^{i\pi/3}\left(\frac{4\zeta}{1-z^2}\right)^{1/4}\left(\frac{\mathrm{Ai}(e^{-2\pi i/3}\nu^{2/3}\zeta)}{\nu^{1/3}}\sum_{k=0}^{\infty}\frac{A_k(\zeta)}{\nu^{2k}} + \frac{\mathrm{Ai}'(e^{-2\pi i/3}\nu^{2/3}\zeta)}{\nu^{5/3}}\sum_{k=0}^{\infty}\frac{B_k(\zeta)}{\nu^{2k}}\right). \quad (10)$$

Here A_k and B_k are real and infinitely differentiable for $\zeta \in \mathbb{R}$. This expansion is uniform in $|\arg z| \le \pi - \delta$ for fixed $\delta > 0$. In particular, for large enough ν, the zeros of the Hankel function are well-approximated by zeros of the Airy function and we may identify each zero $h_{\nu,k}$ of the Hankel function $H_\nu^{(2)}$ with a zero of the Airy function $\mathrm{Ai}(-z)$.

Let a_k denote the k-th zero of the Airy function $\mathrm{Ai}(-z)$; all a_k are positive and

$$a_k = \left[\tfrac{3}{2}\left(k\pi - \tfrac{\pi}{4}\right)\right]^{2/3} + O(k^{-4/3}).$$

We now define $\lambda_{\nu,k}$ and $\tilde{\lambda}_{\nu,k}$ via the Airy zeros and their leading approximations:

$$\lambda_{\nu,k} = \nu\zeta^{-1}(\nu^{-2/3}e^{-i\pi/3}a_k) = \nu\rho^{-1}\left(-i\tfrac{2}{3}a_k^{3/2}\nu^{-1}\right),$$
$$\tilde{\lambda}_{\nu,k} = \nu\rho^{-1}\left(-i\left(k - \tfrac{1}{4}\right)\pi\nu^{-1}\right),$$

where $k = 1, \dots, \left\lfloor \frac{1}{2}\nu + \frac{1}{4} \right\rfloor$. By the Hankel expansion (10), $|h_{\nu,k} - \lambda_{\nu,k}| \le C/\nu$ for large enough ν, while $|h_{\nu,k} - \tilde{\lambda}_{\nu,k}| \le C/\nu$ for large enough ν and k. As we have identified $\left\lfloor \frac{1}{2}\nu + \frac{1}{4} \right\rfloor$ approximate zeros, we can conclude that these account for all $h_{\nu,k}$.

We now divide our attention into those zeros with small argument and those with large argument. We introduce the auxiliary counting function

$$N(r, \theta_1, \theta_2) = \#\{\sigma : \sigma \text{ is a resonance with } |\sigma| \le r,\ \arg\sigma \in [\theta_1, \theta_2]\}.$$

We first address those with small argument. Fix $\epsilon > 0$ and consider those zeros with $|z| < r$ and $\arg z \in [0, \epsilon]$. We need count those $\lambda_{\nu,k}$ with $\arg \lambda_{\nu,k} \in [0, \epsilon]$ and $|\lambda_{\nu,k}| \le r$. As $|\lambda_{\nu,k}|$ is comparable to ν, we can over-count these zeros by counting all $\lambda_{\nu,k}$ with argument in $[0, \epsilon]$ and $\nu \le Cr$.

Because $|\rho| \le C\epsilon^{3/2}$ for those $\lambda_{\nu,k}$ with $\arg \lambda_{\nu,k} \in [0, \epsilon]$, we must only count those a_k with $a_k \le C\nu^{2/3}\epsilon$. The leading order asymptotic [DLMF 2018, 9.9.6] for the zeros of the Airy function shows that this number is $O(\nu\epsilon^{3/2})$.

We now count those resonances with argument in $[0, \epsilon]$. Putting together the asymptotic for ν_j in (9) with the previous two paragraphs, we have (with $m(\nu_j)$ denoting the multiplicity of ν_j)

$$\begin{aligned} N(r, 0, \epsilon) &= \sum_{j=1}^{\infty} m(\nu_j)\#\{h_{\nu_j,k} : |h_{\nu_j,k}| \le r, \arg h_{\nu_j,k} \in [0, \epsilon]\} \\ &\le \sum_{j=1}^{Cr} m(\nu_j)C\nu_j\epsilon^{3/2} \le C\epsilon^{3/2}\sum_{\rho=0}^{Cr}\sum_{\nu_j\in[\rho,\rho+1]} m(\nu_j)\rho \le C\epsilon^{3/2}r^n. \end{aligned} \quad (11)$$

We now consider those resonances with argument in $\left[\epsilon, \frac{\pi}{2}\right]$. For large enough ν, the approximations $\tilde{\lambda}_{\nu,k}$ are valid for these resonances. We count those approximate resonances with $\nu_j \in [\rho, \rho+1)$ and $\arg\lambda_{\nu,k} \in [\theta, \theta+\Delta\theta]$. We start by introducing, for fixed ν, the number Δk_ν of $\tilde{\lambda}_{\nu,k}$ with argument lying in $[\theta, \theta+\Delta\theta]$. Observe that the definition of $\tilde{\lambda}_{\nu,k}$ relates Δk_ν with Δt by

$$\Delta k_\nu = \frac{\nu}{\pi}\Delta t + O(1),$$

where Δt denotes the change in t corresponding to $\Delta\theta$ in the parametrizations above. Note that Δt is *independent* of the choice of ν. We can then write

$$\#\{\tilde{\lambda}_{\nu,k} : \nu_j \in [\rho, \rho+1), \arg\tilde{\lambda}_{\nu,k} \in [\theta, \theta+\Delta\theta]\} = \sum_{\rho\le\nu_j\le\rho+1} m(\nu_j)\Delta k_\nu = \sum_{\rho\le\nu_j<\rho+1} m(\nu_j)\left(\frac{\nu_j}{\pi}\Delta t + O(1)\right).$$

By the definition of the approximate zeros $\tilde{\lambda}_{\nu,k}$, we can estimate their size $|\tilde{\lambda}_{\nu,k}|$ in terms of $|z(\theta)|$, provided that $\arg\tilde{\lambda}_{\nu,k} \in [\theta, \theta+\Delta\theta]$, yielding

$$|\tilde{\lambda}_{\nu,k}| = \nu(|z(\theta)| + O(\Delta\theta)).$$

In particular, if $\nu_j|z(\theta)| \ge r$ but $|\lambda_{\nu,k}| \le r$, then

$$\nu_j \in \left[\frac{r}{|z(\theta)|}(1 - c\Delta\theta), \frac{r}{|z(\theta)|}\right].$$

We may thus rewrite our counting function as

$$\begin{aligned}\#\{\tilde{\lambda}_{\nu,k} : |\tilde{\lambda}_{\nu,k}| \le r, \arg\tilde{\lambda}_{\nu,k} \in [\theta, \theta+\Delta\theta]\} &= \sum_{\substack{|\tilde{\lambda}_{\nu,k}|\le r \\ \arg\tilde{\lambda}_{\nu,k}\in[\theta,\theta+\Delta\theta]}} m(\nu_j) \\ &= \sum_{\substack{\nu_j|z(\theta)|\le r \\ \arg\tilde{\lambda}_{j,k}\in[\theta,\theta+\Delta\theta]}} m(\nu_j) + \sum_{\substack{\nu_j\in[(r/|z(\theta)|)(1-c\Delta\theta), r/|z(\theta)|] \\ \arg\tilde{\lambda}_{\nu,k}\in[\theta,\theta+\Delta\theta]}} m(\nu_j).\end{aligned}$$

By our improved Weyl's law (9), the second term is $O(r^{n-2})$.

We now focus our attention on the first term (here $\lfloor\cdot\rfloor$ denotes the "floor" function):

$$\begin{aligned}\sum_{\substack{\nu_j|z(\theta)|\le r \\ \arg\tilde{\lambda}_{j,k}\in[\theta,\theta+\Delta\theta]}} m(\nu_j) &= \sum_{\rho=0}^{\lfloor r/|z|-1\rfloor} \sum_{\nu_j\in[\rho,\rho+1)} \sum_{\arg\tilde{\lambda}_{\nu,k}\in[\theta,\theta+\Delta\theta]} m(\nu_j) + \sum_{\nu_j\in[\lfloor r/z\rfloor, r/z]} \sum_{\arg\tilde{\lambda}_{\nu,k}\in[\theta,\theta+\Delta\theta]} m(\nu_j) \\ &= \sum_{\rho=0}^{\lfloor r/|z|-1\rfloor} \sum_{\nu_j\in[\rho,\rho+1)} m(\nu_j)\Delta k_\nu + \sum_{\nu_j\in[\lfloor r/z\rfloor, r/z]} \sum_{\arg\tilde{\lambda}_{\nu,k}\in[\theta,\theta+\Delta\theta]} m(\nu_j).\end{aligned}$$

Again by Weyl's law, we observe that the second term is $O(r^{n-2})$. By relating Δt and Δk_ν we can rewrite the first term:

$$\sum_{\rho=0}^{\lfloor r/|z|-1\rfloor} \sum_{\nu_j\in[\rho,\rho+1)} m(\nu_j)\Delta k_\nu = \sum_{\rho=0}^{\lfloor r/|z|-1\rfloor} \sum_{\nu_j\in[\rho,\rho+1)} m(\nu_j)\frac{\nu_j}{\pi}\Delta t + \sum_{\nu_j\le\lfloor r/|z|\rfloor} m(\nu_j)O(1).$$

By Weyl's law (9), the second term is $O(r^{n-1})$, so we again consider the first term.

As Δt is independent of ν_j, we may use Weyl's law as well on the first term:

$$\sum_{\rho=0}^{\lfloor r/|z|-1\rfloor} \sum_{\nu_j\in[\rho,\rho+1)} m(\nu_j)\frac{\nu_j}{\pi}\Delta t = \sum_{\rho=0}^{\lfloor r/|z|-1\rfloor}\left[\frac{n-1}{2^{n-1}\pi^n}\operatorname{Vol}(B_{n-1})\operatorname{Vol}(Y,h)\rho^{n-1}\Delta t + O(\rho^{n-2}) + o(\rho^{n-1})\Delta t\right]$$
$$= \frac{2(n-1)}{(2\pi)^n}\operatorname{Vol}(B_{n-1})\operatorname{Vol}(Y,h)\Delta t \sum_{\rho=0}^{\lfloor r/|z|-1\rfloor}\rho^{n-1} + O(r^{n-1}) + o(r^n)\Delta t$$
$$= \frac{2(n-1)}{(2\pi)^n n}\operatorname{Vol}(B_{n-1})\operatorname{Vol}(Y,h)\frac{1}{n}\left(\frac{r}{|z(\theta)|}\right)^n\Delta t + O(r^{n-1}) + o(r^n)\Delta t.$$

We finally introduce a Riemann sum in t to understand this main term:

$$\begin{aligned}
\#\{\tilde\lambda_{\nu,k} : |\tilde\lambda_{\nu,k}| \le r,\ \arg\tilde\lambda_{\nu,k}\in[\epsilon,\tfrac{\pi}{2}]\}
&= \int_{t^{-1}(\epsilon)}^{\pi/2}\left(\frac{2(n-1)\operatorname{Vol}(B_{n-1})}{(2\pi)^n n}\operatorname{Vol}(Y,h)\right)\frac{r^n}{|z(\theta)|^n}\,dt + O(r^{n-1}) + o(r^n)\\
&= \frac{(n-1)\operatorname{Vol}(B_{n-1})}{(2\pi)^n n}\operatorname{Vol}(Y,h)r^n\int_{\partial K_+}\frac{1}{|z(\theta)|^n}\,dt + O(\epsilon r^n) + o(r^n)\\
&= \left(\frac{(n-1)\operatorname{Vol}(B_{n-1})}{(2\pi)^n n}\operatorname{Vol}(Y,h)\int_{\partial K_+}\frac{|1-z^2|^{1/2}}{|z|^{n+1}}\,d|z|\right)r^n + O(\epsilon r^n) + o(r^n)\\
&= A_n\operatorname{Vol}(Y,h)r^n + O(\epsilon r^n) + o(r^n). \qquad (12)
\end{aligned}$$

Here the prefactor of 2 disappeared because the first integral parametrizes only half of ∂K_+. It reappears in the statement of Theorem 1.1 because each zero there corresponds to two resonances (one on each sheet). We further observe that the constant $A_n\operatorname{Vol}(Y,h)$ agrees with the leading term found in the Euclidean case found in [Stefanov 2006].

Sending ϵ to 0 establishes the theorem for the approximate zeros $\lambda_{\nu,k}$. Because each $\lambda_{\nu,k}$ is in a C/ν neighborhood of a zero $h_{\nu,k}$, this finishes the proof of the theorem.

Acknowledgments

Part of this research formed the core of Yang's master's project at Texas A&M University. Baskin acknowledges partial support from NSF grants DMS-1500646 and DMS-1654056. The authors also thank David Borthwick, Tanya Christiansen, Colin Guillarmou, and Jeremy Marzuola for helpful conversations.

References

[Borthwick 2010] D. Borthwick, "Sharp upper bounds on resonances for perturbations of hyperbolic space", *Asymptot. Anal.* **69**:1-2 (2010), 45–85. MR Zbl

[Borthwick 2012] D. Borthwick, "Sharp geometric upper bounds on resonances for surfaces with hyperbolic ends", *Anal. PDE* **5**:3 (2012), 513–552. MR Zbl

[Borthwick and Philipp 2014] D. Borthwick and P. Philipp, "Resonance asymptotics for asymptotically hyperbolic manifolds with warped-product ends", *Asymptot. Anal.* **90**:3-4 (2014), 281–323. MR Zbl

[Cheeger and Taylor 1982a] J. Cheeger and M. Taylor, "On the diffraction of waves by conical singularities, I", *Comm. Pure Appl. Math.* **35**:3 (1982), 275–331. MR Zbl

[Cheeger and Taylor 1982b] J. Cheeger and M. Taylor, "On the diffraction of waves by conical singularities, II", *Comm. Pure Appl. Math.* **35**:4 (1982), 487–529. MR Zbl

[DLMF 2018] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, and B. V. Saunders (editors), *NIST digital library of mathematical functions*, 2018. release 1.0.21.

[Duistermaat and Guillemin 1975] J. J. Duistermaat and V. W. Guillemin, "The spectrum of positive elliptic operators and periodic bicharacteristics", *Invent. Math.* **29**:1 (1975), 39–79. MR Zbl

[Ivrii 1980] V. Y. Ivrii, "The second term of the spectral asymptotics for a Laplace–Beltrami operator on manifolds with boundary", *Funktsional. Anal. i Prilozhen.* **14**:2 (1980), 25–34. In Russian; translated in *Funct. Anal. Appl.* **14**:2 (1980), 98–106. MR Zbl

[Ivrii 1982] V. Y. Ivrii, "Exact spectral asymptotics for elliptic operators acting in vector bundles", *Funktsional. Anal. i Prilozhen.* **16**:2 (1982), 30–38. In Russian; translated in *Funct. Anal. Appl.* **16**:2 (1982), 101–108. MR Zbl

[Olver 1974] F. W. J. Olver, *Asymptotics and special functions*, Academic Press, New York, 1974. MR Zbl

[Stefanov 2006] P. Stefanov, "Sharp upper bounds on the number of the scattering poles", *J. Funct. Anal.* **231**:1 (2006), 111–142. MR Zbl

[Watson 1944] G. N. Watson, *A treatise on the theory of Bessel functions*, Cambridge University Press, 1944. MR Zbl

Received 23 Apr 2019. Revised 30 Sep 2019. Accepted 19 Nov 2019.

DEAN BASKIN: dbaskin@math.tamu.edu
Department of Mathematics, Texas A&M University, College Station, TX, United States

MENGXUAN YANG: mxyang@math.northwestern.edu
Department of Mathematics, Northwestern University, Evanston, IL, United States

PURE and APPLIED
ANALYSIS

Vol. 2, No. 2, 2020
https://doi.org/10.2140/paa.2020.2.397

THE KÄHLER GEOMETRY OF CERTAIN OPTIMAL TRANSPORT PROBLEMS

GABRIEL KHAN AND JUN ZHANG

Let X and Y be domains of $\mathbb{R}^n$ equipped with probability measures μ and ν, respectively. We consider the problem of optimal transport from μ to ν with respect to a cost function $c : X \times Y \to \mathbb{R}$. To ensure that the solution to this problem is smooth, it is necessary to make several assumptions about the structure of the domains and the cost function. In particular, Ma, Trudinger, and Wang established regularity estimates when the domains are strongly *relatively c-convex* with respect to each other and the cost function has nonnegative *MTW tensor*. For cost functions of the form $c(x, y) = \Psi(x - y)$ for some convex function $\Psi : \mathcal{M} \to \mathbb{R}$, we find an associated Kähler manifold on $T\mathcal{M}$ whose orthogonal antibisectional curvature is proportional to the MTW tensor. We also show that relative c-convexity geometrically corresponds to geodesic convexity with respect to a dual affine connection on $\mathcal{M}$. Taken together, these results provide a geometric framework for optimal transport which is complementary to the pseudo-Riemannian theory of Kim and McCann (*J. Eur. Math. Soc.* **12**:4 (2010), 1009–1040).

We provide several applications of this work. In particular, we find a complete Kähler surface with nonnegative orthogonal antibisectional curvature that is not a Hermitian symmetric space or biholomorphic to $\mathbb{C}^2$. We also address a question in mathematical finance raised by Pal and Wong (2018, arXiv:1807.05649) on the regularity of *pseudoarbitrages*, or investment strategies which outperform the market.

1. Introduction

Optimal transport is a classic field of mathematics combining ideas from geometry, probability, and analysis. The problem was first formalized by Gaspard Monge [1781]. In his work, he considered a worker who is tasked with moving a large pile of sand into a prescribed configuration and wants to minimize the total effort required to complete the job. Trying to determine the optimal way of transporting the sand leads into deep and subtle mathematical phenomena and is a thriving field of research to this day. Furthermore, optimal transport has many practical applications. Monge's work was originally inspired by a problem in engineering, but these same ideas can be applied to logistics, economics, computer imaging processing, and many other fields [Peyré and Cuturi 2019].

The modern framework for optimal transport, due to Kantorovich [1958], considers arbitrary couplings between two probability measures. In this formulation, we consider X and Y as Borel subsets of two metric spaces equipped with probability measures μ and ν, respectively. Intuitively, $d\mu$ is the shape of the original sand pile and $d\nu$ is the target configuration. To transport the sand from μ to ν, we consider a *coupling* of μ and ν, which is a nonnegative measure on $X \times Y$ whose marginal distributions are μ and ν,

MSC2010: primary 49Q20, 53C55; secondary 46N10, 46N30.
Keywords: optimal transportation, Kähler metrics, regularity of optimal maps, complex geometry, MTW condition, curvature, Monge–Kantorovich, tube domains, tangent bundle.

respectively. To measure the efficiency of a plan for transport μ to ν, we consider a lower-semicontinuous cost function $c : X \times Y \to \mathbb{R}$. The solution to the Kantorovich optimal transport problem is the coupling γ which achieves the smallest total cost

$$\min_{\gamma \in \Gamma(\mu,\nu)} \int_{X \times Y} c(x, y)\, d\gamma(x, y).$$

Here $\Gamma(\mu, \nu)$ is the set of all couplings of μ and ν. In this case, a minimizing measure γ is referred to as the *optimal coupling*. An optimal coupling exists for very general measures and cost functions, so the Kantorovich approach is a flexible and powerful framework to study optimal transport.

In Monge's work, it is assumed that the mass at a given point will not be subdivided and sent to multiple locations. This is known as *deterministic* optimal transport, which seeks to find a measurable map $\mathbb{T} : X \to Y$ so that the optimal coupling is entirely contained within the graph of $\mathbb{T}$. When this occurs, the map $\mathbb{T}$ is known as the *optimal map*. A priori, there is no guarantee that optimal transport is deterministic, so a Monge solution may not exist for a given optimal transport problem. We will discuss certain sufficient conditions for the optimal transport to be deterministic in Section 2.

For deterministic optimal transport, it is natural to ask whether the optimal map is continuous or even smooth. This is known as the *regularity problem for optimal transport*. Historically, most of the work on this problem was done in Euclidean space for the cost $c(x, y) = \|x - y\|^2$, better known as the quadratic cost.

For more general cost functions (such as quadratic costs on Riemannian manifolds), the groundbreaking work was done by Ma, Trudinger and Wang [Ma et al. 2005], who proved that the transport map is smooth under the assumptions that

(1) a certain nonlinear fourth-order quantity, known as the MTW tensor (denoted by $\mathfrak{S}$), is nonnegative, and that

(2) the sets X and Y are *relatively c-convex* with respect to each other.[1]

These results were refined by Loeper [2009], who showed that the nonnegativity of $\mathfrak{S}$ is necessary to establish continuity for the optimal transport between smooth measures. Furthermore, he gave some insight into the geometric significance of the MTW tensor. Later work of Kim and McCann [2010] furthered this understanding by presenting a pseudo-Riemannian framework for optimal transport in which the MTW tensor is the curvature of certain light-like planes.

1.1. ***Our results.*** In this paper, we primarily consider Ψ-costs, which we define as follows.

Definition (Ψ-cost). Let $\Psi : \mathcal{M} \to \mathbb{R}$ be a locally strongly convex C^4 function[2] on an open domain $\mathcal{M}$ in Euclidean space.

For open domains X and Y in $\mathbb{R}^n$, a Ψ-cost is a cost function of the form

$$c : X \times Y \to \mathbb{R}, \quad c(x, y) = \Psi(x - y).$$

[1] More precisely, the assumption is that the supports of μ and ν are relatively c-convex.

[2] Here, a function being locally strongly convex means that the Hessian is positive definite. Furthermore, it is possible to work with less regular convex functions, but we will not do so in this paper.

These costs were previously studied by Gangbo and McCann [1995] and by Ma, Trudinger and Wang [2005]. For such a cost to be well-defined, $\mathcal{M}$ must contain the difference set $X - Y$, defined as

$$X - Y := \{z \in \mathbb{R}^n \mid \text{there exists } x \in X,\ y \in Y \text{ such that } z = x - y\}.$$

We can now summarize the main results of our work, which associate a complex manifold to a given Ψ-cost. To do so, we consider $\mathcal{M}$ as a Hessian manifold, using Ψ as its potential function (i.e., setting $g_{ij} = \partial^2\Psi/(\partial u^i \partial u^j)$). Such manifolds naturally admit a dual pair of flat connections, which we denote by D and D^* [Shima 2007].

Using the primal flat connection D and the metric g, there is a canonical Kähler metric on the tangent bundle, known as the Sasaki metric and denoted by $(T\mathcal{M}, g^D, J^D)$. Our main result shows the following correspondence between the curvature of this metric and the MTW tensor.

Theorem. *Let X and Y be open sets in $\mathbb{R}^n$ and c be a Ψ-cost. Then the MTW tensor $\mathfrak{S}$ satisfies the identity*

$$\tfrac{1}{2}\mathfrak{S}(\eta, \xi) = \mathfrak{R}_{g^D}(\xi, J^D\eta^\sharp, \xi, J^D\eta^\sharp) - \mathfrak{R}_{g^D}(\eta^\sharp, \xi, \eta^\sharp, \xi),$$

where ξ and η are an orthogonal real vector-covector pair (which we extend[3] to $T\mathcal{M}$) and $\mathfrak{R}_{g^D}$ is the curvature of $(T\mathcal{M}, g^D, J^D)$ (where the metric is induced by the potential Ψ).

For reasons that we will explain later, we call the right-hand expression the *orthogonal antibisectional curvature*. We furthermore show that relative c-convexity of sets is geodesic convexity with respect to the dual affine connection on $\mathcal{M}$.

Proposition. *For a Ψ-cost, a set Y is c-convex relative to X if and only if, for all $x \in X$, the set $x - Y$ is geodesically convex with respect to the dual connection D^*. Here, D^* is the connection on $\mathcal{M}$ satisfying*

$$\mathcal{X}(g(\mathcal{Y}, \mathcal{Z})) = g(D_{\mathcal{X}}\mathcal{Y}, \mathcal{Z}) + g(\mathcal{Y}, D^*_{\mathcal{X}}\mathcal{Z})$$

for all vector fields $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$.

Apart from providing a new geometric framework for the regularity problem, we can use these results to address several questions of independent interest.

1.1.1. *Applications to complex geometry.* This approach can be used to construct several examples of interesting metrics with subtle nonnegativity properties. In particular, we find a complete complex surface which is neither biholomorphic to $\mathbb{C}^2$ nor Hermitian symmetric but whose orthogonal antibisectional curvature is nonnegative. Many of the complex manifolds constructed using this approach are of independent interest, and we will provide a few examples which we will study in depth in future work.

1.1.2. *Applications to mathematical finance.* Our second main application is to establish regularity for a certain problem in portfolio design theory. The recent work [Pal and Wong 2016] studies the problem of finding *pseudoarbitrages*, which are investment strategies that outperform the market almost surely in the long run under mild and realistic assumptions on the stock market. Their work shows that this is

[3] To be more precise, we consider certain lifts of these vectors to $T\mathcal{M}$. For a more formal statement, see Theorem 6.

equivalent to solving an optimal transport problem where the cost function is a divergence function (in information-geometric language) that is closely related to the free energy in statistical physics.

For this problem, our approach relates the MTW tensor of this cost to a Kähler manifold with constant positive holomorphic sectional curvature. As such, this cost function satisfies the MTW(0) condition (and also satisfies a stronger condition known as *nonnegative cost-curvature* [Figalli et al. 2011]). We further show that relative c-convexity corresponds precisely to the standard notion of convexity on the probability simplex. Combining these calculations, we can apply the results of [Trudinger and Wang 2009] to obtain a regularity theory of portfolio maps and their associated displacement interpolations. This addresses a question asked in [Pal and Wong 2018b], and intuitively shows that when the market conditions change slightly, the investment strategy similarly does not change by much.

A preliminary announcement of some of these results (stated in terms of the so-called $\mathcal{D}_{\Psi}^{(\alpha)}$-divergences) appeared in [Khan and Zhang 2019].

1.2. ***Layout of the paper.*** In Section 2 we discuss some background information on optimal transport. In Section 3 we review some complex and Kähler geometry. Section 4 discusses some background information on Hessian manifolds and the curvature of the Sasaki metric. In Section 5, we state our main results, which show the precise interaction between complex/information geometry and the regularity theory of optimal transport. In Section 6, we explore various applications of this result. In Section 7, we conclude with a section of open questions, which we hope to explore in future work.

1.3. ***Notation.*** We have attempted to preserve the notation from [De Philippis and Figalli 2014; Satoh 2007] as much as possible, while minimizing abuse of notation or overlap. For clarity, we introduce some notational conventions now.

Throughout the paper, X and Y will denote open domains in $\mathbb{R}^n$. Invariably, these will be smooth and bounded. We will use $\{x^i\}_{i=1}^n$ as coordinates on X and $\{y^i\}_{i=1}^n$ as coordinates on Y. To study optimal transport, we will use $c(x, y)$ to denote a cost function $c : X \times Y \to \mathbb{R}$, which will generally be C^4 in this paper. Often times, the domain of c will be larger than $X \times Y$, but we will ignore this. To avoid confusion with coordinate functions and the notation for tangent spaces, we denote the solutions to equations of Monge–Ampère type by U, and the associated optimal map by $\mathbb{T}_{\mathrm{U}}$.

For the most part, $\mathcal{M}$ will be an open domain in Euclidean space which contains $X - Y$, and Ψ will denote a convex function $\Psi : \mathcal{M} \to \mathbb{R}$. It is instructive to also consider $\mathcal{M}$ as an affine manifold, and we will use $\{u^i\}_{i=1}^n$ as its coordinates. When considering the tangent bundle of $\mathcal{M}$ (denoted by $T\mathcal{M}$), we will use bundle coordinates $\{(u^i, v^i)\}_{i=1}^n$. This notation is a change from [Satoh 2007] and is done to avoid overusing x and y.

In order to prescribe $T\mathcal{M}$ with a Hermitian structure, it is necessary to consider a flat affine connection on $\mathcal{M}$, which we denote by D. More precisely, this will be the affine connection induced by differentiation within the u-coordinates. Furthermore, we use $\mathcal{W}$, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$ and ξ to denote tangent vectors on $\mathcal{M}$ (i.e., elements of $T\mathcal{M}$). This is the convention of [Satoh 2007], except with calligraphic font to avoid confusion with our notation for domains. When computing the MTW tensor, we will denote the vectors in the MTW tensor by ξ and the covectors by η.

To simplify the derivative notation, for a two-variable function $c(x, y)$, we use $c_{I,J}$ to denote $\partial_{x^I}\partial_{y^J}c$ for multi-indices I and J. Furthermore, $c^{i,j}$ denotes the matrix inverse of the mixed derivative $c_{i,j}$. For a convex function Ψ, we use the notation Ψ_J to denote $\partial_{u^J}\Psi$ for a multi-index J and the notation Ψ^{ij} to denote the matrix inverse of Ψ_{ij}. Finally, we will use Einstein summation notation throughout the paper.

2. Background on the regularity theory of optimal transport

The main focus of our paper is to study the assumptions needed to ensure optimal transport is regular. In order to understand these, we first review several preliminary results on the regularity theory of optimal transport.

As our primary interest is the geometric structure of the regularity problem, we will not make use of the sharpest possible regularity estimates. The material in this section is based on the survey [De Philippis and Figalli 2014], which provides a more complete background for the regularity theory. For a more thorough overview on optimal transport, see [Villani 2009].

The regularity problem arises when the optimal coupling in the Kantorovich optimal transport problem is induced by a deterministic transport map. As such, we first discuss some conditions which ensure the Kantorovich optimal transport problem has a deterministic solution. The following theorem was originally proven in [Brenier 1987] for the quadratic cost and in more generality in [Gangbo and McCann 1996]. It gives sufficient conditions for deterministic transport and shows that the optimal maps can be found by solving an equation of Monge–Ampère type.

Theorem 1. *Let X and Y be two open domains of $\mathbb{R}^n$ and consider a cost function $c : X \times Y \to \mathbb{R}$. Suppose that $d\mu$ is a smooth probability density supported on X and that $d\nu$ is a smooth probability density supported on Y. Suppose that the following conditions hold*:

1. *The cost function c is of class C^4 with $\|c\|_{C^4(X\times Y)} < \infty$.*
2. *For any $x \in X$, the map $Y \ni y \to c_x(x, y) \in \mathbb{R}^n$ is injective.*
3. *For any $y \in Y$, the map $X \ni x \to c_y(x, y) \in \mathbb{R}^n$ is injective.*
4. $\det(c_{x,y})(x, y) \neq 0$ *for all* $(x, y) \in X \times Y$.

Then there exists a c-convex function $\mathrm{U} : X \to \mathbb{R}$ such that the map $\mathbb{T}_{\mathrm{U}} : X \to Y$ defined by $\mathbb{T}_{\mathrm{U}}(x) := c\text{-}\exp_x(\nabla \mathrm{U}(x))$ is the unique optimal transport map sending μ onto ν. Furthermore, $\mathbb{T}_{\mathrm{U}}$ is injective $d\mu$-a.e.,

$$|\det(\nabla \mathbb{T}_{\mathrm{U}}(x))| = \frac{d\mu(x)}{d\nu(\mathbb{T}_{\mathrm{U}}(x))} \quad d\mu\text{-a.e.}, \tag{1}$$

and its inverse is given by the optimal transport map sending ν onto μ.

In order to express (1) more concretely, we recall the notion of the c-exponential map (denoted by $c\text{-}\exp_x$).

Definition (c-exponential map). For any $x \in X$, $y \in Y$, $p \in \mathbb{R}^n$, the c-exponential map satisfies the identity

$$c\text{-}\exp_x(p) = y \quad \iff \quad p = -c_x(x, y).$$

For the squared-distance cost on a Riemannian manifold, the c-exponential is exactly the standard exponential map, which motivates its name. For this cost in Euclidean space, (1) becomes the standard Monge–Ampère equation

$$\det(\nabla^2 \mathrm{U}(x)) = \frac{f(x)}{g(\nabla \mathrm{U}(x))}. \tag{2}$$

Due to the comparatively simple form for (2), much of the initial work on the regularity problem was done for the quadratic cost in Euclidean space. In this setting, Caffarelli [1992] and others proved a priori estimates under certain convexity and smoothness assumptions on the measures (for a more complete history, see [De Philippis and Figalli 2014]). Caffarelli also observed there is no hope of proving interior regularity for $\mathbb{T}_{\mathrm{U}}$ without assuming that the support of the target measure is convex.

For more general cost functions, Ma, Trudinger and Wang's breakthrough work [Ma et al. 2005] gave three conditions that ensure C^2 regularity for the solutions of the Monge–Ampère equation (1). In this paper, we will use a stronger version of this result, originally proven in [Trudinger and Wang 2009].

Theorem 2. *Suppose that $c : X \times Y \to \mathbb{R}$, μ, and ν satisfy the hypothesis of Theorem 1, and the densities $d\mu$ and $d\nu$ are bounded away from zero and infinity on their respective supports X and Y. Suppose further that the following hold*:

(1) *X and Y are smooth.*

(2) *The domain X is strictly c-convex relative to the domain Y.*

(3) *The domain Y is strictly c^*-convex relative to the domain X.*

(4) *For all vectors $\xi, \eta \in \mathbb{R}^n$ with $\xi \perp \eta$, the following inequality holds*:

$$\mathfrak{S}(\xi, \eta) := \sum_{i,j,k,l,p,q,r,s} (c_{ij,p} c^{p,q} c_{q,rs} - c_{ij,rs}) c^{r,k} c^{s,l} \xi^i \xi^j \eta^k \eta^l \geq 0. \tag{3}$$

Then $\mathrm{U} \in C^\infty(\bar{X})$ and $\mathbb{T}_{\mathrm{U}} : \bar{X} \to \bar{Y}$ is a smooth diffeomorphism, where $\mathbb{T}_{\mathrm{U}}(x) = c\text{-}\exp_x(\nabla \mathrm{U}(x))$.

While we will not discuss the proof in detail, we note that the main challenge is obtaining an a priori C^2 estimate on U. Once such an estimate is established, the Monge–Ampère equation can be linearized at U, at which point standard elliptic bootstrapping yields estimates of all orders and implies that $\mathbb{T}_{\mathrm{U}}$ is smooth.

The main results of this paper concern the assumptions of Theorem 2, so we discuss these in more detail. The first condition is self-explanatory, while the second and third provide the proper notions of convexity for the supports. To explain this in detail, we recall the notion of c-convexity for sets.

Definition (c-segment). A c-segment in X with respect to a point y is a solution set $\{x\}$ to $c_{,y}(x, y) \in \ell$ for ℓ a line segment in $\mathbb{R}^n$. A c^*-segment in Y with respect to a point x is a solution set $\{y\}$ to $c_{x,}(x, y) \in \ell$, where ℓ is a line segment in $\mathbb{R}^n$.

Definition (c-convexity). A set E is c-convex relative to a set E^* if for any two points $x_0, x_1 \in E$ and any $y \in E^*$, the c-segment relative to y connecting x_0 and x_1 lies in E. Similarly we say E^* is c^*-convex relative to E if for any two points $y_0, y_1 \in E^*$ and any $x \in E$, the c^*-segment relative to x connecting y_0 and y_1 lies in E^*.

Finally, we discuss inequality (3), which is known as the MTW(0) condition and is a weakened version of the MTW(κ) condition.

Definition (MTW(κ), $\kappa > 0$). A cost function c satisfies the MTW(κ) condition if for any orthogonal vector-covector pair ξ and η we have $\mathfrak{S}(\xi, \eta) \geq \kappa|\xi|^2|\eta|^2$ for $\kappa > 0$.

Ma, Trudinger and Wang's original work relied on the MTW(κ) assumption, and this stronger condition is used in many applications. Although it is not immediately apparent, $\mathfrak{S}(\xi, \eta)$ is tensorial (coordinate-invariant) so long as one considers η as a covector [Kim and McCann 2010], which we will do throughout the rest of the paper. Furthermore, it transforms quadratically in η and ξ, but is highly nonlinear and nonlocal in the cost function.

The geometric significance of the MTW tensor is an active topic of research. On a Riemannian manifold, Loeper [2009] gave some insight into its behavior. His work showed that for the quadratic cost, the MTW tensor is proportional to the sectional curvature on the diagonal $x = y$. In this paper, he also showed that c-convexity and nonnegativity of the MTW tensor are essentially necessary conditions to prove regularity of optimal transport.

Building on Loeper's results, Kim and McCann [2010] gave a geometric framework for optimal transport. In their formulation, optimal transport is expressed in terms of a pseudo-Riemannian metric on the manifold $X \times Y$ and the MTW tensor becomes the curvature of light-like planes. This interpretation holds for arbitrary cost functions, and gives intrinsic geometric structure to the regularity problem. Our geometric interpretation is different, but many of the formulas appear similar, in part due to the fact that Kim and McCann chose notation reminiscent of complex geometry.

Before concluding our background discussion on optimal transport, we will introduce one more strengthening of the MTW(0) condition, known as nonnegative "cross-curvature" [Figalli et al. 2011].

Definition (nonnegative cross-curvature). A cost function c has nonnegative (resp. strictly positive $\kappa > 0$) cross-curvature if, for any vector-covector pair η and ξ,

$$\mathfrak{S}(\xi, \eta) \geq 0 \quad (\text{resp. } \kappa|\xi|^2|\eta|^2 \text{ for some } \kappa > 0).$$

Note that nonnegative cross-curvature is stronger than MTW(0), as the nonnegativity must hold for all pairs η and ξ, not simply orthogonal ones. Cross-curvature was introduced by Figalli, Kim, and McCann [Figalli et al. 2011] to study a problem in microeconomics. In later work, they also showed that stronger regularity for optimal maps can be proven with this assumption [Figalli et al. 2013]. Cross-curvature was also studied in [Sei 2013] for an application in statistics.

3. Background on Kähler geometry

In order to connect optimal transport with Kähler geometry, we will review some background on Kähler manifolds. We will only discuss what is needed for this work, and refer the reader to [Zheng 2000] for a more complete reference on complex geometry.

Given a smooth manifold[4] $\mathbb{X}$, an *almost complex structure* J is a smoothly varying endomorphism of $T\mathbb{X}$ which satisfies $J^2 = -\mathrm{Id}$. In this case, we say that the pair $(\mathbb{X}, J)$ is an *almost-complex manifold*. From the definition, it immediately follows that any almost-complex manifold must be even-dimensional and orientable.

We say that an almost complex manifold $\mathbb{X}$ is *complex* if it admits an atlas of holomorphic coordinate charts satisfying $z^i = u^i + \sqrt{-1}y^i$ such that $J\partial_{x^i} = \partial_{y^i}$ and $J\partial_{y^i} = -\partial_{x^i}$. In other words, around each point in $\mathbb{X}$, a complex manifold admits a local biholomorphism to a subset of $\mathbb{C}^n$ (in which J acts on tangent vectors as multiplication by $\sqrt{-1}$). If this can be done, we say that the almost complex structure is *integrable*. Due to a deep theorem of [Newlander and Nirenberg 1957], integrability of an almost complex structure J is equivalent to the vanishing of the so-called Nijenhuis tensor, which is defined as

$$N_J(\mathcal{X}, \mathcal{Y}) = -J^2[\mathcal{X}, \mathcal{Y}] + J[\mathcal{X}, J\mathcal{Y}] + J[J\mathcal{X}, \mathcal{Y}] - [J\mathcal{X}, J\mathcal{Y}].$$

Showing that this condition is necessary is relatively straightforward,[5] but it highly nontrivial to show that it is also sufficient.

We say that an almost complex structure is *compatible* with a Riemannian metric g if it satisfies $g(\mathcal{X}, \mathcal{Y}) = g(J\mathcal{X}, J\mathcal{Y})$ for all tangent vectors $\mathcal{X}$ and $\mathcal{Y}$. In this case, the triple $(\mathbb{X}, g, J)$ is said to be a *Hermitian* manifold. Furthermore, we say a Hermitian manifold is *Kähler* if J is integrable and the Kähler form $\omega = g(J\,\cdot\,, \cdot\,)$ is closed (i.e., $d\omega = 0$). This closedness has many important consequences for the geometry of Kähler metrics. Most importantly, in any set of holomorphic coordinates $\{z^i\}_{i=1}^n$, we can express the Kähler form as

$$\omega = \sqrt{-1}\frac{\partial^2\Phi}{\partial z^i\,\bar{\partial} z^j} dz^i \wedge d\bar{z}^j$$

for some strictly plurisubharmonic potential Φ. This leads to many important geometric properties, only a few of which we will explore here.

3.1. ***The curvature of Kähler manifolds.*** In this paper, we will study the curvature for a certain class of Kähler manifolds. As such, it is necessary to review some background on the curvature of Kähler metrics. We will specialize our focus to the curvature of the Levi-Civita connection on $\mathbb{X}$, which we denote as ∇. For non-Kähler Hermitian manifolds, there are several canonical connections which have distinct curvature tensors. Fortunately, all these connections coincide for Kähler manifolds. As such, we can unambiguously denote the curvature tensor as $\mathfrak{R}$, which is defined[6] as

$$\mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) = g(\nabla_{\mathcal{X}}\nabla_{\mathcal{Y}}\mathcal{Z} - \nabla_{\mathcal{Y}}\nabla_{\mathcal{X}}\mathcal{Z} - \nabla_{[\mathcal{X},\mathcal{Y}]}\mathcal{Z}, \mathcal{W}). \tag{4}$$

[4]Not to be confused with our notation for domains.

[5]The Nijenhuis tensor vanishes if the Lie bracket of any two holomorphic vectors is holomorphic, which is automatically true for complex structures.

[6]For this definition to be meaningful, we must extend each of the vectors to vector fields, but since the expression is tensorial, the choice of extension does not matter.

Apart from the usual symmetries of the Riemannian curvature, the curvature of a Kähler metric satisfies the following identity (when $\mathbb{X}$ is regarded as a (real) manifold):

$$\mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) = \mathfrak{R}(J\mathcal{X}, J\mathcal{Y}, \mathcal{Z}, \mathcal{W}). \tag{5}$$

After repeatedly applying this identity and using the other symmetries of the curvature tensor, it is possible to show that we can determine the entire curvature tensor from the values $\mathfrak{R}(\mathcal{X}, \overline{\mathcal{Y}}, \mathcal{Z}, \overline{\mathcal{W}})$, where $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}$ are holomorphic[7] vector fields and the overline represents conjugation.[8]

3.1.1. *Sectional and bisectional curvature.* Aside from the full curvature tensor, there are various notions of sectional and bisectional curvature on Kähler manifolds, which are important in the study of complex differential geometry.

As a preliminary, we first recall the definition of sectional curvature for Riemannian manifolds. Let $\mathcal{X}, \mathcal{Y}$ be nonparallel tangent vectors. The sectional curvature is defined as

$$K(\mathcal{X}, \mathcal{Y}) = \frac{\mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Y}, \mathcal{X})}{g(\mathcal{X}, \mathcal{X})g(\mathcal{Y}, \mathcal{Y}) - g(\mathcal{X}, \mathcal{Y})^2}.$$

It is a classic theorem in Riemannian geometry that the sectional curvature completely determines the entire curvature tensor, which can be proven using the polarization formula and a careful application of the Bianchi identity. The sectional curvature is a fundamental concept in Riemannian geometry, and many theorems depend on either upper or lower bounds for it. Furthermore, the assumption that the sectional curvature is nonnegative greatly restricts the topology and geometry of a given Riemannian manifold.

For a Kähler manifold, there are several notions of curvature closely related to the sectional curvature. One natural type of sectional curvature on a Kähler manifold is the *holomorphic sectional curvature*. For a tangent vector $\mathcal{X} \in T\mathbb{X}$, this is defined as

$$\mathfrak{H}(\mathcal{X}) = \frac{\mathfrak{R}(\mathcal{X}, J\mathcal{X}, J\mathcal{X}, \mathcal{X})}{\|\mathcal{X}\|^4}.$$

Similarly to the sectional curvature, the holomorphic sectional curvature determines the entire curvature tensor of a Kähler manifold; see [Ballmann 2006, Proposition 4.51].

Of particular interest are Kähler manifolds whose holomorphic sectional curvature is constant c. In this case, the polarization formula can be used to show that such a Kähler manifold $\mathbb{X}$ is a Hermitian symmetric space whose curvature satisfies

$$\mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}) = \frac{c}{4}\begin{pmatrix} g(\mathcal{X}, \mathcal{Z})g(\mathcal{Y}, \mathcal{W}) - g(\mathcal{X}, \mathcal{W})g(\mathcal{Y}, \mathcal{Z}) + g(\mathcal{X}, J\mathcal{Z})g(\mathcal{Y}, J\mathcal{W}) \\ -g(\mathcal{X}, J\mathcal{W})g(\mathcal{Y}, J\mathcal{Z}) + 2g(\mathcal{X}, J\mathcal{Y})g(\mathcal{Z}, J\mathcal{W}) \end{pmatrix}.$$

Spaces with constant holomorphic sectional curvature serve as the Kähler analogues of manifolds of constant sectional curvature. It is worth noting that when the complex dimension is greater than 1, a complex manifold cannot have constant sectional curvature (unless it is flat), so the definition of a complex space form is not a Kähler manifold with constant sectional curvature, but rather a Kähler manifold with

[7] In other words, vectors which satisfy $J\mathcal{X} + \sqrt{-1}\mathcal{X} = 0$ in a holomorphic coordinate chart.

[8] In fact, this shows that $\mathfrak{R}(\overline{\mathcal{X}}, \overline{\mathcal{Y}}, \mathcal{Z}, \mathcal{W})$, $\mathfrak{R}(\mathcal{X}, \overline{\mathcal{Y}}, \mathcal{Z}, \mathcal{W})$, and $\mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W})$ all vanish on a Kähler manifold.

constant holomorphic sectional curvature. The sectional curvature of such a metric (with $n > 1$) ranges between c and $\frac{1}{4}c$, so the metric is quarter-pinched from the view of Riemannian geometry.

Despite the similarity between the holomorphic sectional curvature and the sectional curvature, the former is a much more subtle invariant than the latter. For instance, nonnegative holomorphic sectional curvature does not imply nonnegative Ricci curvature, but does imply nonnegative scalar curvature [Ni and Zheng 2018].

As such, it is worthwhile to consider other curvature quantities which more directly control the geometry of a Kähler manifold. One such example is the *bisectional curvature*,[9] which was introduced in [Goldberg and Kobayashi 1967]. For two unit vectors $\mathcal{X}$ and $\mathcal{Y}$, it is defined as

$$\mathfrak{B}(\mathcal{X}, \mathcal{Y}) = \mathfrak{R}(\mathcal{X}, J\mathcal{X}, J\mathcal{Y}, \mathcal{Y}).$$

The reason that this is known as the bisectional curvature is that it satisfies the identity

$$\mathfrak{B}(\mathcal{X}, \mathcal{Y}) = \mathfrak{R}(\mathcal{X}, J\mathcal{Y}, J\mathcal{Y}, \mathcal{X}) + \mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Y}, \mathcal{X}),$$

which can be proven using the Bianchi identity.

A metric is said to have nonnegative bisectional curvature if $\mathfrak{B}(\mathcal{X}, \mathcal{Y}) \geq 0$ for all vectors $\mathcal{X}$ and $\mathcal{Y}$. Nonnegative bisectional curvature is a weaker condition than nonnegative sectional curvature (as the bisectional curvature is the sum of two sectional curvatures) but still provides very strong control over the geometry of a Kähler manifold.

There are several further curvatures of interest. For instance, it is possible to consider the bisectional curvature when it is restricted to unit tangent vectors $\mathcal{X}$ and $\mathcal{Y}$ satisfying $g(\mathcal{X}, \mathcal{Y}) = g(\mathcal{X}, J\mathcal{Y}) = 0$. This is known as the orthogonal bisectional curvature. We say that a Kähler manifold has (*NOB*) (nonnegative orthogonal bisectional curvature) if for all unit tangent vectors $\mathcal{X}$ and $\mathcal{Y}$ satisfying $g(\mathcal{X}, \mathcal{Y}) = g(\mathcal{X}, J\mathcal{Y}) = 0$ we have $\mathfrak{B}(\mathcal{X}, \mathcal{Y}) \geq 0$.

In this paper, we will need to consider a curvature tensor we call the *antibisectional curvature*. For totally real vectors[10] $\mathcal{X}$ and $\mathcal{Y}$, we define this to be

$$\mathfrak{A}(\mathcal{X}, \mathcal{Y}) = \mathfrak{R}(\mathcal{X}, J\mathcal{Y}, J\mathcal{Y}, \mathcal{X}) - \mathfrak{R}(\mathcal{X}, \mathcal{Y}, \mathcal{Y}, \mathcal{X}).$$

Similarly, we define the orthogonal antibisectional curvature (denoted by $\mathfrak{O}\mathfrak{A}$) to be the restriction of the antibisectional curvature to vectors $\mathcal{X}, \mathcal{Y}$ satisfying $g(\mathcal{X}, \mathcal{Y}) = g(\mathcal{X}, J\mathcal{Y}) = 0$. More precisely,

$$\mathfrak{O}\mathfrak{A}(\mathcal{X}, \mathcal{Y}) := \mathfrak{A}(\mathcal{X}, \mathcal{Y})|_{\{\mathcal{X}, \mathcal{Y} \mid g(\mathcal{X}, \mathcal{Y}) = g(\mathcal{X}, J\mathcal{Y}) = 0\}}.$$

The reason for the term "antibisectional" curvature is that $\mathfrak{A}$ and $\mathfrak{B}$ differ only in that we subtract rather than add the sectional curvatures. However, these curvatures are very different. For instance, the bisectional curvature is J-invariant, in that we can multiply either $\mathcal{X}$ or $\mathcal{Y}$ by J and get the same result.

[9] This is more commonly called the *holomorphic bisectional curvature*, but we will omit the "holomorphic" for the sake of exposition.

[10] In other words, vectors whose imaginary component is zero within a particular holomorphic chart.

On the other hand, $\mathfrak{A}$ changes sign if we multiply one of the vectors by J.[11] As such, it takes some care to define nonnegative orthogonal antibisectional curvature.

We say that a Kähler metric on a domain in $\mathbb{C}^n$ has (*NAB*) if, for all totally real vectors $\mathcal{X}, \mathcal{Y}$,

$$\mathfrak{A}(\mathcal{X}, \mathcal{Y}) \geq 0.$$

Similarly, we say that a Kähler metric has (*NOAB*) if, for all orthogonal totally real vectors $\mathcal{X}, \mathcal{Y}$,

$$\mathfrak{O}\mathfrak{A}(\mathcal{X}, \mathcal{Y}) \geq 0.$$

Due the previous discussion, we define nonnegativity for antibisectional curvature when restricted to totally real vectors. This definition inherently relies on a canonical decomposition of $T\mathbb{X}$ into real and imaginary vectors (i.e., an embedding into $\mathbb{C}^n$). For the spaces of interest in this paper, this can be done in a natural way. However, for more general Kähler manifolds, formulating nonnegative antibisectional curvature is less clear.

It is worth observing that if a Kähler manifold $\mathbb{X}$ has constant holomorphic sectional curvature, then the orthogonal antibisectional curvature identically vanishes. In fact, if we require that $\mathfrak{O}\mathfrak{A}(\mathcal{X}, \mathcal{Y}) \geq 0$ for all orthogonal vectors $\mathcal{X}$ and $\mathcal{Y}$, then the polarization formula shows that Hermitian symmetric spaces are the only spaces satisfying this property.[12]

3.2. ***Positively curved Kähler metrics.*** One question of considerable interest in complex geometry is to understand complete Kähler metrics with various nonnegativity properties. Most famously, Frankel conjectured that if a compact Kähler manifold has positive holomorphic bisectional curvature, it is biholomorphic to the complex projective space $\mathbb{CP}^n$. This conjecture was independently proven in [Mori 1979; Siu and Yau 1980].

For compact Kähler manifolds, it is possible to obtain this result under weaker curvature assumptions. For instance, all compact Kähler manifolds with positive orthogonal bisectional curvature are biholomorphic to $\mathbb{CP}^n$; see [Chen 2007; Feng et al. 2017; Gu and Zhang 2010]. Furthermore, all compact irreducible Kähler manifolds with nonnegative isotropic curvature are either Hermitian symmetric or else biholomorphic to $\mathbb{CP}^n$ [Seshadri 2009]. For complex surfaces, nonnegative orthogonal bisectional curvature is equivalent to nonnegative isotropic curvature[13] [Li and Ni 2019] so the previous result gives a classification of such surfaces.

For noncompact manifolds, it is natural to ask whether similar results hold. The most famous conjecture in this direction is Yau's uniformization conjecture [1994], which states that any complete irreducible noncompact Kähler metric with nonnegative bisectional curvature is biholomorphic to $\mathbb{C}^n$. Although the full conjecture is still open, Liu [2019] proved it under certain volume growth assumptions.

Although these results are not directly related to the work in this paper, they provide much of the intuition for how positive curvature of a Kähler metric imposes strong restrictions on the geometry of

---

[11] This was pointed out to us by Fangyang Zheng.

[12] Thanks to Fangyang Zheng for this observation.

[13] In higher dimensions, nonnegative isotropic curvature is a stronger assumption than nonnegative orthogonal bisectional curvature.

that manifold. In this spirit, by drawing a connection between the curvature of Kähler metrics and the MTW tensor in optimal transport, we hope to provide strong restrictions on cost functions which have nonnegative MTW tensor. We plan to develop this idea further in future work.

4. Hessian manifolds and the Sasaki metric

In order to interpret the MTW tensor as a complex-geometric curvature, we study the Sasaki metric, which is an almost-Hermitian metric on the tangent bundle of a Riemannian manifold. We discuss some background on this metric, focusing on the case of Hessian manifolds, in which case the Sasaki metric is Kähler.

4.1. ***The Sasaki metric on the tangent bundle.*** On a Riemannian manifold $(\mathcal{M}, g)$ with a flat[14] affine connection D, the tangent bundle naturally inherits a Hermitian structure $(T\mathcal{M}, g^D, J^D)$ [Dombrowski 1962]. The metric g^D is known as the *Sasaki metric* and the complex structure J^D is called the *canonical complex structure*. For completeness, we present a brief overview of this construction. For a more complete reference, we refer to [Satoh 2007].

Since D is flat, we can find local coordinates $\{u^i\}_{i=1}^n$ on $\mathcal{M}$ in which the Christoffel symbols of D vanish. Using these coordinates, define smooth functions $v^1, \ldots, v^n$ on the tangent bundle $T\mathcal{M}$ by $v^j(\mathcal{X}) = \mathcal{X}^j$ for a vector $\mathcal{X} = \mathcal{X}^i \partial_{u^i}$. The collection of functions $\{(u^i, v^i)\}_{i=1}^n$ then forms local coordinates for $T\mathcal{M}$. Then, for a tangent vector $\xi \in T_u\mathcal{M}$ (which we consider as a point in the tangent bundle $T\mathcal{M}$) and a tangent vector $\mathcal{X} = \mathcal{X}^i \partial_{u^i} \in T_u\mathcal{M}$, we can define vertical and horizontal lifts of $\mathcal{X}$ at ξ, denoted by $\mathcal{X}_\xi^V$ and $\mathcal{X}_\xi^H$, respectively. These are elements of $T_\xi(T\mathcal{M})$, which are defined as

$$\mathcal{X}_\xi^V = \mathcal{X}^i \partial_{v^i}, \quad \mathcal{X}_\xi^H = \mathcal{X}^i \partial_{u^i}. \tag{6}$$

This yields a decomposition of $T_\xi(T\mathcal{M})$ into horizontal and vertical subspaces, which depends on the choice of connection D:

$$T_\xi(T\mathcal{M}) = H_\xi(T\mathcal{M}) \oplus V_\xi(T\mathcal{M}).$$

As such, there is a natural identification $H_\xi(T\mathcal{M}) \cong V_\xi(T\mathcal{M}) \cong T_u\mathcal{M}$, which we use to construct the Sasaki metric [Satoh 2007, Definition 2.1].

Definition (the Sasaki metric and canonical complex structure). Let $(\mathcal{M}^n, g)$ be a Riemannian manifold with a flat affine connection D. For $\mathcal{X}, \mathcal{Y} \in T_u\mathcal{M}$ and $\xi \in T\mathcal{M}$ with $\xi = (u, v)$ in bundle coordinates, the canonical complex structure J^D is defined as

$$J^D \mathcal{X}_\xi^H = \mathcal{X}_\xi^V, \quad J^D \mathcal{X}_\xi^V = -\mathcal{X}_\xi^H.$$

Furthermore, the Sasaki metric g^D is defined as

$$\tilde{g}^D(\mathcal{X}_\xi^H, \mathcal{Y}_\xi^H) = \tilde{g}^D(\mathcal{X}_\xi^V, \mathcal{Y}_\xi^V) = g(\mathcal{X}, \mathcal{Y}), \quad \tilde{g}^D(\mathcal{X}_\xi^H, \mathcal{Y}_\xi^V) = 0.$$

[14] It is possible to define the Sasaki metric for arbitrary connections, but that will not be necessary for this paper.

This induces a Hermitian structure on $T\mathcal{M}$, which depends on both the choice of metric and flat connection on $\mathcal{M}$. To see that this is indeed a Hermitian manifold (and not merely almost Hermitian), we rely on the following result.

Theorem 3 [Dombrowski 1962]. *Let $(\mathcal{M}, g)$ be a Riemannian manifold with an affine connection D. The almost-Hermitian manifold $(T\mathcal{M}, g^D, J^D)$ satisfies the following*:

(1) *The almost-complex structure J^D is integrable whenever the connection D is flat.*

(2) *$(T\mathcal{M}, g^D, J^D)$ is Kähler if and only if D and D^* are both flat connections, which further implies that g is a Hessian metric.*

4.2. ***Hessian manifolds.*** We are primarily interested in the case where $T\mathcal{M}$ is Kähler, for which we must study Hessian manifolds (also known as *affine-Kähler* manifolds, due to the parallel with Kähler geometry). There are two equivalent definitions for such manifolds; with the former definition primarily used in differential geometry and the latter primarily used in information geometry.

Definition (Hessian manifold: differential-geometric). A Riemannian manifold $(\mathcal{M}, g)$ is Hessian if there is an atlas of local coordinates $\{u^i\}_{i=1}^n$ so that for each coordinate chart there is a convex potential Ψ such that

$$g_{ij} = \frac{\partial^2 \Psi}{\partial u^i \, \partial u^j}.$$

Furthermore, the transition maps between these coordinate charts are affine (i.e., $\mathcal{M}$ is an affine manifold).

Definition (Hessian manifold: information-geometric). A Riemannian manifold $(\mathcal{M}, g)$ is said to be *Hessian* if it admits dually flat connections. That is to say, it admits two flat (torsion- and curvature-free) connections D and D^* satisfying

$$\mathcal{X}(g(\mathcal{Y}, \mathcal{Z})) = g(D_{\mathcal{X}}\mathcal{Y}, \mathcal{Z}) + g(\mathcal{Y}, D^*_{\mathcal{X}}\mathcal{Z}) \tag{7}$$

for all vector fields $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$. Because of these dual flat connections, a Hessian manifold is often said to be *dually flat*.

Although these definitions initially appear different, they are actually equivalent. If we choose an atlas of coordinate charts (we will abuse notation and refer to this chart as u) in which the metric g is of Hessian form, we can induce a flat connection D by differentiation with respect to the u-coordinates. The requirement that the transition maps be affine is exactly what is necessary for this connection to be well-defined when we switch coordinates.

Before moving on, we take a moment to discuss the properties of the connection D in more detail. Firstly, by definition we have that the Christoffel symbols of D vanish in the u-coordinates. As a result, the D-geodesic equations

$$\frac{d^2u^i}{ds^2} + \Gamma^i_{jk}\frac{du^j}{ds}\frac{du^k}{ds} = 0$$

simplify to the equations

$$\frac{d^2u^i}{ds^2} = 0.$$

As a result, D-geodesics correspond to straight lines in the u-coordinates. It is worth noting that these geodesics are distinct from the geodesics with respect to the Levi-Civita connection.

We now turn our attention to the dual connection D^*. We can induce the dual connection by differentiation with respect to the so-called dual coordinates θ, which are defined as

$$\theta^i := \frac{\partial \Psi}{\partial u^i}. \tag{8}$$

It is a straightforward calculation to show that the connection induced by the u-coordinates and the connection induced by the θ-coordinates are indeed dual, in the sense of (7). For more information, see Chapter 2 of [Shima 2007]. Similarly to the situation for the primal connection, the Christoffel symbols of D^* vanish in the θ-coordinates and D^*-geodesics correspond to straight lines in the θ-coordinates. Furthermore, we can define D^*-convexity for subsets of $\mathcal{M}$ in terms of whether a subset entirely contains the D^*-geodesics between its points. This can be extended to define strict convexity as well. Strict convexity will be important in Proposition 8, so we draw attention to it here.

In the dual coordinates, the metric is also of Hessian form, where the potential is the Legendre transform Ψ^*, defined as

$$\Psi^*(\theta) = \sup_{u \in \mathcal{M}} \langle \theta, u \rangle - \Psi(u).$$

When Ψ is a convex function, Ψ^* is as well. In this case, we say that Ψ and Ψ^* are Legendre duals.[15] For further details on this correspondence, we refer the reader to [Shima 2007, Chapter 2].

There are topological and geometric obstructions for a given Riemannian manifold to admit a Hessian structure. In dimensions 4 and higher, there are local curvature obstructions as well; see [Amari and Armstrong 2014]. As all of the manifolds of interest in this paper are open domains in $\mathbb{R}^n$ (which admit a global coordinate chart), we can construct Hessian metrics simply by choosing a convex potential.

4.3. ***The curvature of the Sasaki metric.*** We now calculate the curvature of a Kähler Sasaki metric. To do so, we use the curvature formulas for a general Sasaki metric [Satoh 2007, Proposition 2.3] and then simplify them using the dually flat structure. Applying Satoh's Proposition 2.3 in the case where D is a flat connection, we have the following.

Proposition 4. *Let $(\mathcal{M}, g, D)$ be an affine manifold with flat connection D and Levi-Civita connection ∇. Let $\mathfrak{R}_{g^D}$ be the Riemannian curvature tensor of the Sasaki metric g^D on $T\mathcal{M}$. For vectors $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}, \xi \in T_u\mathcal{M}$,*

$$\mathfrak{R}_{g^D}(\mathcal{Z}^H_\xi, \mathcal{W}^H_\xi, X^H_\xi, \mathcal{Y}^H_\xi) = R^\nabla_g(\mathcal{Z}, \mathcal{W}, \mathcal{X}, \mathcal{Y}),$$

$$\mathfrak{R}_{g^D}(\mathcal{Z}^V_\xi, \mathcal{W}^V_\xi, \mathcal{X}^V_\xi, \mathcal{Y}^V_\xi) = -\frac{1}{4}\sum_i [(D_{e_i} g)(\mathcal{X}, \mathcal{Z})(D_{e_i} g)(\mathcal{Y}, \mathcal{W}) - (D_{e_i} g)(\mathcal{Y}, \mathcal{Z})(D_{e_i} g)(\mathcal{X}, \mathcal{W})],$$

$$\mathfrak{R}_{g^D}(\mathcal{Z}^H_\xi, \mathcal{W}^V_\xi, \mathcal{X}^V_\xi, \mathcal{Y}^V_\xi) = \mathfrak{R}_{g^D}(\mathcal{Z}^H_\xi, \mathcal{W}^V_\xi, \mathcal{X}^H_\xi, \mathcal{Y}^H_\xi) = 0,$$

$$\mathfrak{R}_{g^D}(\mathcal{Z}^H_\xi, \mathcal{W}^V_\xi, \mathcal{X}^H_\xi, \mathcal{Y}^V_\xi) = -\frac{1}{2}(D^2_{\mathcal{X}\mathcal{Z}} g)(\mathcal{Y}, \mathcal{W}) - \frac{1}{2}(D_{\gamma(\mathcal{X}, \mathcal{Z})} g)(\mathcal{Y}, \mathcal{W}) + \frac{1}{4}\sum_i (D_{\mathcal{X}} g)(\mathcal{W}, e_i) \cdot (D_{\mathcal{Z}} g)(\mathcal{Y}, e_i).$$

[15] This is indeed a duality: for a convex function Ψ, $\Psi = \Psi^{**}$.

Here, $\{e_i\}$ *is an orthonormal basis of* $T_u\mathcal{M}$ *and* γ^D *is the difference between* D *and the Levi-Civita connection on* $\mathcal{M}$:

$$\gamma^D(\mathcal{X}, \mathcal{Y}) = D_{\mathcal{X}}\mathcal{Y} - \nabla_{\mathcal{X}}\mathcal{Y}.$$

When $(\mathcal{M}, g, D)$ is dually flat (i.e., Hessian), the situation simplifies further. To ease the computations, recall that we are working in coordinates $\{u^i\}_{i=1}^n$ where the Christoffel symbols of D vanish. Doing so, we find the following identities:

(1) The Riemannian metric g is given by the Hessian of a convex potential Ψ.

(2) In the induced coordinates $\{(u^i, v^i)\}_{i=1}^n$ on the tangent bundle $(T\mathcal{M}, g^D, J^D)$, the complex structure can be written as $J^D\partial_{u^i} = \partial_{v^i}$ and $J^D\partial_{v^i} = -\partial_{u^i}$. As such, the coordinate chart $(u^1, v^1, \ldots, u^n, v^n)$ is biholomorphic to an open set in $\mathbb{C}^n$ under the natural identification.

(3) There are simple expressions for the Riemannian curvature and the Christoffel symbols of the Levi-Civita connection.

(a) The Riemannian curvature of the $(\mathcal{M}, g)$ is (see [Shima 2007, Proposition 3.2])

$$R_g^{\nabla}(\partial_{u^i}, \partial_{u^j}, \partial_{u^k}, \partial_{u^l}) = -\tfrac{1}{4}\Psi^{pq}(\Psi_{jlp}\Psi_{ikq} - \Psi_{ilp}\Psi_{jkq}).$$

(b) The Christoffel symbols of the Levi-Civita connection satisfy the identity

$$\Gamma_{ijk} = \tfrac{1}{2}\Psi_{ijk}, \quad \Gamma_{ji}^k = \tfrac{1}{2}\Psi_{ijm}\Psi^{km}.$$

(4) Using these formulas for the Christoffel symbols, we obtain a simple formula for $D_{\gamma^D(\mathcal{X},\mathcal{Z})}$ for two vector fields $\mathcal{X} = \mathcal{X}^i\partial_{u^i}$ and $\mathcal{Z} = \mathcal{Z}^k\partial_{u^k}$:

$$D_{\gamma^D(\mathcal{X},\mathcal{Z})} = -\mathcal{X}^i\mathcal{Z}^k\Gamma_{ik}^r D_{\partial_{u^r}} = -\mathcal{X}^i\mathcal{Z}^k\Psi_{iks}\Psi^{sr}D_{\partial_{u^r}}$$

Combining these identities with the curvature formulas for the Sasaki metric, we find the following proposition.

Proposition 5 (curvature of a Kähler Sasaki metric). *Let* $(\mathcal{M}, g, D)$ *be a Hessian manifold. The Riemannian curvature of the Sasaki metric on* $(T\mathcal{M}, g^D, J^D)$ (*in the* (u, v)-*coordinates defined on page* 16) *is*

$$\mathfrak{R}_{g^D}(\partial_{u^i}, \partial_{u^j}, \partial_{u^k}, \partial_{u^l}) = \mathfrak{R}_{g^D}(\partial_{v^i}, \partial_{v^j}, \partial_{v^k}, \partial_{v^l}) = -\tfrac{1}{4}\Psi^{rs}(\Psi_{jlr}\Psi_{iks} - \Psi_{ilr}\Psi_{jks}), \tag{9}$$

$$\mathfrak{R}_{g^D}(\partial_{u^i}, \partial_{v^j}, \partial_{u^k}, \partial_{v^l}) = -\tfrac{1}{2}\Psi_{ijkl} + \tfrac{1}{4}(\Psi_{iks}\Psi^{sr}\Psi_{jlr}) + \tfrac{1}{4}(\Psi^{sr}\Psi_{jks}\Psi_{ilr}). \tag{10}$$

Furthermore, when stated in terms of holomorphic vectors, the curvature of $T\mathcal{M}$ *satisfies the identity*

$$(\mathfrak{R}_{g^D})_{i\bar{j}k\bar{l}} = \mathfrak{R}_{g^D}(\partial_{u^i}, \partial_{v^j}, \partial_{u^k}, \partial_{v^l}) - \mathfrak{R}_{g^D}(\partial_{u^i}, \partial_{u^j}, \partial_{u^k}, \partial_{u^l}) = -\tfrac{1}{2}\Psi_{ijkl} + \tfrac{1}{2}\Psi^{rs}\Psi_{iks}\Psi_{jlr}. \tag{11}$$

We remark that for a Hessian manifold, Shima [2007] defined the Hessian curvature to be the negative of formula (11). We will not use this convention and instead work in terms of complex geometry.

4.4. ***The geometry when $\mathcal{M}$ is a domain.*** The previous calculation provides the curvature of $T\mathcal{M}$ for an arbitrary Hessian manifold. However, in the special case where $\mathcal{M}$ is a domain in Euclidean space and g is induced by a global potential Ψ, it is possible to construct this Kähler manifold explicitly as a subset of $\mathbb{C}^n$.

In this case, we consider the associated global coordinates $\{u\}_{i=1}^n$ as being defined on domain in $\mathbb{R}^n$. By a slight abuse of notation, we will denote the domain of the u-coordinates as $\mathcal{M}$. We can then write $z^i = u^i + \sqrt{-1}v^i$ as the standard holomorphic coordinate on $\mathbb{C}^n$, and $T\mathcal{M}$ can be identified with the domain

$$T\mathcal{M} = \{(u, v) \mid u \in \mathcal{M}, v \in \mathbb{R}^n\} \subset \mathbb{C}^n.$$

This class of domains are known as *tube domains* and have been studied in various contexts. For an introduction on these spaces, we refer the reader to page 41 of [Hörmander 1973] and for a more detailed study of their geometry, we refer the reader to [Yang 1982].

As for the associated Kähler metric, we can write the Kähler form as

$$\omega = \sqrt{-1}\Psi_{ij} dz^i \wedge d\bar{z}^j$$

and the Sasaki metric as

$$g = \begin{pmatrix} \Psi_{ij}(u) & 0 \\ 0 & \Psi_{ij}(u) \end{pmatrix}.$$

As can be seen from these formulas, the Kähler Sasaki metric is translation symmetric in its fibers (since Ψ does not depend on the fiber coordinates v). However, from (9) we can see that the fiber directions are not "flat" unless the underlying Hessian manifold $\mathcal{M}$ is Riemannian curvature free.

5. Optimal transport and complex geometry

With the background concluded, we can now state the central results of this paper, which relate the regularity theory of optimal transport to the complex geometry of the Sasaki metric.

5.1. ***The MTW tensor and the curvature of $T\mathcal{M}$.*** When $c : X \times Y \to \mathbb{R}$ is a Ψ-cost (as in the definition on page 398), Ma, Trudinger and Wang observed that the MTW tensor takes the form

$$\mathfrak{S}_{(x,y)}(\xi, \eta) = (\Psi_{ijp}\Psi_{rsq}\Psi^{pq} - \Psi_{ijrs})\Psi^{rk}\Psi^{sl}\xi^i\xi^j\eta^k\eta^l. \tag{12}$$

In this formula, k and l are summed over, despite the double superscript resulting from the vector-covector ambiguity.

To make the connection between (11) and (12) precise, we do the following. Firstly, we induce $\mathcal{M}$ with the structure of a Hessian manifold. To do so, we use Ψ as a potential for a Riemannian metric and let D be the flat connection induced by differentiation with respect to the u-coordinates. This then induces $(T\mathcal{M}, g^D, J^D)$ with a Kähler metric. Secondly, given a tangent vector $\xi \in T_x X$ and a point $y \in Y$, we induce a tangent vector (also denoted by ξ) in $T_{x-y}\mathcal{M}$ by shifting the base by y and leaving the components unchanged. We also use this same construction to induce cotangent vectors in $T^*_{x-y}\mathcal{M}$, given a point $x \in X$ and a cotangent vector $\eta \in T^*_y Y$.

After doing so, by comparing (9), (10) and (12), we obtain the following result.

Theorem 6. *Let X and Y be open sets in $\mathbb{R}^n$ and c be a Ψ-cost. Furthermore, let (ξ, η) be a vector-covector pair in $T_x X \times T_y^* Y$ such that the associated vector-covector pair on $T_{x-y}\mathcal{M} \times T_{x-y}^*\mathcal{M}$ is orthogonal. Finally, let ζ be an arbitrary vector in $T_{x-y}\mathcal{M}$.*

Then the MTW tensor satisfies the identity

$$\mathfrak{S}(\xi, \eta) = 2\mathfrak{R}_{g^D}(\xi_\zeta^H, (\eta^\sharp)_\zeta^V, \xi_\zeta^H, (\eta^\sharp)_\zeta^V) - 2\mathfrak{R}_{g^D}((\eta^\sharp)_\zeta^H, \xi_\zeta^H, (\eta^\sharp)_\zeta^H, \xi_\zeta^H). \tag{13}$$

Here, $\mathfrak{R}_{g^D}$ is the curvature of the Sasaki metric on $(T\mathcal{M}, g^D, J^D)$ after sharping η (recall that the η in the MTW tensor is a covector with $\eta(\xi) = 0$). Furthermore, the cross curvature satisfies the same identity when we allow ξ and η to be an arbitrary vector-covector pair.

Here, due to the symmetries of the Kähler Sasaki metric, the choice of ζ is arbitrary. We will discuss this fact in Section 7. Note that it is important to be careful with the indices in the previous result.[16]

Recalling our previous discussion on the curvature of Kähler metrics, if we consider $T\mathcal{M}$ as a tube domain, then the right-hand side of (13) is twice the orthogonal antibisectional curvature, which implies the following corollary.

Remark 7. The MTW tensor for a Ψ-cost is nonnegative if and only if $T\mathcal{M}$ has (NOAB) on the set $T(X - Y) \subset T\mathcal{M}$.

5.2. *Relative c-convexity of sets and dual geodesic convexity.*

In order to establish regularity for optimal transport (as done in Theorem 2), not only is it necessary to assume that the MTW tensor is nonnegative, there are also assumptions about the relative c-convexity of the supports of μ and ν. For Ψ-costs, there is a natural geometric interpretation for this notion, which we establish here.

Proposition 8. *For a Ψ-cost, a set Y is c-convex relative to X if and only if, for all $x \in X$, the set $x - Y \subset \mathcal{M}$ is geodesically convex with respect to the dual connection D^*.*

Proof. Recall that for $x \in X$, a c-segment in Y is the curve $c\text{-}\exp_x(\ell)$ for some line segment ℓ and a set Y is c-convex relative to a set X if, for all $x \in X$, Y contains all c-segments between points in Y. For a Ψ-cost, relative c-convexity corresponds with geodesic convexity with respect to the dual connection[17] D^*.

We now apply (8) to see

$$-c_{i,} = -\Psi_i(x - y) = -\theta^i(x - y),$$

where $\theta^i(x - y)$ is the point $x - y \in \mathcal{M}$ in terms of the dual coordinates θ^i. By the definition on page 402, c-segments correspond to straight lines in the θ-coordinates. From the discussion after (8), this shows that c-segments are geodesics with respect to the dual connection D^*.

As such, a set Y contains all its c-segments if and only if, for all $x \in X$, $x - Y$ contains all of D^* geodesics, which is another way of saying that $x - Y$ is geodesically convex with respect to D^*. □

[16] A previous version of this paper [Khan and Zhang 2019] mistakenly switched the roles of j and k, leading to an incorrect claim of a correspondence of the MTW tensor to the orthogonal bisectional curvature.

[17] Recall that the dual connection D^* satisfies (7), where D is the flat connection induced by differentiation with respect to the u-coordinates.

An analogous result holds for relative c-convexity of X relative to Y. Combining the previous two results, we can restate Theorem 2 in this new language.

Theorem. *Suppose X and Y are smooth bounded domains in $\mathbb{R}^n$ and that $d\mu$ and $d\nu$ are smooth probability densities supported on X and Y, respectively, bounded away from zero and infinity on their supports. Consider a Ψ-cost for some convex function $\Psi : \mathcal{M} \to \mathbb{R}$ and suppose the following conditions hold*:

(1) *Ψ is C^4 and locally strongly convex (i.e., its Hessian is positive definite).*

(2) *For all $x \in X$, $x - Y \subset \mathcal{M}$ is strictly geodesically convex with respect to the dual connection D^*.*

(3) *For all $y \in Y$, $X - y \subset \mathcal{M}$ is strictly geodesically convex with respect to the dual connection D^*.*

(4) *The Kähler manifold $(T\mathcal{M}, g^D, J^D)$ has (NOAB) on the subset $T(X - Y)$.*

Let $\mathbb{T}_{\mathrm{U}}$ be the c-optimal transport map carrying μ to ν as in Theorem 1. Then $\mathrm{U} \in C^\infty(\overline{X})$ and $\mathbb{T}_{\mathrm{U}} : \overline{X} \to \overline{Y}$ is a smooth diffeomorphism.

We should note that for many Ψ-costs of interest, Ψ will not be uniformly strongly convex over its entire domain. This is no issue for the regularity theory, as we will restrict our attention to bounded sets X and Y, so that $X - Y$ is precompact. As such, Ψ will be strongly convex on $X - Y$.

5.3. ***Information-geometric interpretation.*** The previous results provide new interpretations of the regularity theory, but we can also use this approach to find new examples of cost functions which satisfy MTW(0). Before doing so, we will briefly discuss information geometry, which is where many of these examples originate.

Information geometry studies the geometry of parametrized statistical models. For a more complete background, we refer readers to [Amari 2016]. All of the examples in this paper are constructed from exponential families, so we will focus only on the information geometry of exponential families.

Given a sample space S (about which we make no assumptions), an exponential family is a parametrized family of probability distributions whose probability density functions are of the form

$$f_S(s \mid u) = h(s)\exp(\boldsymbol{\eta}(u) \cdot \boldsymbol{T}(s) - A(u)) \tag{14}$$

for some known functions $h : S \to \mathbb{R}$, $\boldsymbol{\eta} : U \to \mathbb{R}^n$, $\boldsymbol{T} : S \to \mathbb{R}^n$, and $A : U \to \mathbb{R}$. Here, $u \in U$ serve as parameters of the distributions, and they are generally defined on an open domain U in $\mathbb{R}^n$. Note that $\boldsymbol{\eta}$ should not be confused with the covector notation; it is instead a function of the parameters. This class of statistical models includes many commonly used parametrized families, such as the univariate normal and multinomial distributions (both of which we consider in the application section).

An exponential family is said to be in *canonical form* if $\boldsymbol{\eta}(u) = u$, in which case u is said to be the *natural parameters*. In this case, we can rewrite (14) as

$$f_S(s \mid u) = h(s)\exp(u \cdot \boldsymbol{T}(s) - \Psi(u)). \tag{15}$$

Here, $\Psi(u)$ is known as the *log-partition function*, and serves to preserve the probability normalization (i.e., $\int_S f_S(s \mid u)\, ds = 1$). When an exponential family is written in terms of its natural parameters, $\Psi(u)$ is a convex function whose domain is also convex.

5.3.1. *The Fisher metric.* Given any parametrized statistical model (not just an exponential family), there is a canonical Riemannian metric that can be induced on the parameter space. This metric, known as the *Fisher metric*, takes the form

$$g\left(\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j}\right) = \int_S \frac{\partial \log f(s \mid u)}{\partial u^i} \frac{\partial \log f(s \mid u)}{\partial u^j} f(s \mid u)\, ds. \tag{16}$$

There are several reasons to consider the Fisher metric as a canonical metric, and this is just one of several equivalent definitions for it. However, a more complete discussion of this topic would take us too far from the central aim of the this project. For more information, we refer the reader to the paper by Ay, Jost, Vân Lê, and Schwachhöfer [Ay et al. 2015].

For an exponential family in canonical form, the Fisher metric takes a special form. More precisely, it can be written as a Hessian metric

$$g\left(\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j}\right) = \frac{\partial^2}{\partial u^i \partial u^j} \Psi(u), \tag{17}$$

where $\Psi(u)$ is the log-partition function (which is guaranteed to be convex). From this, there is a natural statistical reason to consider Hessian manifolds, which we can further use to construct cost functions for optimal transport.

This paper is not the first work to consider using the log-partition function to find interesting Ψ-costs. This construction was first introduced in [Pal 2017], who developed some of the optimal transport theory for costs of this form. We thank the reviewer for bringing this paper to our attention.

5.3.2. *The case of $\mathcal{D}_\Psi^{(\alpha)}$-divergences.* Although our main results are stated in terms of Ψ-costs, they also hold (with minor modifications) for cost functions that are $\mathcal{D}_\Psi^{(\alpha)}$-divergences, which were previously studied by the second author [Zhang 2004] in the context of information geometry.

Definition ($\mathcal{D}_\Psi^{(\alpha)}$-divergence). Let $\Psi : \mathcal{M} \to \mathbb{R}$ be a convex function on a convex domain $\mathcal{M}$ in Euclidean space. For two points $x, y \in \mathcal{M}$ and $\alpha \in \mathbb{R}$, a $\mathcal{D}_\Psi^{(\alpha)}$-divergence is a function of the form

$$\mathcal{D}_\Psi^{(\alpha)}(x, y) = \frac{4}{1-\alpha^2}\left[\frac{1-\alpha}{2}\Psi(x) + \frac{1+\alpha}{2}\Psi(y) - \Psi\left(\frac{1-\alpha}{2}x + \frac{1+\alpha}{2}y\right)\right].$$

For cost functions of this form, we use Ψ to construct a Hessian metric on $\mathcal{M}$ and consider X and Y as subsets of $\mathcal{M}$. Theorem 6 relates[18] the MTW tensor of a $\mathcal{D}_\Psi^{(\alpha)}$-divergence on $\mathcal{M}$ to the orthogonal antibisectional curvature of $T\mathcal{M}$ and Proposition 8.

There are several reasons to extend our results to the case of a $\mathcal{D}_\Psi^{(\alpha)}$-divergence. Firstly, they are a natural class of divergences which interpolate between dual Bregman divergences [1967] (as α approaches ± 1).

[18]There is a scaling factor of $\frac{1}{2}(1-\alpha^2)$ between the curvature and the MTW tensor for a $\mathcal{D}_\Psi^{(\alpha)}$-divergence.

Secondly, such divergences satisfy a natural “biduality”, which is important to the study of information geometry [Zhang 2004]. Thirdly, $\mathcal{D}_\Psi^{(\alpha)}$-divergences are often more natural than Ψ-costs for optimal transport on statistical manifolds for the following reason.

Both Ψ-costs and $\mathcal{D}_\Psi^{(\alpha)}$-divergences involve a convex function defined on an open domain of $\mathbb{R}^n$. When choosing to use one versus the other as a cost function, the primary difference is whether to assume that $X - Y \subset \mathcal{M}$ (as for the former), or that $X,\ Y,\ \frac{1}{2}(1-\alpha)X + \frac{1}{2}(1+\alpha)Y \subset \mathcal{M}$ (as for the latter).

For a $\mathcal{D}_\Psi^{(\alpha)}$-divergence induced by a log-partition function Ψ, the domain $\mathcal{M}$ is convex. As such, if we consider X and Y to be subsets of the natural parameters of an exponential family and let $\alpha \in (-1, 1)$, we are assured that $\frac{1}{2}(1-\alpha)X + \frac{1}{2}(1+\alpha)Y \subset \mathcal{M}$. Because of this, the $\mathcal{D}_\Psi^{(\alpha)}$-divergence is guaranteed to be well-defined. We will give an example of such a divergence function and prove the regularity for an associated optimal transport problem in Section 6.1.3.

More broadly, divergences are a generalization of distance functions, where the assumptions of symmetry and the triangle inequality are dropped. Such functions are widely used in statistics and information geometry because they can be seen as generalizations of the relative entropy. Using divergences as cost functions in order to connect information geometry with optimal transport is an active field of research [Wong and Yang 2019], and we expect that there are interesting connections yet to be found.

6. Applications

As the results in the previous section give a new interpretation for prior work, it is natural to ask for original results that can be found using this approach. In this section, we give several such applications. We will not provide the derivations of the identities in this section, as they are very involved but otherwise routine. In order to compute the associated curvature tensors, we have written a Mathematica notebook, which is available online [Khan 2018].

6.1. ***A complete, complex surface with (NOAB).*** Since the antibisectional curvature appears similar to the bisectional curvature, we can also ask how much control nonnegative antibisectional curvature or (NOAB) provides over the geometry of a Kähler manifold. Using the polarization formula [Hawley 1953], it can be shown that any metric of constant holomorphic sectional curvature has vanishing orthogonal antibisectional curvature, so any Hermitian symmetric space satisfies (NOAB). One can then ask whether there are other examples.

The following example gives a very interesting metric which satisfies (NOAB) and completeness but is neither Hermitian symmetric nor biholomorphic to $\mathbb{C}^n$.

Example 9 (a complete surface with (NOAB)). Consider the negative half-plane

$$\mathcal{M} = \mathbb{H} := \{(u^1, u^2) \mid u^2 < 0\}.$$

Prescribe a Hessian metric associated with the potential function $\Psi : \mathbb{H} \to \mathbb{R}$ given by

$$\Psi(u) = -\frac{(u^1)^2}{4u^2} - \frac{1}{2}\log(-2u^2).$$

For a vector $\xi = \partial_{u^1} + a\partial_{u^2}$ and a covector $\eta = a\,du^1 - du^2$, the associated orthogonal antibisectional curvature[19] on $T\mathbb{H}$ is given by

$$\mathfrak{O}\mathfrak{A}(\eta^\sharp, \xi) = \frac{6a^2(-a(u^1)^2 + u^2)^2}{(u^2)^2}.$$

As such, the metric has (NOAB). This metric is of independent interest, and for a more complete discussion, we refer the reader to [Molitor 2014]. We will note in passing a few of its curvature properties. For a vector $\xi = \partial_{u^1} + a\partial_{u^2}$ and a covector $\eta = du^1 + a\,du^2$, the antibisectional curvature is given by

$$\mathfrak{A}(\eta, \xi) = 2 - 12a^2 - 12\frac{au^1}{u^2} + 6\left(\frac{au^1}{u^2}\right)^2.$$

As such, the antibisectional curvature does not have a definite sign. It can similarly be shown that the orthogonal bisectional curvature also does not have a definite sign. However, the metric does have constant negative scalar curvature. This manifold is complete and Stein (it is biholomorphic to an open set in $\mathbb{C}^2$). However, it has the standard complex structure on a half-space in $\mathbb{R}^4$, so is *not* biholomorphic to $\mathbb{C}^2$.

6.1.1. *The Fisher metric of the normal distribution $\mathcal{N}(\mu, \sigma)$.* Although this example has interesting theoretical properties, it may appear to be a somewhat ad hoc construction without context. In fact, it is a natural example from information geometry. If we consider u^1 and u^2 as the natural parameters of the normal distribution with mean μ and standard deviation σ (i.e., $u^1 = \mu/\sigma^2$ and $u^2 = -1/(2\sigma^2)$), then the Riemannian metric $g_{ij} = \Psi_{ij}$ is the Fisher metric on the statistical manifold of univariate normal distributions (with unknown mean and variance). As a Riemannian manifold, $(\mathbb{H}, g)$ is a complete hyperbolic surface (which motivated our choice of notation). Note, however, that the (u^1, u^2)-coordinates do *not* induce the standard half-plane model of hyperbolic space.[20]

6.1.2. *A closely related example.* Using the normal statistical manifold, it is possible to construct another Kähler metric which satisfies (MTW). This space is actually Hermitian symmetric and was first constructed by Shima [2007, Example 6.7].

Consider the domain

$$\widetilde{\mathcal{M}} := \{(\theta^1, \theta^2) \mid \theta^2 - (\theta^1)^2 > 0\}$$

and prescribe it with a Hessian metric with potential $\Psi^*(\theta) = -\frac{1}{2} - \log(\theta^2 - (\theta^1)^2)$.

This potential arises from the parametrization of the univariate normal distribution in terms of its dual parameters $\theta^1 = \mu$ and $\theta^2 = \mu^2 + \sigma^2$ and the potential $\Psi^*(\theta)$ is the Legendre dual of the above potential in Example 9. Computing the antibisectional curvature for a vector ξ and covector η, we find that it satisfies

$$\mathfrak{A}(\xi, \eta^\sharp) = -\eta(\xi)^2.$$

As such, the orthogonal antibisectional curvature vanishes and the holomorphic sectional curvature is a negative constant. From this, we can see that the geometry of $T\widetilde{\mathcal{M}}$ is of independent interest, as it is a

[19]By computing antibisectional curvature solely on real vectors and covectors, we are slightly abusing notation. To formalize this, extend ξ and η to their real counterparts on $T\mathcal{M}$.

[20]In (μ, σ)-coordinates, the Fisher metric is $ds^2 = (1/\sigma^2)(d\mu^2 + 2d\sigma^2)$, which is much closer to the standard model.

complete Hermitian symmetric space with constant negative holomorphic sectional curvature. This is an example of a Siegel upper half-space.

We note that it is possible to construct other Kähler metrics with (NOAB) that are very similar to $T\widetilde{\mathcal{M}}$. Using a similar construction for round multivariate Gaussian distributions, it is possible to construct such a Hermitian symmetric space in arbitrary dimensions. For another example, we can consider the potential $\Psi(\theta^1, \theta^2) = -\frac{1}{2} - \log(\theta^2 - (\theta^1)^4)$, which also has (NOAB).

6.1.3. *Regularity for an associated cost function.* We can also use the potential

$$\Psi(u) = -\frac{(u^1)^2}{4u^2} - \frac{1}{2}\log(-2u^2)$$

to construct a cost function with a natural regularity theory. Due to the fact that the domain of Ψ is convex, it is more natural to consider a $\mathcal{D}_\Psi^{(\alpha)}$-divergence rather than a Ψ-cost. As such, we will consider the cost function

$$\begin{aligned} c(x, y) &= \mathcal{D}_\Psi^{(0)}(x, y) \\ &= 2\left(-\frac{(x^1)^2}{4x^2} - \frac{1}{2}\log(-2x^2)\right) + 2\left(\frac{(y^1)^2}{4y^2} + \frac{1}{2}\log(-2y^2)\right) + 4\left(\frac{(x^1+y^1)^2}{8(x^2+y^2)} + \frac{1}{2}\log(-x^2-y^2)\right). \end{aligned}$$

For this cost function, we can apply our previous calculations to obtain the following result.

Corollary. *Suppose μ and ν are probability measures supported on bounded subsets X and Y of the normal statistical manifold $\mathcal{M}$. Suppose further that the following regularity assumptions hold.*

(1) *μ and ν are absolutely continuous with respect to the Lebesgue measure. Furthermore, $d\mu$ and $d\nu$ are smooth and bounded away from zero and infinity on their respective supports.*

(2) *For all $x \in X$, $\frac{1}{2}(x + Y)$ is strictly convex with respect to the coordinates $\theta^1 = \mu$ and $\theta^2 = \mu^2 + \sigma^2$. Furthermore, the same property holds for $\frac{1}{2}(X + y)$ for all $y \in Y$.*

Let $c(x, y)$ be the cost function given by

$$c(x, y) = \mathcal{D}_\Psi^{(0)}(x, y) = 2\Psi(x) + 2\Psi(y) - 4\Psi\left(\frac{x+y}{2}\right),$$

where Ψ is the convex function given in Example 9. Then the c-optimal map $\mathbb{T}_{\mathrm{U}}$ taking μ to ν is smooth.

6.2. ***The regularity of pseudoarbitrages.*** Recently, a series of papers [Pal and Wong 2016; 2018a; 2018b; Wong 2018; 2019] has studied the problem of finding *pseudoarbitrages*, which are investment strategies which outperform the market portfolio under "mild and realistic assumptions". Their work combines information geometry with optimal transport and mathematical finance to reduce the problem to solving optimal transport problems where the cost function is given by a so-called log-divergence.

A central result in [Pal and Wong 2018a] shows that a portfolio map π outperforms the market portfolio almost surely in the long run if and only if it is a solution to the Monge problem for the cost function $c : \mathbb{R}^{n-1} \times \mathbb{R}^{n-1} \to \mathbb{R}$ given by

$$c(x, y) := \log\left(1 + \sum_{i=1}^{n-1} e^{x^i - y^i}\right) - \log(n) - \frac{1}{n}\sum_{i=1}^{n-1} x^i - y^i. \tag{18}$$

To give some context for this cost function, it is instructive to consider x and y as the natural parameters of the multinomial distribution. For natural parameters $\{x^i\}_{i=1}^{n-1}$, we can compute the probability p_i of the i-th event (in this context, the i-th market weight) using the formulas

$$p_i = \frac{e^{x^i}}{1+\sum_{j=1}^{n-1} e^{x^j}} \quad \text{for } 1 \le i < n, \tag{19}$$

$$p_n = \frac{1}{1+\sum_{j=1}^{n-1} e^{x^j}}. \tag{20}$$

To write this cost function in a more familiar form, we similarly find probabilities q_i associated to the y-parameters and fix $\boldsymbol{\pi} = (1/n, \dots, 1/n) \in \triangle^n$. Rewriting our cost in these terms,[21] we have

$$\hat{c}(p,q) := \log\left(\sum_{i=1}^{n} \boldsymbol{\pi}_i \frac{p_i}{q_i}\right) - \sum_{i=1}^{n} \boldsymbol{\pi}_i \log\left(\frac{p_i}{q_i}\right).$$

This quantity is known as the free energy in statistical physics [Pal and Wong 2018a] and by various different names in finance (such as the “diversification return”, the “excess growth rate”, the “rebalancing premium” and the “volatility return”). Since Pal and Wong refer to this as a *logarithmic divergence*, we refer to this cost as the logarithmic cost. This cost function is not symmetric, so is not induced by any distance function. However, Jensen’s inequality shows that it is a divergence.

The main focus of Pal and Wong’s work is to study the information-geometric properties of divergence functions induced by exponentially concave functions, of which $\hat{c}$ is only a single example. For any exponentially concave function, one can define a corresponding divergence which has a self-dual representation in terms of the logarithmic cost; see [Pal and Wong 2018a, Proposition 3.7]. In order to study optimal transport, we do not specify the exponentially concave function a priori. In fact, such a function induces the *solution* to an optimal transport problem.

For the logarithmic cost, only the first term affects optimal transport. As such, we instead consider the cost function

$$\tilde{c}(x,y) := \log\left(1 + \sum_{i=1}^{n-1} e^{x^i - y^i}\right).$$

This is now a Ψ-cost for the convex function

$$\Psi(u) = \log\left(1 + \sum_{i=1}^{n-1} e^{u^i}\right).$$

As such, we can apply Theorem 6 to compute the MTW tensor for the cost $\tilde{c}$. For a vector ξ and a covector η, the antibisectional curvature of $T\mathbb{R}^{n-1}$ (denoted by $\mathfrak{A}$) with Hessian metric induced by Ψ is

$$\mathfrak{A}(\xi, \eta^\sharp) = 2(g(\eta^\sharp, \xi))^2.$$

As such, the MTW tensor identically vanishes and the cost has nonnegative cross-curvature. A proof for this identity can be found in [Shima 2007, Proposition 3.9]. From the curvature formulas, we see that

[21] This is what Pal and Wong denote by $T(p \mid q)$.

this potential induces a Kähler metric on $\mathbb{C}^n$ with constant positive holomorphic sectional curvature. As such, it is a Hermitian symmetric space (although it is not complete).

In order to apply the result of [Trudinger and Wang 2009], we must also determine what relative c-convexity means in this context. To do so, we solve for the dual coordinates to the natural parameters u^i by calculating $\partial_{u^i}\Psi$ for $i = 1, \dots, n-1$. Doing so, we find that the dual coordinates are

$$\theta^i = \frac{e^{u^i}}{1+\sum_{j=1}^{n-1} e^{u^j}} \quad \text{for } 1 \le i < n,$$

which are exactly the formulas for the market weights p_i (i.e., $\theta^i = p_i$). This is initially surprising, but has a natural interpretation in terms of information geometry.

6.2.1. *The information geometry of the multinomial distribution.* It is worth discussing the geometry of this example in more detail. It turns out that if we consider the $\{x^i\}$ as natural parameters, the potential Ψ induces the Fisher metric of the multinomial distribution, which is an important exponential family in statistics. Geometrically, this is the round metric on the positive orthant of a sphere, which immediately shows that neither the underlying Hessian metric nor the Sasaki metric is complete. It is worth mentioning that this metric cannot be extended to a Kähler Sasaki metric on the tangent bundle of the entire sphere, due to the fact that the sphere is *not* an affine manifold.

For an exponential family of probability distributions, the dual coordinates are the expected values of the natural sufficient statistics. More specifically, for the multinomial distribution the dual coordinates are precisely the original market weights, which explains the relationship between the market weights and the partial derivatives of the potential function. As such, if we let $\mathcal{P}$ be the coordinate transformation from the natural parameters x to the market weights p (i.e., $\mathcal{P}(x)$ is as given in (19)), a subset $X \subset \mathbb{R}^{n-1}$ is relatively c-convex if and only if the set $\mathcal{P}(X)$ is convex as a subset of the probability simplex in the usual sense. Using this transformation, we say that a subset $\mathcal{P}(X)$ of the probability simplex has *uniform probability* if X is a precompact set. More concretely, a subset $\mathcal{P}(X)$ has uniform probability if and only if there exists $\delta > 0$ so that for all $p \in \mathcal{P}(X)$ and $1 \le i \le n$, $p_i > \delta$.

6.2.2. *Regularity of optimal transport.* From these observations and the previous identity for the MTW tensor (13), we can derive the following regularity result.

Corollary 10. *Suppose μ and ν are smooth probability measures supported respectively on subsets X and Y of the probability simplex $\triangle^n$. Suppose further that the following regularity assumptions hold*:

(1) *X and Y are smooth and strictly convex. Furthermore, both have uniform probability* (*as defined above*).

(2) *μ and ν are absolutely continuous with respect to the Lebesgue measure and $d\mu$ and $d\nu$ are bounded away from zero and infinity on their supports.*

Let $\hat{c}(p,q)$ be the cost function given by

$$\hat{c}(p,q) = \log\left(\frac{1}{n}\sum_{i=1}^{n}\frac{q_i}{p_i}\right) - \frac{1}{n}\sum_{i=1}^{n}\log\frac{q_i}{p_i}.$$

Then the $\hat{c}$-optimal map $\mathbb{T}_{\mathrm{U}}$ taking μ to ν is smooth.

Pal and Wong [2018b] study the cost function $\hat{c}$ and use it to define a displacement interpolation between two probability measures. In their paper, they inquire about the regularity problem for this interpolation. We can now answer this question using the previous result.

Corollary 11. *Suppose that μ and ν are smooth probability measures satisfying the assumptions of Corollary 10 and that $\mathbb{T}_{\mathrm{U}}$ is the $\hat{c}$-optimal map transporting μ to ν. Suppose further that $\mathbb{T}(t)\mu$ is the displacement interpolation from μ to ν induced by the one-parameter family of exponentially concave potentials $\varphi(t)$, with*

$$\varphi(t) = t\mathrm{U} + (1-t)\varphi_0, \quad \textit{where } \varphi_0 = \frac{1}{n}\sum_{i=1}^{n} \log p_i.$$

Then $\mathbb{T}_{\varphi(t)}$ is smooth, both as a map on the probability simplex for fixed t and also in terms of the t-parameter.

For $t = 1$, the solution to the interpolation problem is simply $\mathbb{T}_{\mathrm{U}}$ and so Corollary 10 shows that the potential U is smooth. Since the displacement interpolation linearly interpolates between smooth potential functions, the associated displacement interpolation is also smooth for $0 \leq t \leq 1$. The displacement induced by linearly interpolating the potential functions is exactly the transport considered in [Pal and Wong 2018b] (see Definition 6), and so this establishes regularity for this transport.

In closing, we note that the cost function considered here is very similar, but not identical, to the radial antennae cost, which was studied in [Wang 2004]. It is of interest to determine whether there is some deeper connection between these two costs which explains their apparent similarity.

6.3. ***Other examples in complex geometry and optimal transport.*** While writing this paper, we were able to find several more examples of Hessian manifolds whose tangent bundles have nonnegative bisectional curvature or (NOAB).

Relatively few examples of positively curved metrics are known (for some examples, see [Wu and Zheng 2011]), so this method may be helpful for finding new ones. One limitation of this approach is that many of the manifolds are not complete as metric spaces. It would be of interest to determine which convex functions induce complete Kähler metrics with nonnegative or positive orthogonal bisectional curvature, and we plan to study this problem in future work.

Each of these examples further induces a cost with nonnegative MTW tensor. Furthermore, since many of these examples are obtained from statistical manifolds, it may be possible to use them to induce meaningful statistical divergences.

(1) $\Psi(u) = -\log\bigl(1 - \sum_{i=1}^{n} e^{u^i}\bigr)$ defined on the set $\mathcal{M} = \bigl\{u \bigm| \sum_{i=1}^{n} e^{u^i} < 1\bigr\}$. This potential induces $T\mathcal{M}$ with a Sasaki metric of constant negative holomorphic sectional curvature:

$$\mathfrak{A}(\xi, \eta^\sharp) = -\eta(\xi)^2.$$

From an information-geometric point of view, this is the Fisher metric of the negative multinomial distribution. As a Hessian manifold, $\mathcal{M}$ is a noncompact metric of constant negative holomorphic sectional curvature, but is not complete.

(2) $\Psi(u) = (e^{u^1} + e^{u^2})^p$ for $0 < p < 1$. For a vector $\xi = \partial_{u^1} + a\partial_{u^2}$ and covector $\eta = a\,du^1 - du^2$, the associated orthogonal antibisectional curvature of the Sasaki metric is given by

$$\mathfrak{O}\mathfrak{A}(\xi, \eta^\sharp) = \frac{2(1/p - 1)(a-1)^2(e^{u^1} + ae^{u^2})^2}{(e^{u^1} + e^{u^2})^{2+p}}.$$

This is nonnegative, and so the Kähler metric has (NOAB). As Hessian manifolds, this family of metrics is neither compact nor complete.

(3) $\Psi(u) = \log(\cosh(u^1) + \cosh(u^2))$. This potential induces a Sasaki metric on $T\mathbb{R}^2$ whose bisectional curvature is nonnegative. For a vector $\xi = \xi_1 \partial_{u^1} + \xi_2 \partial_{u^2}$ and covector $\eta = \eta_1\, du^1 + \eta_2\, du^2$, the associated bisectional curvature is given by

$$\mathfrak{B}(\xi, \eta^\sharp) = |\xi|^2 |\eta|^2 + 4\xi_1\xi_2\eta_1\eta_2,$$

where $|\xi|^2 = \xi_1^2 + \xi_2^2$ and $|\eta|^2 = \eta_1^2 + \eta_2^2$.

Furthermore, the antibisectional curvature also satisfies the same formula:

$$\mathfrak{A}(\xi, \eta^\sharp) = |\xi|^2 |\eta|^2 + 4\xi_1\xi_2\eta_1\eta_2.$$

As such, this metric has (NOAB) and nonnegative bisectional curvature. As a Hessian manifold, this metric is bounded, and so is not complete. Note that the curvature of this metric is in fact parallel with respect to D, which makes it an interesting example. We will explore this metric further in future work.

7. Open questions

7.1. ***The complex geometry of optimal maps.*** It is of interest to understand the geometry of optimal maps from the perspective of the complex geometry. Although we were able to give a complex-/information-geometric interpretation of the Ma–Trudinger–Wang conditions, we do not have a complex-geometric interpretation for Theorem 1.

It is worth comparing the situation to the pseudo-Riemannian theory of optimal transport of [Kim and McCann 2010]. One striking feature in this theory is the natural geometric interpretation for optimal maps. More precisely, if one deforms the pseudometric by a particular conformal factor (which is determined by the respective densities), the optimal map is induced by a maximal codimension-n surface with respect to the conformal pseudometric [Kim et al. 2010].

For a Ψ-cost, we hope that [Gangbo and McCann 1995] will allow us to encode optimal transport problems within $T\mathcal{M}$ so that the solution corresponds to some submanifold (or a less regular subset when the transport is discontinuous). Intuitively, this should indicate the "direction" in which the mass is transported from (X, μ) to (Y, ν). However, at present we cannot make this intuition rigorous, so we leave it for future work.

If we are able to complete the previous step, a natural follow-up question would be to try to establish the regularity theory for optimal transport with Ψ-costs in terms of complex Monge–Ampère equations. For an overview on complex Monge–Ampère equations, we refer the reader to the paper by Phong, Song and Sturm [Phong et al. 2012].

7.2. ***A potential non-Kähler generalization.*** While Ψ-costs yield many interesting examples, there are many relevant cost functions which are not of this form. As such, one natural generalization of the construction considered here is to instead consider a Lie group $\mathcal{G}$ and cost functions of the form $\Psi(x \cdot y^{-1})$ for $x, y \in \mathcal{G}$. Our work thus far can be interpreted as doing this calculation in the special case where $\mathcal{G}$ is Abelian. For non-Abelian groups, we hope it is possible to recover the MTW tensor as a curvature tensor in almost-complex geometry. In this case, there would be correction terms due to the non-Abelian nature of the group. Furthermore, the natural connections on a non-Abelian Lie group are not flat, so the associated almost-complex structure on the tangent bundle $T\mathcal{G}$ would fail to be integrable.[22]

There are several key difficulties in making this program rigorous. Firstly the curvature of an almost-Hermitian manifold is much more complicated than that of a Kähler manifold, in that it does not satisfy (5). Furthermore, there is not one, but several canonical connections to choose from, and it's not immediately clear which is the right one to use. Finally, the cut locus of a Lie group can be nontrivial, which plays an important role in the regularity theory of optimal transport on manifolds. It seems that before any of these issues can be addressed, it will be necessary to understand the optimal map in terms of complex geometry. Hopefully this will provide insight into the correct generalization in the almost-complex setting.

However, there is reason to be hopeful about this approach, as there are several examples of MTW(κ) costs induced from this construction. Most strikingly, it is known that the squared distance cost on $\mathbb{RP}^3$ with its round metric satisfies a stronger version of the MTW condition which implies regularity [Loeper and Villani 2010]. However, $\mathbb{RP}^3$ is diffeomorphic to SO(3) and the round metric is one example of a left-invariant Berger metric (see [Brown et al. 2007] for a more complete discussion of left-invariant metrics). As such, we hope that this approach can be used to find other examples of cost functions satisfying MTW(κ), either by considering other Lie groups or by considering other left-invariant metrics on SO(3).

7.3. ***Implications for optimal transport.*** The primary focus of this work is to use optimal transport theory to study complex geometry and information geometry. However, it remains an open question what can be proven about optimal transport using this approach. For instance, it is hard to find cost functions which satisfy MTW(0). Similarly, relatively few Kähler metrics of positive curvature are known and there are various stability and gap theorems about them (see, e.g., [Liu 2019; Ni and Niu 2019]). In optimal transport, we expect that similar results can be proven using complex geometry. We plan to explore this topic in future work.

7.4. ***The complex geometry of the antibisectional curvature.*** The orthogonal antibisectional curvature is a very subtle invariant, and its geometry is quite mysterious. It does not determine the full curvature tensor of the manifold, since all spaces with constant holomorphic sectional curvature have vanishing orthogonal antibisectional curvature. However, it is nonnegative for some important metrics which do not have any other obvious nonnegativity properties. In future work, we hope to understand this curvature more fully and to understand what sort of control it exerts over the geometry of a complex manifold.

[22] If we restrict our attention to torsion-free connections, we may be able to recover an almost-Kähler theory (see [Satoh 2007, Theorem 1.1]), but it's not clear that this is the best option.

Acknowledgments

Khan would like to thank Fangyang Zheng and Bo Guan for their helpful discussions with this project. He would also like to thank the referee for helpful suggestions for the paper. The project is supported by DARPA/ARO Grant W911NF-16-1-0383 ("Information Geometry: Geometrization of Science of Information", PI: Zhang). A preliminary announcement of these results appeared [Khan and Zhang 2019].

References

[Amari 2016] S. Amari, *Information geometry and its applications*, Applied Mathematical Sciences **194**, Springer, 2016. MR Zbl

[Amari and Armstrong 2014] S. Amari and J. Armstrong, "Curvature of Hessian manifolds", *Differential Geom. Appl.* **33** (2014), 1–12. MR Zbl

[Ay et al. 2015] N. Ay, J. Jost, H. Vân Lê, and L. Schwachhöfer, "Information geometry and sufficient statistics", *Probab. Theory Related Fields* **162**:1-2 (2015), 327–364. MR Zbl

[Ballmann 2006] W. Ballmann, *Lectures on Kähler manifolds*, European Mathematical Society (EMS), Zürich, 2006. MR Zbl

[Bregman 1967] L. M. Bregman, "A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming", *Ž. Vyčisl. Mat i Mat. Fiz.* **7** (1967), 620–631. In Russian; translated in *USSR Comput. Math. and Math. Phys.* **7**:3 (1967), 200–217. MR Zbl

[Brenier 1987] Y. Brenier, "Décomposition polaire et réarrangement monotone des champs de vecteurs", *C. R. Acad. Sci. Paris Sér. I Math.* **305**:19 (1987), 805–808. MR Zbl

[Brown et al. 2007] N. Brown, R. Finck, M. Spencer, K. Tapp, and Z. Wu, "Invariant metrics with nonnegative curvature on compact Lie groups", *Canad. Math. Bull.* **50**:1 (2007), 24–34. MR Zbl

[Caffarelli 1992] L. A. Caffarelli, "The regularity of mappings with a convex potential", *J. Amer. Math. Soc.* **5**:1 (1992), 99–104. MR Zbl

[Chen 2007] X. X. Chen, "On Kähler manifolds with positive orthogonal bisectional curvature", *Adv. Math.* **215**:2 (2007), 427–445. MR Zbl

[De Philippis and Figalli 2014] G. De Philippis and A. Figalli, "The Monge–Ampère equation and its link to optimal transportation", *Bull. Amer. Math. Soc.* (*N.S.*) **51**:4 (2014), 527–580. MR Zbl

[Dombrowski 1962] P. Dombrowski, "On the geometry of the tangent bundle", *J. Reine Angew. Math.* **210** (1962), 73–88. MR Zbl

[Feng et al. 2017] H. Feng, K. Liu, and X. Wan, "Compact Kähler manifolds with positive orthogonal bisectional curvature", *Math. Res. Lett.* **24**:3 (2017), 767–780. MR Zbl

[Figalli et al. 2011] A. Figalli, Y.-H. Kim, and R. J. McCann, "When is multidimensional screening a convex program?", *J. Econom. Theory* **146**:2 (2011), 454–478. MR Zbl

[Figalli et al. 2013] A. Figalli, Y.-H. Kim, and R. J. McCann, "Hölder continuity and injectivity of optimal maps", *Arch. Ration. Mech. Anal.* **209**:3 (2013), 747–795. MR Zbl

[Gangbo and McCann 1995] W. Gangbo and R. J. McCann, "Optimal maps in Monge's mass transport problem", *C. R. Acad. Sci. Paris Sér. I Math.* **321**:12 (1995), 1653–1658. MR Zbl

[Gangbo and McCann 1996] W. Gangbo and R. J. McCann, "The geometry of optimal transportation", *Acta Math.* **177**:2 (1996), 113–161. MR Zbl

[Goldberg and Kobayashi 1967] S. I. Goldberg and S. Kobayashi, "Holomorphic bisectional curvature", *J. Differential Geometry* **1** (1967), 225–233. MR Zbl

[Gu and Zhang 2010] H. Gu and Z. Zhang, "An extension of Mok's theorem on the generalized Frankel conjecture", *Sci. China Math.* **53**:5 (2010), 1253–1264. MR Zbl

[Hawley 1953] N. S. Hawley, "Constant holomorphic curvature", *Canad. J. Math.* **5** (1953), 53–56. MR Zbl

[Hörmander 1973] L. Hörmander, *An introduction to complex analysis in several variables*, 2nd ed., North-Holland Mathematical Library **7**, North-Holland, Amsterdam, 1973. MR Zbl

[Kantorovitch 1958] L. Kantorovitch, "On the translocation of masses", *Management Sci.* **5** (1958), 1–4. MR Zbl

[Khan 2018] G. Khan, "MTW notebook", Mathematica notebook, 2018, available at https://sites.google.com/a/umich.edu/gabekhan/code/optimaltransport.

[Khan and Zhang 2019] G. Khan and J. Zhang, "Hessian curvature and optimal transport", pp. 423–430 in *Geometric science of information*, edited by F. Nielsen and F. Barbaresco, Lecture Notes in Comput. Sci. **11712**, Springer, 2019. MR

[Kim and McCann 2010] Y.-H. Kim and R. J. McCann, "Continuity, curvature, and the general covariance of optimal transportation", *J. Eur. Math. Soc.* (*JEMS*) **12**:4 (2010), 1009–1040. MR Zbl

[Kim et al. 2010] Y.-H. Kim, R. J. McCann, and M. Warren, "Pseudo-Riemannian geometry calibrates optimal transportation", *Math. Res. Lett.* **17**:6 (2010), 1183–1197. MR Zbl

[Li and Ni 2019] X. Li and L. Ni, "Kähler–Ricci shrinkers and ancient solutions with nonnegative orthogonal bisectional curvature", preprint, 2019. arXiv

[Liu 2019] G. Liu, "On Yau's uniformization conjecture", *Camb. J. Math.* **7**:1-2 (2019), 33–70. MR Zbl

[Loeper 2009] G. Loeper, "On the regularity of solutions of optimal transportation problems", *Acta Math.* **202**:2 (2009), 241–283. MR Zbl

[Loeper and Villani 2010] G. Loeper and C. Villani, "Regularity of optimal transport in curved geometry: the nonfocal case", *Duke Math. J.* **151**:3 (2010), 431–485. MR Zbl

[Ma et al. 2005] X.-N. Ma, N. S. Trudinger, and X.-J. Wang, "Regularity of potential functions of the optimal transportation problem", *Arch. Ration. Mech. Anal.* **177**:2 (2005), 151–183. MR Zbl

[Molitor 2014] M. Molitor, "Gaussian distributions, Jacobi group, and Siegel–Jacobi space", *J. Math. Phys.* **55**:12 (2014), art. id. 122102. MR Zbl

[Monge 1781] G. Monge, "Mémoire sur la théorie des déblais et des remblais", pp. 666–704 in *Histoire de l'Académie Royale des Sciences de Paris*, De l'Imprimerie Royale, Paris, 1781.

[Mori 1979] S. Mori, "Projective manifolds with ample tangent bundles", *Ann. of Math.* (2) **110**:3 (1979), 593–606. MR Zbl

[Newlander and Nirenberg 1957] A. Newlander and L. Nirenberg, "Complex analytic coordinates in almost complex manifolds", *Ann. of Math.* (2) **65** (1957), 391–404. MR Zbl

[Ni and Niu 2019] L. Ni and Y. Niu, "Gap theorem on Kähler manifolds with nonnegative orthogonal bisectional curvature", *J. Reine Angew. Math.* (online publication April 2019).

[Ni and Zheng 2018] L. Ni and F. Zheng, "Comparison and vanishing theorems for Kähler manifolds", *Calc. Var. Partial Differential Equations* **57**:6 (2018), art. id. 151. MR Zbl

[Pal 2017] S. Pal, "Embedding optimal transports in statistical manifolds", *Indian J. Pure Appl. Math.* **48**:4 (2017), 541–550. MR Zbl

[Pal and Wong 2016] S. Pal and T.-K. L. Wong, "The geometry of relative arbitrage", *Math. Financ. Econ.* **10**:3 (2016), 263–293. MR Zbl

[Pal and Wong 2018a] S. Pal and T.-K. L. Wong, "Exponentially concave functions and a new information geometry", *Ann. Probab.* **46**:2 (2018), 1070–1113. MR Zbl

[Pal and Wong 2018b] S. Pal and T.-K. L. Wong, "Multiplicative Schrödinger problem and the Dirichlet transport", preprint, 2018. arXiv

[Peyré and Cuturi 2019] G. Peyré and M. Cuturi, "Computational optimal transport", *Found. Trends Mach. Learn.* **11**:5–6 (2019), 355–607.

[Phong et al. 2012] D. H. Phong, J. Song, and J. Sturm, "Complex Monge–Ampère equations", pp. 327–410 in *Algebra and geometry: in memory of C. C. Hsiung*, edited by H.-D. Cao and S.-T. Yau, Surv. Differ. Geom. **17**, Int. Press, Boston, 2012. MR Zbl

[Satoh 2007] H. Satoh, "Almost Hermitian structures on tangent bundles", pp. 105–118 in *Proceedings of the Eleventh International Workshop on Differential Geometry*, edited by Y. J. Suh and B. H. Kim, Kyungpook Nat. Univ., Taegu, 2007. MR Zbl

[Sei 2013] T. Sei, "A Jacobian inequality for gradient maps on the sphere and its application to directional statistics", *Comm. Statist. Theory Methods* **42**:14 (2013), 2525–2542. MR Zbl

[Seshadri 2009] H. Seshadri, "Manifolds with nonnegative isotropic curvature", *Comm. Anal. Geom.* **17**:4 (2009), 621–635. MR Zbl

[Shima 2007] H. Shima, *The geometry of Hessian structures*, World Scientific, Hackensack, NJ, 2007. MR Zbl

[Siu and Yau 1980] Y. T. Siu and S. T. Yau, "Compact Kähler manifolds of positive bisectional curvature", *Invent. Math.* **59**:2 (1980), 189–204. MR Zbl

[Trudinger and Wang 2009] N. S. Trudinger and X.-J. Wang, "On the second boundary value problem for Monge–Ampère type equations and optimal transportation", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **8**:1 (2009), 143–174. MR Zbl

[Villani 2009] C. Villani, *Optimal transport: old and new*, Grundlehren der Mathematischen Wissenschaften **338**, Springer, 2009. MR Zbl

[Wang 2004] X.-J. Wang, "On the design of a reflector antenna, II", *Calc. Var. Partial Differential Equations* **20**:3 (2004), 329–341. MR Zbl

[Wong 2018] T.-K. L. Wong, "Logarithmic divergences from optimal transport and Rényi geometry", *Inf. Geom.* **1**:1 (2018), 39–78. MR Zbl

[Wong 2019] T.-K. L. Wong, "Information geometry in portfolio theory", pp. 105–136 in *Geometric structures of information*, edited by F. Nielsen, Springer, 2019. MR Zbl

[Wong and Yang 2019] T.-K. L. Wong and J. Yang, "Optimal transport and information geometry", preprint, 2019. arXiv

[Wu and Zheng 2011] H.-H. Wu and F. Zheng, "Examples of positively curved complete Kähler manifolds", pp. 517–542 in *Geometry and analysis, Vol. 1*, edited by L. Ji, Adv. Lect. Math. (ALM) **17**, International Press, Somerville, MA, 2011. MR Zbl

[Yang 1982] P. C. Yang, "Automorphism of tube domains", *Amer. J. Math.* **104**:5 (1982), 1005–1024. MR Zbl

[Yau 1994] S.-T. Yau, "Open problems in geometry", pp. v+235 in *Lectures on differential geometry*, edited by R. Schoen and S.-T. Yau, Conference Proceedings and Lecture Notes in Geometry and Topology **1**, International Press, Cambridge, MA, 1994.

[Zhang 2004] J. Zhang, "Divergence function, duality, and convex analysis", *Neural Comput.* **16**:1 (2004), 159–195.

[Zheng 2000] F. Zheng, *Complex differential geometry*, AMS/IP Studies in Advanced Mathematics **18**, American Mathematical Society, Providence, RI, 2000. MR Zbl

Received 3 Jun 2019. Revised 17 Oct 2019. Accepted 19 Nov 2019.

GABRIEL KHAN: gabekhan@umich.edu
University of Michigan, Ann Arbor, MI, United States

JUN ZHANG: junz@umich.edu
University of Michigan, Ann Arbor, MI, United States

PURE and APPLIED
ANALYSIS
Vol. 2, No. 2, 2020
https://doi.org/10.2140/paa.2020.2.427

SEMICLASSICAL ASYMPTOTICS FOR NONSELFADJOINT HARMONIC OSCILLATORS

VÍCTOR ARNAIZ AND GABRIEL RIVIÈRE

We consider nonselfadjoint perturbations of semiclassical harmonic oscillators. Under appropriate dynamical assumptions, we establish some spectral estimates such as upper bounds on the resolvent near the real axis when no geometric control condition is satisfied.

1. Introduction

Motivated by earlier work of Lebeau [1996] on the asymptotic properties of the damped wave equation, Sjöstrand [2000] initiated the spectral study of this partial differential equation on compact Riemannian manifolds. He proved that eigenfrequencies satisfy Weyl asymptotics in the high-frequency limit [Sjöstrand 2000, Theorem 0.1] — see also [Markus 1988; Markus and Macaev 1979]. Moreover, he showed that eigenfrequencies lie in a strip of the complex plane which can be completely determined in terms of the average of the damping function along the geodesic flow [Sjöstrand 2000, Theorems 0.0 and 0.2] — see also [Lebeau 1996; Rauch and Taylor 1975]. Following [Sjöstrand 2000], showing these results turns out to be the particular case of a more systematic study of a nonselfadjoint semiclassical problem which has since then been the object of several works. More precisely, it was investigated how these generalized eigenvalues are asymptotically distributed inside the strip determined by Sjöstrand and how the dynamics of the underlying classical Hamiltonian influences this asymptotic distribution. Mostly two questions have been considered in the literature. First, one can ask about the precise distribution of eigenvalues inside the strip and this question was addressed both in the completely integrable framework [Hitrik 2002; Hitrik and Sjöstrand 2004; 2005; 2008a; 2008b; 2012; 2018; Hitrik et al. 2007] and in the chaotic one [Anantharaman 2010]. Second, it is natural to focus on how eigenfrequencies can accumulate at the boundary of the strip and also to get resolvent estimates near the boundary of the strip. Again, this question has been explored both in the integrable case [Asch and Lebeau 2003; Hitrik and Sjöstrand 2004; Burq and Hitrik 2007; Anantharaman and Léautaud 2014; Burq and Gérard 2018] and in the chaotic one [Christianson 2007; Schenck 2010; Nonnenmacher 2011; Christianson et al. 2014; Rivière 2014; Jin 2017].

The purpose of this work is to consider the second question for simple models of completely integrable systems. Via these models, we aim at illustrating the influence of the subprincipal symbol of the selfadjoint part of our semiclassical operators on the asymptotic distribution of eigenvalues but also on resolvent estimates near the real axis. As briefly reminded below, this is related to the decay of the corresponding semigroup [Lebeau 1996]. Among other things, our study is motivated by [Asch and Lebeau 2003,

MSC2010: 35P20.
Keywords: semiclassical analysis, nonselfadjoint operators, resolvent estimates, integrable systems, averaging method.

Theorem 2.3]. In that reference, they indeed showed how a selfadjoint perturbation of the principal symbol of the damped wave operator on the 2-sphere can create a spectral gap inside the spectrum in the high-frequency limit. Theorem 7 below shows how this result can be extended to our context.[1] A major ingredient in the proof of [Asch and Lebeau 2003] but also in [Hitrik and Sjöstrand 2004; 2005; 2008a; 2008b; 2012; 2018; Hitrik et al. 2007] is the *analyticity* of the involved operators. One of the novelties of the present article compared with these references is Theorem 3, where we only suppose that the operators are *smooth*, i.e., quantizing $\mathcal{C}^\infty$ symbols. This theorem shows what can be said under these lower regularity assumptions and how this is influenced by the subprincipal symbols of the selfadjoint part, as was the case in [Asch and Lebeau 2003]. This will be achieved by building on the dynamical construction used by the first author and Macià [Arnaiz and Macià 2018] for studying Wigner measures of semiclassical harmonic oscillators — see also [Macià and Rivière 2016; 2019] in the case of Zoll manifolds. As in [Arnaiz and Macià 2018], we restrict ourselves to the case of nonselfadjoint perturbations of semiclassical harmonic oscillators on $\mathbb{R}^d$. Yet it is most likely that the methods presented here can be adapted to deal with semiclassical operators associated with more general completely integrable systems, including damped wave equations on Zoll manifolds.

1.1. ***Nonselfadjoint harmonic oscillators.*** Let us now describe the spectral framework in which we are interested. We fix $\omega = (\omega_1, \dots, \omega_d)$ to be an element of $(\mathbb{R}_+^*)^d$ and we set $\widehat{H}_\hbar$ to be the semiclassical harmonic oscillator given by

$$\widehat{H}_\hbar := \frac{1}{2}\sum_{j=1}^{d} \omega_j(-\hbar^2\partial_{x_j}^2 + x_j^2). \tag{1}$$

We want to understand the spectral properties of nonselfadjoint perturbations of $\widehat{H}_\hbar$. Before being more precise on that issue, let us recall that the symbol H of $\widehat{H}_\hbar$ is given by the classical harmonic oscillator

$$H(x,\xi) = \frac{1}{2}\sum_{j=1}^{d} \omega_j(\xi_j^2 + x_j^2), \qquad (x,\xi) \in \mathbb{R}^{2d}, \tag{2}$$

whose induced Hamiltonian flow will be denoted by ϕ_t^H. A brief account on the dynamical properties of this flow is given in Section 2. For any smooth function $a \in \mathcal{C}^\infty(\mathbb{R}^{2d})$, we define its average $\langle a\rangle$ by the Hamiltonian flow ϕ_t^H as

$$\langle a\rangle(x,\xi) := \lim_{T\to\infty} \frac{1}{T}\int_0^T a\circ\phi_t^H(x,\xi)\,dt \in \mathcal{C}^\infty(\mathbb{R}^{2d}), \tag{3}$$

whose properties are related to the Diophantine properties of ω — see Section 2 for details.

Fix now two smooth functions A and V in $\mathcal{C}^\infty(\mathbb{R}^{2d}, \mathbb{R})$ all of whose derivatives (at any order) are bounded. Following [Zworski 2012, Chapter 4], one can define the Weyl quantization of these smooth symbols:

$$\widehat{A}_\hbar := \mathrm{Op}_\hbar^w(A) \quad\text{and}\quad \widehat{V}_\hbar := \mathrm{Op}_\hbar^w(V).$$

[1] Observe that, compared with [Asch and Lebeau 2003], our operators are not necessarily associated with a periodic flow.

These are selfadjoint operators which are bounded on $L^2(\mathbb{R}^d)$ thanks to the Calderón–Vaillancourt theorem. We aim at describing the asymptotic properties of the following nonselfadjoint operators in the semiclassical limit $\hbar \to 0^+$:

$$\widehat{P}_\hbar := \widehat{H}_\hbar + \delta_\hbar \widehat{V}_\hbar + i\hbar \widehat{A}_\hbar,$$

where $\delta_\hbar \to 0$ as $\hbar \to 0^+$. More precisely, we focus on sequences of (pseudo-)eigenvalues $\lambda_\hbar = \alpha_\hbar + i\hbar\beta_\hbar$ such that there exist $\beta \in \mathbb{R}$ and $(v_\hbar)_{\hbar\to 0^+}$ in $L^2(\mathbb{R}^d)$ for which

$$(\alpha_\hbar, \beta_\hbar) \to (1, \beta) \quad \text{as } \hbar \to 0^+ \qquad \text{and} \qquad \widehat{P}_\hbar v_\hbar = \lambda_\hbar v_\hbar + r_\hbar, \quad \|v_\hbar\|_{L^2} = 1. \tag{4}$$

Here $r_\hbar$ should be understood as a small remainder term which will be typically of order $o(\hbar)$. This remainder term allows us to encompass the case of quasimodes, which is important to get resolvent estimates.

Remark 1. Throughout this work, we shall consider subsequences $\hbar_n \to 0^+$ such that the above convergence property holds. In order to alleviate notation, we will omit the index n and just write $\hbar \to 0^+$, $\lambda_\hbar = \lambda_{\hbar_n}$, $v_\hbar = v_{\hbar_n}$, etc. For a similar reason, we do not relabel subsequences. This kind of convention is standard when working with semiclassical parameters.

Recall from [Markus 1988; Markus and Macaev 1979; Sjöstrand 2000, Theorem 5.2] that true eigenvalues exist and that, counted with their algebraic multiplicity, they satisfy Weyl asymptotics as $\hbar \to 0^+$. It also follows from [Rauch and Taylor 1975; Lebeau 1996; Sjöstrand 2000, Lemma 2.1] that:

Proposition 2. *Let $(\lambda_\hbar = \alpha_\hbar + i\hbar\beta_\hbar)_{\hbar\to 0^+}$ be a sequence satisfying (4) with $\beta_\hbar \to \beta$ and $r_\hbar = o(\hbar)$. Then, one has*

$$\beta \in \Big[\min_{z\in H^{-1}(1)} \langle A\rangle(z),\ \max_{z\in H^{-1}(1)} \langle A\rangle(z)\Big]. \tag{5}$$

Note that one always has

$$\min_{z\in H^{-1}(1)} A(z) \le A_- := \min_{z\in H^{-1}(1)} \langle A\rangle(z) \le A_+ := \max_{z\in H^{-1}(1)} \langle A\rangle(z) \le \max_{z\in H^{-1}(1)} A(z),$$

where the inequalities may be strict. For the sake of completeness and as it will be instructive for our proof, we briefly recall the proof of this proposition[2] in Section 3.1. One can verify that the quantum propagator $(e^{it\widehat{P}_\hbar/\hbar})_{t\ge 0}$ defines a bounded operator on $L^2(\mathbb{R}^d)$ whose norm is bounded by $e^{|t|\|\operatorname{Op}_\hbar(A)\|_{\mathcal{L}(L^2)}}$. Moreover, if we suppose in addition that $\langle A\rangle \ge a_0 > 0$ on $\mathbb{R}^{2d}$, we say that the damping term is geometrically controlled and one gets exponential decay of the quantum propagator in time [Lebeau 1996; Helffer and Sjoestrand 2010]. More generally, controlling the way pseudoeigenvalues accumulate on the real axis provides information on the decay rate of the quantum propagator [Lebeau 1996; Helffer and Sjoestrand 2010], and this is precisely the question we are aiming at when $\langle A\rangle$ may vanish.

[2]In the case where the nonselfadjoint perturbation is $\gg \hbar$ and where the symbols enjoy some extra analytical properties, this proposition remains true (after a proper renormalization) when $r_\hbar = 0$ and when ω satisfies appropriate Diophantine properties, such as (9) below.

1.2. ***The smooth case.*** Let us now explain our main results, which show how the selfadjoint term $\widehat{V}_\hbar$ influences the way that the eigenvalues may accumulate on the boundary of the interval given by Proposition 2. In the smooth case, our main result reads as follows:

Theorem 3. *Suppose that* $A \geq 0$ *and that, for every* $(x, \xi) \in H^{-1}(1) \cap \langle A \rangle^{-1}(0)$, *there exists* $T > 0$ *such that*

$$\langle A \rangle \circ \phi_T^{\langle V \rangle}(x, \xi) > 0, \tag{6}$$

where $\phi_t^{\langle V \rangle}$ *is the Hamiltonian flow generated by* $\langle V \rangle$. *For every* $R > 0$, *there exists*[3] $\varepsilon_R > 0$ *such that, for*

$$\delta_\hbar \geq \varepsilon_R^{-1} \hbar^2,$$

and, for every sequence $(\lambda_\hbar = \alpha_\hbar + i\hbar\beta_\hbar)_{\hbar \to 0^+}$ *satisfying* (4) *with* $\|r_\hbar\| \leq \varepsilon_R \hbar \delta_\hbar$, *we have*

$$\liminf_{\hbar \to 0^+} \frac{\beta_\hbar}{\delta_\hbar} > R.$$

Remark 4. If $\delta_\hbar \gg \hbar^2$ and $\|r_\hbar\| \ll \hbar\delta_\hbar$, then this theorem shows that

$$\lim_{\hbar \to 0^+} \frac{\beta_\hbar}{\delta_\hbar} = +\infty.$$

In other words, under the geometric control condition (6), eigenvalues cannot accumulate too fast on the real axis as $\hbar \to 0^+$. We emphasize that, compared with the analytic case treated in [Asch and Lebeau 2003], our result applies a priori to quasimodes. Hence, it also yields the following resolvent estimate in the smooth case. For every $R > 0$, there exists some constant $\varepsilon_R > 0$ such that, for $\hbar > 0$ small enough and for $\delta_\hbar \geq \varepsilon_R^{-1} \hbar^2$,

$$\frac{\operatorname{Im} \lambda}{\hbar} \leq R\delta_\hbar \quad \Longrightarrow \quad \|(\widehat{P}_\hbar - \lambda)^{-1}\|_{L^2 \to L^2} \leq \frac{1}{\varepsilon_R \hbar \delta_\hbar}, \tag{7}$$

which is useful regarding energy decay estimates and asymptotic expansion of the corresponding semigroup — see, e.g., [Helffer and Sjoestrand 2010].

Note that the assumption that $A \geq 0$ makes the proof a little bit simpler but we could deal with more general functions by using the (nonselfadjoint) averaging method from [Sjöstrand 2000] and by making some appropriate Diophantine assumptions — see, e.g., Section 4. Our proof will crucially use the Fefferman–Phong inequality (hence the Weyl quantization) and this allows us to reach perturbations of size $\delta_\hbar \gtrsim \hbar^2$. If we had used another choice (say for instance the standard one), we would have only been able to use the Gårding inequality and it would have led us to the stronger restriction $\delta_\hbar \gtrsim \hbar$.

In the case where $V = 0$ and under some analyticity assumptions in dimension 2, it was shown in [Hitrik and Sjöstrand 2004, Theorem 6.7] that one can find some eigenvalues such that $\beta_\hbar$ is exactly of order $\hbar$ provided that ϕ_t^H is periodic and that $\langle A \rangle$ vanishes on finitely many closed orbits. Hence, our hypothesis (6) on the subprincipal V is crucial here. Note that this geometric condition is similar to the one appearing in [Arnaiz and Macià 2018] for the study of semiclassical measures of the Schrödinger equation — see also [Macià and Rivière 2016; 2019] in the case of Zoll manifolds. As we shall see,

[3] The (more or less explicit) constant ε_R coming out from our proof satisfies $\lim_{R \to +\infty} \varepsilon_R = 0$.

ensuring this dynamical property depends on the Diophantine properties of ω. Recall that, to each ω, one can associate the submodule

$$\Lambda_\omega := \{k \in \mathbb{Z}^d : \omega \cdot k = 0\}. \tag{8}$$

When the resonance module satisfies $\Lambda_\omega = \{0\}$, we will see in Section 2 that our geometric control condition (6) can only be satisfied if $\langle A\rangle > 0$. A typical case in which our dynamical condition holds is when $H^{-1}(1) \cap \langle A\rangle^{-1}(0)$ consists of a disjoint union of a finite number of minimal ϕ_t^H-invariant tori $(\mathcal{T}_k)_{k=1,\dots,N}$. In this case, our dynamical condition is equivalent to saying that the Hamiltonian vector field $X_{\langle V\rangle}$ satisfies

$$\text{for all } 1 \le k \le N,\ \text{for all } z \in \mathcal{T}_k, \quad X_{\langle V\rangle}(z) = \frac{d}{dt}(\phi_t^{\langle V\rangle}(z))|_{t=0} \notin T_z\mathcal{T}_k.$$

1.3. *The analytic case.* We now discuss the case where the functions A and V enjoy some analyticity properties. To that aim, we follow a method introduced in [Asch and Lebeau 2003] in the case of the damped wave equation on the 2-sphere. We will explain how to adapt this strategy in the framework of harmonic oscillators which are not necessarily periodic. The upcoming results should be viewed as an extension of Asch and Lebeau's construction to semiclassical harmonic oscillators and as an illustration on what can be gained via analyticity compared with the purely dynamical approach used to prove Theorem 3. We emphasize that the argument presented here only holds for true eigenmodes, i.e., $r_\hbar = 0$ in (4). In particular, it does not seem to yield any resolvent estimate like (7), which is crucial to deducing some results on the semigroup generated by $\widehat{P}_\hbar$.

We now assume some extra conditions on the symbols H, V and A. First, given the vector of frequencies $\omega := (\omega_1, \dots, \omega_d)$ of the harmonic oscillator H, we shall say that $\omega \in \mathbb{R}^d$ is *partially Diophantine* [de la Llave 2001, equation (2.19)] if one has

$$|\omega \cdot k|^{-1} \le C|k|^\nu \quad \text{for all } k \in \mathbb{Z}^d \setminus \Lambda_\omega. \tag{9}$$

This restriction is due to the fact that, in the process of averaging, we will deal with the classical problem of small denominators in KAM theory. To keep an example in mind, note that $\omega = (1, \dots, 1)$ is obviously partially Diophantine.[4]

We will make use of some analyticity assumptions on the symbols V and A in the following sense:

Definition 5. Let $s > 0$. We say that $a \in L^1(\mathbb{R}^{2d})$ belongs to the space $\mathcal{A}_s$ if

$$\|a\|_s := \int_{\mathbb{R}^{2d}} |\hat{a}(w)| e^{s\|w\|}\, dw < \infty,$$

where $\hat{a}$ denotes the Fourier transform of a and $\|w\|$ the Euclidean norm on $\mathbb{R}^{2d}$.

Let $\rho, s > 0$. We introduce the space $\mathcal{A}_{\rho,s}$ of functions $a \in L^1(\mathbb{R}^{2d})$ such that

$$\|a\|_{\rho,s} := \frac{1}{(2\pi)^d} \sum_{k\in\mathbb{Z}^d} \|a_k\|_s\, e^{\rho|k|} < \infty, \tag{10}$$

[4]In that example, the flow is periodic and we are in the same situation as in [Asch and Lebeau 2003].

where

$$a_k(z) = \int_{\mathbb{T}^d} a \circ \Phi_\tau^H(z) e^{-ik\cdot\tau}\, d\tau, \quad k \in \mathbb{Z}^d,$$

with Φ_τ^H defined by (15).

Remark 6. Observe that, for any a element in $\mathcal{A}_s$ and for every multi-index $\alpha \in \mathbb{Z}_+^d$, $\widehat{\partial^\alpha a}$ belongs to L^1. Hence, a is smooth and one has $\partial^\alpha a \in L^\infty$ for every $\alpha \in \mathbb{Z}_+^d$. Hence, any element in $\mathcal{A}_s$ belongs to the class $S(1)$ of symbols that are amenable to semiclassical calculus on $\mathbb{R}^d$. In particular, by [Zworski 2012, Lemma 4.10], one has,

$$\text{for all } a \in \mathcal{A}_s, \quad \| \operatorname{Op}_\hbar^w(a)\|_{\mathcal{L}(L^2)} \leq C_{d,s} \|a\|_s. \tag{11}$$

As a consequence of (30), one can show that $\|a\|_s \leq \|a\|_{\rho,s}$ for all $\rho > 0$.

Our next result reads:

Theorem 7. *Suppose that A and V belong to the space $\mathcal{A}_{\rho,s}$ for some fixed $\rho, s > 0$ and that $\langle A\rangle \geq 0$. Assume also that ω is partially Diophantine and that, for every $(x, \xi) \in H^{-1}(1) \cap \langle A\rangle^{-1}(0)$, there exists $T > 0$ such that*

$$\langle A\rangle \circ \phi_T^{\langle V\rangle}(x, \xi) > 0.$$

Then there exists $\varepsilon := \varepsilon(A, V) > 0$ such that, for

$$\delta_\hbar = \hbar,$$

and for any sequence of solutions to (4) with $r_\hbar = 0$,

$$\beta \geq \varepsilon. \tag{12}$$

This theorem shows that eigenvalues of the nonselfadjoint operator $\widehat{P}_\hbar$ cannot accumulate on the boundary of the strip given by Proposition 2. Compared with Theorem 3, it only deals with the case of true eigenvalues and it does not seem that a good resolvent estimate can be easily deduced from the proof below. Finally, for the sake of simplicity, we also supposed that $\delta_\hbar = \hbar$ but it is most likely that the argument can be applied when $\delta_\hbar$ does not go to 0 too slowly.

2. The classical harmonic oscillator

The Hamiltonian equations corresponding to H are given by

$$\begin{cases} \dot{x}_j = \omega_j \xi_j, \\ \dot{\xi}_j = -\omega_j x_j, \quad j = 1, \dots, d. \end{cases} \tag{13}$$

Hence, we can write the solution to this system as a superposition of d-independent commuting flows:

$$(x(t), \xi(t)) = \phi_t^H(x, \xi) := \phi_{\omega_d t}^{H_d} \circ \cdots \circ \phi_{\omega_1 t}^{H_1}(x, \xi), \quad (x, \xi) \in \mathbb{R}^{2d}, \quad t \in \mathbb{R},$$

where $H_j(x, \xi) = \frac{1}{2}(x_j^2 + \xi_j^2)$ and $\phi_t^{H_j}(x, \xi)$ denotes the associated Hamiltonian flow. In other words, the solution to (13) can be written in terms of the unitary block matrices

$$\begin{pmatrix} x_j(t) \\ \xi_j(t) \end{pmatrix} = \begin{pmatrix} \cos(\omega_j t) & \sin(\omega_j t) \\ -\sin(\omega_j t) & \cos(\omega_j t) \end{pmatrix} \begin{pmatrix} x_j \\ \xi_j \end{pmatrix}, \quad j = 1, \dots, d. \tag{14}$$

Observe that each flow $\phi_t^{H_j}$ is periodic with period 2π. We now introduce the transformation

$$\Phi_\tau^H := \phi_{t_d}^{H_d} \circ \cdots \circ \phi_{t_1}^{H_1}, \quad \tau = (t_1, \dots, t_d) \in \mathbb{R}^d. \tag{15}$$

Note that $\tau \mapsto \Phi_\tau^H$ is $2\pi\mathbb{Z}^d$-periodic; therefore we can view it as a function on the torus $\mathbb{T}^d := \mathbb{R}^d/2\pi\mathbb{Z}^d$. Considering now the submodule

$$\Lambda_\omega := \{k \in \mathbb{Z}^d : k \cdot \omega = 0\},$$

we can define the minimal torus

$$\mathbb{T}_\omega := \Lambda_\omega^\perp/(2\pi\mathbb{Z}^d \cap \Lambda_\omega^\perp),$$

where $\Lambda_\omega^\perp$ denotes the linear space orthogonal to Λ_ω. The dimension of $\mathbb{T}_\omega$ is $d_\omega = d - \operatorname{rk} \Lambda_\omega$. Kronecker's theorem states that the family of probability measures on $\mathbb{T}^d$ defined by

$$\frac{1}{T}\int_0^T \delta_{t\omega}\, dt$$

converges (for the weak-$\star$ topology) to the normalized Haar measure ν_ω on the subtorus $\mathbb{T}_\omega \subset \mathbb{T}^d$.

For any function $a \in \mathcal{C}^\infty(\mathbb{R}^{2d})$, we have $a \circ \phi_t^H = a \circ \Phi_{t\omega}^H$. Thus, we can write the average $\langle a \rangle$ of a by the flow ϕ_t^H as

$$\langle a \rangle(x,\xi) = \lim_{T\to\infty} \frac{1}{T}\int_0^T a \circ \Phi_{t\omega}^H(x,\xi)\, dt = \int_{\mathbb{T}_\omega} a \circ \Phi_\tau^H(x,\xi)\nu_\omega(d\tau) \in \mathcal{C}^\infty(\mathbb{R}^{2d}). \tag{16}$$

Recall that the energy hypersurfaces $H^{-1}(E) \subset \mathbb{R}^{2d}$ are compact for every $E \geq 0$. For $E > 0$, due to the complete integrability of H, these hypersurfaces are foliated by the invariant tori: $\{\Phi_\tau^H(x,\xi) : \tau \in \mathbb{T}_\omega\}$. Note that some invariant tori of the energy hypersurface $H^{-1}(E)$, $E > 0$, may have dimension less than d_ω. For instance, if $\omega = (1, \pi)$ then $d_\omega = 2$, but the torus $\{\Phi_\tau^H(0,1,0,1) : \tau \in \mathbb{T}_\omega\} \subset H^{-1}(\pi)$ has dimension 1.

Observe also that $1 \leq d_\omega \leq d$. In the case $d_\omega = 1$ and $\omega = \omega_1(1, \dots, , 1)$, the flow ϕ_t^H is $(2\pi/\omega_1)$-periodic. On the other hand, if $d_\omega = d$, then, for every $a \in \mathcal{C}^\infty(\mathbb{R}^{2d})$, there exists $\mathcal{I}(a) \in \mathcal{C}^\infty(\mathbb{R}^d)$ such that $\langle a \rangle(z) = \mathcal{I}(a)(H_1(z), \dots, H_d(z))$. In particular, for every a and b in $\mathcal{C}^\infty(\mathbb{R}^{2d})$, one has $\{\langle a \rangle, \langle b \rangle\} = 0$ whenever $d_\omega = d$.

To conclude this section, we prove the following lemma:

Lemma 8. *If $a \in \mathcal{A}_s$ then $\langle a \rangle \in \mathcal{A}_s$ and $\|\langle a \rangle\|_s \leq \|a\|_s$.*

Proof. By (16), we can write the Fourier transform of $\langle a \rangle$ as

$$\widehat{\langle a \rangle}(x,\xi) = \int_{\mathbb{T}_\omega} \widehat{a \circ \Phi_\tau^H}(x,\xi)\nu_\omega(d\tau).$$

Moreover, since $\widehat{a \circ \Phi_\tau^H}(x,\xi) = \hat{a} \circ \Phi_\tau^H(x,\xi)$ thanks to (14), we have $\widehat{\langle a \rangle} = \langle \hat{a} \rangle$. Thus, using (14) one more time, one finds

$$\begin{aligned}\|\langle a \rangle\|_s &= \int_{\mathbb{R}^{2d}} |\langle \hat{a} \rangle(z)| e^{s|z|}\, dz \leq \int_{\mathbb{T}_\omega}\int_{\mathbb{R}^{2d}} |\hat{a} \circ \Phi_\tau^H(z)| e^{s|z|}\, dz\, \nu_\omega(d\tau) \\ &= \int_{\mathbb{R}^{2d}} |\hat{a}(z)| e^{s|z|}\, dz = \|a\|_s.\end{aligned}$$

□

3. Proof of Theorem 3

We now give the proof of our main result in the $\mathcal{C}^\infty$ case. Before doing that, we briefly recall the proof of Proposition 2 in order to make the proof of Theorem 3 more comprehensive. Note that we use the following convention for the scalar product on $\mathbb{R}^d$:

$$\langle u, v\rangle_{L^2} = \int_{\mathbb{R}^d} u(x)\overline{v(x)}\,dx.$$

3.1. ***Proof of Proposition 2.*** Let $\lambda_\hbar = \alpha_\hbar + i\hbar\beta_\hbar$ be a sequence of (pseudo-)eigenvalues satisfying (4). Denote by $(v_\hbar)_{\hbar\to 0^+}$ the corresponding sequence of normalized quasimodes. Introduce the Wigner distribution $W^\hbar_{v_\hbar} \in \mathcal{D}'(\mathbb{R}^{2d})$ associated to the function $v_\hbar$:

$$W^\hbar_{v_\hbar} : \mathcal{C}^\infty_c(\mathbb{R}^{2d}) \ni a \longmapsto W^\hbar_{v_\hbar}(a) := \langle \mathrm{Op}^w_\hbar(a)v_\hbar, v_\hbar\rangle_{L^2(\mathbb{R}^d)}.$$

According to [Zworski 2012, Chapter 5] and modulo extracting a subsequence, there exists a probability measure μ carried by $H^{-1}(1)$ such that $W^\hbar_{v_\hbar} \rightharpoonup \mu$. The measure μ is called the semiclassical measure associated to the (sub-)sequence $(v_\hbar)_{\hbar\to 0^+}$. Note that these properties of the limit points follow from the facts that $v_\hbar$ is normalized and that $\widehat{H}_\hbar v_\hbar = v_\hbar + o_{L^2}(1)$. We will now make use of the eigenvalue equation (4) to derive an invariance property of μ. Using the symbolic calculus for Weyl pseudodifferential operators [Zworski 2012, Chapter 4], we have, for every $a \in \mathcal{C}^\infty_c(\mathbb{R}^{2d}, \mathbb{R})$,

$$\langle [\widehat{H}_\hbar + \delta_\hbar \widehat{V}_\hbar, \mathrm{Op}^w_\hbar(a)]v_\hbar, v_\hbar\rangle_{L^2(\mathbb{R}^d)} = \frac{\hbar}{i}\langle \mathrm{Op}^w_\hbar(\{H, a\})v_\hbar, v_\hbar\rangle_{L^2(\mathbb{R}^d)} + O(\hbar(\delta_\hbar + \hbar)).$$

On the other hand, using that $v_\hbar$ is a quasimode of $\widehat{P}_\hbar$ and the composition rule for the Weyl quantization [Zworski 2012, Chapter 4], we also have

$$\langle [\widehat{H}_\hbar + \delta_\hbar \widehat{V}_\hbar, \mathrm{Op}^w_\hbar(a)]v_\hbar, v_\hbar\rangle_{L^2(\mathbb{R}^d)} = 2i\hbar\langle \mathrm{Op}^w_\hbar(a(A - \beta_\hbar))v_\hbar, v_\hbar\rangle_{L^2(\mathbb{R}^d)} + O(\|r_\hbar\|) + O(\hbar^3).$$

Note that there is no $O(\hbar^2)$-term due to the fact that a is real-valued and to the symmetries of the Weyl quantization. Passing to the limit $\hbar \to 0^+$ and recalling that $\|r_\hbar\| = o(\hbar)$, one finds that $\mu(\{H, a\}) = 2\mu((\beta - A)a)$ for every a in $\mathcal{C}^\infty_c(\mathbb{R}^d)$. This is equivalent to the fact that, for every $t \in \mathbb{R}$ and for every $a \in \mathcal{C}^\infty_c(\mathbb{R}^{2d})$, one has

$$\int_{\mathbb{R}^{2d}} a(z)\mu(dz) = \int_{\mathbb{R}^{2d}} a \circ \phi^H_t(z) e^{2\int_0^t (A-\beta)\circ\phi^H_s(z)\,ds}\mu(dz). \tag{17}$$

Taking a to be equal to 1 in a neighborhood of $H^{-1}(1)$, identity (17) implies

$$e^{2\beta t} = \int_{\mathbb{R}^{2d}} e^{2\int_0^t A\circ\phi^H_s(z)\,ds}\mu(dz) \quad \text{for all } t \in \mathbb{R}, \tag{18}$$

from which Proposition 2 follows thanks to (3). In the case where $\beta = 0$ and $A \geq 0$, one can deduce from (18) that,

$$\text{for all } t \in \mathbb{R}, \quad \mathrm{supp}(\mu) \subset H^{-1}(1) \cap \{z : A \circ \phi^H_t(z) = 0\}.$$

Hence, we can record the following useful lemma:

Lemma 9. *Suppose that $A \geq 0$. Let μ be a semiclassical measure associated to the sequence $(v_\hbar)_{\hbar\to 0^+}$ satisfying (4) with $\beta = 0$ and $r_\hbar = o(\hbar)$. Then*

$$\operatorname{supp}\mu \subset \{z \in H^{-1}(1) : \langle A\rangle(z) = 0\}. \tag{19}$$

3.2. ***Proof of Theorem 3.*** Let us now reproduce the same argument but suppose that $a = \langle a\rangle$, implying in particular that $\{H, \langle a\rangle\} = 0$. From this, we get

$$\langle[\widehat{H}_\hbar + \delta_\hbar \widehat{V}_\hbar, \operatorname{Op}_\hbar^w(\langle a\rangle)]v_\hbar, v_\hbar\rangle_{L^2(\mathbb{R}^d)} = \frac{\hbar\delta_\hbar}{i}\langle \operatorname{Op}_\hbar^w(\{V, \langle a\rangle\})v_\hbar, v_\hbar\rangle_{L^2(\mathbb{R}^d)} + O(\hbar^3).$$

As before, recalling that a is real-valued, one still has

$$\langle[\widehat{H}_\hbar + \delta_\hbar \widehat{V}_\hbar, \operatorname{Op}_\hbar^w(\langle a\rangle)]v_\hbar, v_\hbar\rangle_{L^2(\mathbb{R}^d)} = 2i\hbar\langle \operatorname{Op}_\hbar^w(\langle a\rangle(A-\beta_\hbar))v_\hbar, v_\hbar\rangle_{L^2(\mathbb{R}^d)} + O(\|r_\hbar\|) + O(\hbar^3).$$

Hence, one gets

$$\big\langle \operatorname{Op}_\hbar^w\big((2(A-\beta_\hbar) + \delta_\hbar X_V)\langle a\rangle\big)v_\hbar, v_\hbar\big\rangle_{L^2(\mathbb{R}^d)} = O(\|r_\hbar\|\hbar^{-1}) + O(\hbar^2),$$

where X_V is the Hamiltonian vector field of V. Suppose now that $A \geq 0$ and $\langle a\rangle \geq 0$. From the Fefferman–Phong inequality [Zworski 2012, Chapter 4], one knows that there exists some constant $C > 0$ such that

$$2\beta_\hbar\langle \operatorname{Op}_\hbar^w(\langle a\rangle)v_\hbar, v_\hbar\rangle_{L^2(\mathbb{R}^d)} \geq \delta_\hbar\langle \operatorname{Op}_\hbar^w(X_V\langle a\rangle)v_\hbar, v_\hbar\rangle_{L^2(\mathbb{R}^d)} - C(\hbar^2 + \|r_\hbar\|\hbar^{-1}),$$

where the constant C depends only on A, V, and a. Now, we fix $R > 0$ and we would like to show that $\liminf_{\hbar\to 0^+} \beta_\hbar/\delta_\hbar > R$ provided that $\delta_\hbar \geq \varepsilon_R^{-1}\hbar^2$ and that $\|r_\hbar\| \leq \varepsilon_R\hbar\delta_\hbar$ for some small enough $\varepsilon_R > 0$ (to be determined later on). To that end, we proceed by contradiction and suppose that, up to an extraction, one has $2\beta_\hbar/\delta_\hbar \to c_0 \in [0, 2R]$ (in particular $\beta = 0$). One finally gets, after letting $\hbar \to 0^+$,

$$c_0\mu(\langle a\rangle) \geq \mu(X_V\langle a\rangle) - C\varepsilon_R \tag{20}$$

for some $C \geq 0$ depending on A, V, and a. Using one more time Lemma 9, one can also deduce that μ is invariant by ϕ_t^H. Hence,

$$\mu(\{V, \langle a\rangle\}) = \mu(\{\langle V\rangle, \langle a\rangle\}),$$

which implies

$$c_0\mu(\langle a\rangle) \geq \mu(X_{\langle V\rangle}\langle a\rangle) - C\varepsilon_R. \tag{21}$$

By our geometric control condition (6) and since $H^{-1}(1) \cap \langle A\rangle^{-1}(0)$ is compact, there exist $T_1 > 0$ and $\varepsilon_0 > 0$ such that

$$\int_0^{T_1} \langle A\rangle \circ \phi_t^{\langle V\rangle}(z)\,dt > \varepsilon_0 \quad \text{for all } z \in H^{-1}(1) \cap \langle A\rangle^{-1}(0),$$

where $\phi_t^{\langle V\rangle}$ is the flow generated by $X_{\langle V\rangle}$. Up to the fact that we may have to increase the value of $C > 0$ (in a way that depends only on T_1, A, V, and a), we can suppose that (21) holds uniformly for every function $\langle a\rangle \circ \phi_t^{\langle V\rangle}$ with $0 \leq t \leq T_1$; i.e., for every $t \in [0, T_1]$,

$$c_0\mu(\langle a\rangle \circ \phi_t^{\langle V\rangle}) \geq \mu(\{\langle V\rangle, \langle a\rangle\} \circ \phi_t^{\langle V\rangle}) - C\varepsilon_R.$$

This is equivalent to the fact that $\frac{d}{dt}\big(e^{-c_0 t}\int_{\mathbb{R}^{2d}}\langle a\rangle\circ\phi_t^{\langle V\rangle}\,d\mu\big)\leq C\varepsilon_R e^{-c_0 t}$ for every $t\in[0,T_1]$. Hence, if $c_0\neq 0$, one finds that, for every $t\in[0,T_1]$,

$$\int_{\mathbb{R}^{2d}}\langle a\rangle\circ\phi_t^{\langle V\rangle}(z)\mu(dz)\leq e^{c_0 t}\int_{\mathbb{R}^{2d}}\langle a\rangle(z)\mu(dz)+\frac{C\varepsilon_R(e^{tc_0}-1)}{c_0}. \tag{22}$$

We now apply this inequality with $a=A$ and integrate over the interval $[0,T_1]$. In that way, we obtain

$$\varepsilon_0<\int_0^{T_1}\int_{\mathbb{R}^{2d}}\langle A\rangle\circ\phi_t^{\langle V\rangle}(z)\mu(dz)\,dt\leq\int_0^{T_1}\frac{C\varepsilon_R(e^{tc_0}-1)}{c_0}\,dt\leq\frac{C\varepsilon_R T_1(e^{T_1c_0}-1)}{c_0}.$$

Observe that, for $c_0=0$, we would get the upper bound $C\varepsilon_R T_1^2$. In both cases, this yields the expected contradiction by taking ε_R small enough (in a way that depends only on R, A, and V) and it concludes the proof of Theorem 3.

Remark 10. Note that we could get the conclusion faster under the stronger geometric assumption

$$\text{for all } z\in H^{-1}(1)\cap\langle A\rangle^{-1}(0),\quad \{\langle A\rangle,\langle V\rangle\}(z)\neq 0, \tag{23}$$

which implies (but is not equivalent to) the geometric control condition (6) of Theorem 3. Together with (21), this yields the upper bound

$$\mu(X_{\langle V\rangle}\langle A\rangle)\leq C\varepsilon_R.$$

Hence, provided $\varepsilon_R>0$ is chosen small enough in a way that depends only on A and V (but not R), we get a contradiction. This shows that, for a small enough choice of $\varepsilon_R>0$, one has in fact $\beta_\hbar\gg\delta_\hbar$ under the geometric condition (23).

4. The averaging method

From this point on in the article, we will make the assumption that

$$\delta_\hbar=\hbar.$$

This will slightly simplify the exposition and it should a priori be possible to extend the results provided $\hbar\leq\delta_\hbar$ does not go to 0 too slowly. In this section, we briefly recall how to perform a semiclassical averaging method in the context of nonselfadjoint operators following [Sjöstrand 2000; Hitrik 2002]. For that purpose, we define

$$\widehat{F}_\hbar:=\mathrm{Op}_\hbar^w(F_1+iF_2),$$

where F_1 and F_2 are two real-valued and smooth functions on $\mathbb{R}^{2d}$ that will be determined later on. We make the assumption that all the derivatives (at every order) of F_1 and F_2 are bounded. For every t in $[0,1]$, we set $\mathcal{F}_\hbar(t)=e^{it\widehat{F}_\hbar}$. By [Engel and Nagel 2000, Theorem III.1.3], the family $\mathcal{F}_\hbar(t)$ defines a strongly continuous group (note that $\mathcal{F}_\hbar$ is invertible) on $L^2(\mathbb{R}^d)$ such that

$$\|\mathcal{F}_\hbar(t)\|_{\mathcal{L}(L^2)}\leq e^{|t|\|\mathrm{Op}_\hbar^w(F_2)\|_{\mathcal{L}(L^2)}}. \tag{24}$$

For simplicity, we shall set $\mathcal{F}_\hbar = \mathcal{F}_\hbar(1)$ and we will study the properties of the conjugated operator

$$\widehat{Q}_\hbar := \mathcal{F}_\hbar \widehat{P}_\hbar \mathcal{F}_\hbar^{-1}$$

for appropriate choices of F_1 and F_2. Using the conventions of [Zworski 2012, Chapter 4], symbols of order $m \in \mathbb{R}$ are defined by

$$S(\langle z\rangle^m) := \{(a_\hbar)_{0\leq\hbar\leq 1} \in \mathcal{C}^\infty(\mathbb{R}^{2d}, \mathbb{C}) : \text{for all } \alpha \in \mathbb{N}^{2d},\ |\partial^\alpha a(z)| \leq C_\alpha \langle z\rangle^m\},$$

where $\langle z\rangle = (1+\|z\|^2)^{1/2}$. We shall denote by $\Psi_\hbar^m$ the set of all operators of the form $\operatorname{Op}_\hbar^w(a)$ with $a \in S(\langle z\rangle^m)$.

4.1. ***Semiclassical conjugation.*** Writing the Taylor expansion, one knows that, for every a in $S(\langle z\rangle^m)$,

$$\mathcal{F}_\hbar \operatorname{Op}_\hbar^w(a)\mathcal{F}_\hbar^{-1} = \operatorname{Op}_\hbar^w(a) + i[\widehat{F}_\hbar, \operatorname{Op}_\hbar^w(a)] + \int_0^1 (1-t)\mathcal{F}_\hbar(t)[\widehat{F}_\hbar, [\widehat{F}_\hbar, \operatorname{Op}_\hbar^w(a)]]\mathcal{F}_\hbar(-t)\, dt. \tag{25}$$

Observe from the composition rules for semiclassical pseudodifferential operators [Zworski 2012, Chapter 4] that $[\widehat{F}_\hbar, [\widehat{F}_\hbar, \operatorname{Op}_\hbar^w(a)]]$ is an element of $\hbar^2\Psi_\hbar^m$. Moreover, a direct extension of the Egorov theorem [Zworski 2012, Theorem 11.1] to the nonselfadjoint framework shows that the third term in the right-hand side is in fact an element of $\hbar^2\Psi_\hbar^m$. Then one can verify from the composition rules for pseudodifferential operators that

$$\mathcal{F}_\hbar \operatorname{Op}_\hbar^w(a)\mathcal{F}_\hbar^{-1} = \operatorname{Op}_\hbar^w(a) + \hbar \operatorname{Op}_\hbar^w(\{F_1, a\}) + i\hbar \operatorname{Op}_\hbar^w(\{F_2, a\}) + \hbar^2 \widehat{R}_\hbar,$$

where $\widehat{R}_\hbar$ is an element in $\Psi_\hbar^m$. Applying this equality to the operator $\widehat{P}_\hbar$, one finds

$$\widehat{Q}_\hbar = \widehat{P}_\hbar + \hbar \operatorname{Op}_\hbar^w(\{F_1, H\}) + i\hbar \operatorname{Op}_\hbar^w(\{F_2, H\}) + \hbar^2 \widehat{R}_\hbar, \tag{26}$$

where $\widehat{R}_\hbar$ is now an element in $\Psi_\hbar^2$. We now aim at choosing F_1 and F_2 in such a way that

$$\{F_1, H\} + V = \langle V\rangle \quad \text{and} \quad \{F_2, H\} + A = \langle A\rangle. \tag{27}$$

If we are able to do so, then we will have

$$\mathcal{F}_\hbar \widehat{P}_\hbar \mathcal{F}_\hbar^{-1} = \widehat{H}_\hbar + \hbar \operatorname{Op}_\hbar^w(\langle V\rangle) + i\hbar \operatorname{Op}_\hbar^w(\langle A\rangle) + \hbar^2 \widehat{R}_\hbar. \tag{28}$$

4.2. ***Solving cohomological equations.*** In order to solve cohomological-type equations like (27), we need to make a few Diophantine restrictions on ω. Let $g \in \mathcal{C}^\infty(\mathbb{R}^{2d})$ be any smooth function such that $\langle g\rangle = 0$ and all of whose derivatives (at any order) are bounded. We look for another function $f \in \mathcal{C}^\infty(\mathbb{R}^{2d})$ all of whose derivatives (at any order) are bounded and which solves the cohomological equation

$$\{H, f\} = g. \tag{29}$$

We then apply this result with $g = V - \langle V\rangle$ (resp. $A - \langle A\rangle$) in order to find $f = F_1$ (resp. F_2).

For any $f \in \mathcal{C}^\infty(\mathbb{R}^{2d})$ all of whose derivatives (at any order) are bounded, we can write $f \circ \Phi_\tau^H$ as a Fourier series in $\tau \in \mathbb{T}^d$:

$$f \circ \Phi_\tau^H(x, \xi) = \sum_{k\in\mathbb{Z}^d} f_k(x, \xi)\frac{e^{ik\cdot\tau}}{(2\pi)^d}, \quad f_k(x, \xi) := \int_{\mathbb{T}^d} f \circ \Phi_\tau^H(x, \xi)e^{-ik\cdot\tau}\, d\tau. \tag{30}$$

Notice that $f_k \circ \Phi_\tau^H = f_k\, e^{ik\cdot\tau}$ and that, for $\tau = 0$, we have $f = (2\pi)^{-d}\sum_k f_k$. Recalling (16) and the definition (8) of Λ_ω, one has

$$\langle f\rangle \circ \Phi_\tau^H(x,\xi) = \sum_{k\in\mathbb{Z}^d} f_k(x,\xi)\left(\lim_{T\to+\infty}\frac{1}{T(2\pi)^d}\int_0^T e^{ik.(\tau+t\omega)}\,dt\right) = \frac{1}{(2\pi)^d}\sum_{k\in\Lambda_\omega} f_k(x,\xi)e^{ik\cdot\tau}.$$

In particular, as $\langle g\rangle\circ\Phi_\tau^H = 0$ for every $\tau\in\mathbb{T}^d$, one finds that $g_k = 0$ for every $k\in\Lambda_\omega$ and thus,

$$\text{for all } \tau\in\mathbb{T}^d,\quad g\circ\Phi_\tau^H(x,\xi) = \frac{1}{(2\pi)^d}\sum_{k\in\mathbb{Z}^d\setminus\Lambda_\omega} g_k(x,\xi)e^{ik\cdot\tau}.$$

Observe also that, if f is a solution of (29), then so is $f+\lambda\langle f\rangle$ for any $\lambda\in\mathbb{R}$, since $\{H,\langle f\rangle\}=0$ thanks to (16). Thus, we can try to solve the cohomological equation (29) by supposing $f\circ\Phi_\tau^H$ to be of the form

$$f\circ\Phi_\tau^H(x,\xi) = \frac{1}{(2\pi)^d}\sum_{k\in\mathbb{Z}^d\setminus\Lambda_\omega} f_k(x,\xi)e^{ik\cdot\tau},$$

and writing

$$\{H, f\circ\Phi_\tau^H\} = \frac{d}{dt}(f\circ\Phi_{\tau+t\omega}^H)|_{t=0} = \frac{1}{(2\pi)^d}\sum_{k\in\mathbb{Z}^d\setminus\Lambda_\omega} ik\cdot\omega\, f_k e^{ik\cdot\tau}.$$

Hence, if we set

$$f\circ\Phi_\tau^H(x,\xi) = \frac{1}{(2\pi)^d}\sum_{k\in\mathbb{Z}^d\setminus\Lambda_\omega}\frac{1}{ik\cdot\omega}\,g_k(x,\xi)e^{ik\cdot\tau}, \tag{31}$$

then f will solve (29) (at least formally). It is not difficult to see that, unless we impose some quantitative restriction on how fast $|k\cdot\omega|^{-1}$ can grow, the solutions given formally by (31) may fail to be even distributions — see for instance [de la Llave 2001, Exercise 2.16]. On the other hand, if ω is partially Diophantine, and $g\in\mathcal{C}^\infty(\mathbb{R}^{2d})$ has all its derivatives (at any order) bounded and is such that $\langle g\rangle = 0$, then (31) defines a smooth solution $f\in\mathcal{C}^\infty(\mathbb{R}^{2d})$ of (29) all of whose derivatives (at any order) are bounded. As a special case, we observe that, if $\omega = (1,\dots,1)$, then an explicit solution of (29) is given by

$$f = -\frac{1}{2\pi}\int_0^{2\pi}\int_0^t g\circ\phi_s^H\,ds\,dt. \tag{32}$$

4.3. ***Proof of Theorem 7.*** We now turn to the proof of Theorem 7 and, to that aim, we should exploit the analyticity assumptions on A and V in order to improve the result of Theorem 3 when $r_\hbar = 0$ in (4). It means that we are not considering anymore quasimodes but true eigenmodes. Hence, from this point on in the article,

$$r_\hbar = 0.$$

The point of using analyticity is that the symbolic calculus on the family of spaces $\mathcal{A}_s$ is extremely well behaved — see the Appendix for a brief review. This will allow us to construct a second normal form for the operator $\widehat{P}_\hbar$ via conjugation by a second operator so that the nonselfadjoint part of the operator is averaged by the two flows ϕ_t^H and $\phi_t^{\langle V\rangle}$.

Recall from (28) that

$$\mathcal{F}_\hbar \widehat{P}_\hbar \mathcal{F}_\hbar^{-1} = \widehat{H}_\hbar + \hbar \operatorname{Op}_\hbar^w(\langle V\rangle) + i\hbar \operatorname{Op}_\hbar^w(\langle A\rangle) + \hbar^2 \widehat{R}_\hbar. \tag{33}$$

Let us now make a few additional comments using the fact that A and V belong to some space $\mathcal{A}_s$. First of all, according to Lemma 8, we know that, as soon as A and V belong to the space $\mathcal{A}_s$, both $\langle A\rangle$ and $\langle V\rangle$ belong[5] to the space $\mathcal{A}_s$. Moreover the functions F_1 and F_2 used to define $\mathcal{F}_\hbar$ are constructed from A and V using (31). In particular, by (9) and for every $0<\sigma<\rho$, the following inequalities hold:

$$\|F_1\|_s \le \|F_1\|_{\rho-\sigma,s} \lesssim_{\rho,s} \|V\|_{\rho,s} \quad \text{and} \quad \|F_2\|_s \le \|F_2\|_{\rho-\sigma,s} \lesssim_{\rho,s} \|A\|_{\rho,s}.$$

We can make use of this regularity information to analyze the regularity of the remainder term $\widehat{R}_\hbar$ in (33). Recall that part of this term comes from the remainder term when we apply the composition formula to $[\operatorname{Op}_\hbar^w(A), \operatorname{Op}_\hbar^w(F_j)]$ and to $[\operatorname{Op}_\hbar^w(V), \operatorname{Op}_\hbar^w(F_j)]$ for $j=1,2$. In that case, Lemma 15 from the Appendix tells us that the remainder is a pseudodifferential operator whose symbol belongs to $\mathcal{A}_{s-\sigma}$ for every $0<\sigma<s$. There is another contribution coming from the integral term in the Taylor formula (25) with $\operatorname{Op}_\hbar^w(a)$ replaced by $\widehat{P}_\hbar$. For that term, we first make use of Lemma 15 and of the fact that F_j solve cohomological equations[6] (27) in order to verify that the double bracket is a pseudodifferential operator whose symbol belongs to $\mathcal{A}_{s-\sigma}$ for every $0<\sigma<s$. Then, an application of the analytic Egorov lemma from the Appendix (point (1) of Lemma 11 with $G=\hbar(F_1+iF_2)$) shows that this remainder term is still a pseudodifferential operator whose symbol now belongs to $\mathcal{A}_{s-\sigma}$ for every $0<\sigma<s$. To summarize, we have verified that $\widehat{R}_\hbar = \operatorname{Op}_\hbar^w(R_\hbar)$ with $\|R_\hbar\|_{s-\sigma} \le C_{s,\sigma,\rho}$ for every $0<\sigma<s$ and uniformly for $0<\hbar\le\hbar_0$.

We now perform a second conjugation whose effect will be to replace $\langle A\rangle$ in (33) by a term involving V. Let F_3 be some real-valued element in $\mathcal{A}_{s-\sigma}$ for some $0<\sigma<s$ satisfying $\langle F_3\rangle = F_3$. We set, for $\varepsilon>0$ small enough (independent of $\hbar$),

$$\widetilde{\mathcal{F}}_\hbar(t) := e^{(t/\hbar)\widehat{F}_{3,\hbar}}, \quad t\in[-\varepsilon,\varepsilon],$$

where $\widehat{F}_{3,\hbar} = \operatorname{Op}_\hbar^w(\langle F_3\rangle)$. We can define the new conjugate of $\widehat{H}_\hbar$:

$$\widetilde{\mathcal{F}}_\hbar(-\varepsilon)\mathcal{F}_\hbar \widehat{P}_\hbar \mathcal{F}_\hbar^{-1}\widetilde{\mathcal{F}}_\hbar(\varepsilon) = \widehat{H}_\hbar + \hbar\widetilde{\mathcal{F}}_\hbar(-\varepsilon)\big(\operatorname{Op}_\hbar^w(\langle V\rangle) + i\operatorname{Op}_\hbar^w(\langle A\rangle) + \hbar\widehat{R}_\hbar\big)\widetilde{\mathcal{F}}_\hbar(\varepsilon),$$

where we used that $[\widehat{H}_\hbar, \operatorname{Op}_\hbar^w(\langle F_3\rangle)]=0$. In fact, as H is quadratic in (x,ξ) and as we used the Weyl-quantization, the fact that H and $\langle F_3\rangle$ (Poisson-)commute implies that $[\widehat{H}_\hbar, \operatorname{Op}_\hbar^w(\langle F_3\rangle)]=0$. Suppose now that $\varepsilon\|\langle F_3\rangle\|_{s-\sigma} \le \frac{1}{2}\sigma^2$ so that we can use the (analytic) Egorov lemma, Lemma 11, with $G=iF_3$. This tells us that

$$\widetilde{\mathcal{F}}_\hbar(-\varepsilon)\widehat{R}_\hbar\widetilde{\mathcal{F}}_\hbar(\varepsilon) = \operatorname{Op}_\hbar^w(R_\hbar(\varepsilon)), \tag{34}$$

with $R_\hbar(\varepsilon)$ belonging to $\mathcal{A}_{s-\sigma}$ uniformly for $\hbar$ small enough. Using the conventions of the Appendix, one also has

$$\widetilde{\mathcal{F}}_\hbar(\varepsilon)(\operatorname{Op}_\hbar^w(\langle V\rangle) + i\operatorname{Op}_\hbar^w(\langle A\rangle))\widetilde{\mathcal{F}}_\hbar(-\varepsilon) = \operatorname{Op}_\hbar^w(\Psi_\varepsilon^{iF_3,\hbar}(\langle V\rangle + i\langle A\rangle)). \tag{35}$$

[5]Recall also that $\mathcal{A}_s \subset S(1)$.

[6]This comment is to handle the contribution coming from $\widehat{H}_\hbar$.

Consider now a sequence $(\lambda_\hbar = \alpha_\hbar + i\hbar\beta_\hbar)_{0<\hbar\leq 1}$ solving (4) with $r_\hbar = 0$ and $\beta_\hbar \to \beta$. In particular, one can find a sequence of normalized eigenvectors $(\tilde{v}_\hbar)_{0<\hbar\leq 1}$ such that

$$\widetilde{\mathcal{F}}_\hbar(-\varepsilon)\mathcal{F}_\hbar \widehat{P}_\hbar \mathcal{F}_\hbar^{-1}\widetilde{\mathcal{F}}_\hbar(\varepsilon)\tilde{v}_\hbar = \lambda_\hbar \tilde{v}_\hbar.$$

Implementing (34) and (35), one obtains

$$\operatorname{Im}\langle \operatorname{Op}_\hbar^w(\Psi_\varepsilon^{iF_3,\hbar}(\langle V\rangle + i\langle A\rangle))\tilde{v}_\hbar, \tilde{v}_\hbar\rangle + O(\hbar) = \frac{1}{\hbar}\operatorname{Im}\langle \widetilde{\mathcal{F}}_\hbar(-\varepsilon)\mathcal{F}_\hbar \widehat{P}_\hbar \mathcal{F}_\hbar^{-1}\widetilde{\mathcal{F}}_\hbar(\varepsilon)\tilde{v}_\hbar, \tilde{v}_\hbar\rangle = \beta_\hbar.$$

From point (3) of Lemma 11, one then finds

$$\beta_\hbar = \langle \operatorname{Op}_\hbar^w(\langle A\rangle - \varepsilon\{\langle F_3\rangle, \langle V\rangle\})\tilde{v}_\hbar, \tilde{v}_\hbar\rangle + O(\varepsilon^2) + O(\hbar).$$

Up to another extraction, we can suppose that the sequence $(\tilde{v}_\hbar)_{\hbar>0}$ has a unique semiclassical measure $\tilde{\mu}$ which is still a probability measure carried by $H^{-1}(1)$. Letting $\hbar \to 0^+$, one finds

$$\beta = \tilde{\mu}(\langle A\rangle + \varepsilon\{\langle V\rangle, \langle F_3\rangle\}) + O(\varepsilon^2).$$

Given $0<\sigma<s$, suppose now that we can pick F_3 in $\mathcal{A}_{s-\sigma}$ such that $\{\langle F_3\rangle, \langle V\rangle\} < 0$ on $\langle A\rangle^{-1}(0)\cap H^{-1}(1)$. Then, one can find some $c_0 > 0$ such that $c_0\varepsilon + O(\varepsilon^2) \leq \beta$. In particular, β cannot be taken equal to 0, which concludes the proof of Theorem 7 except for the proof of the existence of F_3.

Let us now show that the geometric control assumption (6) of Theorem 7 implies the existence of F_3. Since $\langle A\rangle$ and $\langle V\rangle$ belong to $\mathcal{A}_s$, Remark 12 from the Appendix and the compactness of the set $H^{-1}(1)\cap\langle A\rangle^{-1}(0)$ show that, for every $0<\sigma<s$, there exists some small enough $t_0>0$ such that

$$F_3(z) := \int_0^{t_0}\left(\int_0^t \langle A\rangle\circ\phi_\tau^{\langle V\rangle}(z)\,d\tau\right)dt$$

belongs to $\mathcal{A}_{s-\sigma}$. One has, for every $z\in H^{-1}(1)\cap\langle A\rangle^{-1}(0)$,

$$\{\langle V\rangle, F_3\}(z) = \int_0^{t_0}\langle A\rangle\circ\phi_t^{\langle V\rangle}(z)\,dt.$$

It remains to verify that this quantity is positive for every z_0 in $\langle A\rangle^{-1}(0)\cap H^{-1}(1)$. Still using Remark 12, one has the analytic expansion

$$\langle A\rangle\circ\phi_t^{\langle V\rangle}(z) = \sum_{j=0}^{\infty}\frac{t^j}{j!}\operatorname{Ad}^j_{\langle V\rangle}(\langle A\rangle)(z), \tag{36}$$

uniformly for $t\in[-t_0,t_0]$ and $z\in H^{-1}(1)$. This implies that, if we fix some z_0 in $H^{-1}(1)$, then the map $t\mapsto\langle A\rangle\circ\phi_t^{\langle V\rangle}(z)$ is analytic on $\mathbb{R}$. Now, given some $z_0\in\langle A\rangle^{-1}(0)\cap H^{-1}(1)$, there exists some z_1 in the orbit of z_0 such that $\langle A\rangle(z_1)>0$ thanks to our geometric control assumption (6). In particular, the analytic map $t\mapsto\langle A\rangle\circ\phi_t^{\langle V\rangle}(z_0)$ is nonconstant and there exists some $j\geq 1$ such that $\operatorname{Ad}^j_{\langle V\rangle}(\langle A\rangle)(z_0)\neq 0$. Hence, $\{\langle V\rangle, F_3\}(z_0)>0$, which concludes the proof.

Appendix: Symbolic calculus on the spaces $\mathcal{A}_s$

We collect some basic lemmas about the quantization of the spaces $\mathcal{A}_s$. We fix $s > 0$ throughout this appendix. Let $a, b \in \mathcal{A}_s$. The operator given by the composition $\mathrm{Op}^w_\hbar(a)\,\mathrm{Op}^w_\hbar(b)$ is another pseudodifferential operator with symbol c given by the Moyal product $c = a \sharp_\hbar b$, which can be written by the following integral formula [Dimassi and Sjöstrand 1999, Chapter 7, p. 79]:

$$c(z) = a \sharp_\hbar b(z) = \frac{1}{(2\pi)^{4d}} \int_{\mathbb{R}^{4d}} \hat{a}(w^*)\hat{b}(z^* - w^*) e^{(i\hbar/2)\varsigma(w^*, z^* - w^*)} e^{i z^* \cdot z} \, dw^* \, dz^*, \tag{37}$$

where $\varsigma(x, \xi, y, \eta) := \xi \cdot y - x \cdot \eta$ is the standard symplectic product and where

$$\hat{a}(w) := \int_{\mathbb{R}^{2d}} e^{-i w \cdot z} a(z)\, dz.$$

We set $[a, b]_\hbar := a \sharp_\hbar b - b \sharp_\hbar a$. Given now $a, G \in \mathcal{A}_s$, the following conjugation formula holds formally:

$$e^{i(t/\hbar)\mathrm{Op}^w_\hbar(G)}\,\mathrm{Op}^w_\hbar(a) e^{-i(t/\hbar)\mathrm{Op}^w_\hbar(G)} = \mathrm{Op}^w_\hbar(\Psi^{G,\hbar}_t a),$$

where

$$\Psi^{G,\hbar}_t a := \sum_{j=0}^{\infty} \frac{1}{j!} \left(\frac{it}{\hbar}\right)^j \mathrm{Ad}^{\sharp_\hbar, j}_G(a), \quad t \in \mathbb{R}, \tag{38}$$

and

$$\mathrm{Ad}^{\sharp_\hbar, j}_G(a) = [G, \mathrm{Ad}^{\sharp_\hbar, j-1}_G(a)]_\hbar, \quad \mathrm{Ad}^{\sharp_\hbar, 0}_G(a) = a.$$

One of the aims of this appendix is to prove the following analytic version of Egorov's theorem:

Lemma 11 (analytic Egorov's lemma). *Let $0 < \sigma < s$. Consider the family of Fourier integral operators $\{\mathcal{G}_\hbar(t) : t \in \mathbb{R}\}$ defined by*

$$\mathcal{G}_\hbar(t) := e^{-(it/\hbar)\widehat{G}_\hbar},$$

where $\widehat{G}_\hbar = \mathrm{Op}^w_\hbar(G)$ for some $G \in \mathcal{A}_s$. Assume

$$|t| < \frac{\sigma^2}{2\|G\|_s}. \tag{39}$$

Then, there exists a constant $C_\sigma > 0$ (depending only on σ) such that, for every $a \in \mathcal{A}_s$,

(1) $\Psi^{G,\hbar}_t a \in \mathcal{A}_{s-\sigma}$;

(2) $\|\Psi^{G,\hbar}_t a - a\|_{s-\sigma} \le C_\sigma |t| \|G\|_s \|a\|_s$;

(3) $\|\Psi^{G,\hbar}_t a - a + t\{G, a\}\|_{s-\sigma} \le C_\sigma |t|^2 \|G\|_s \|a\|_s$ *for some $C_\sigma > 0$ depending only on σ.*

Remark 12. With the hypothesis of Lemma 11, one also has that $a \circ \phi^G_t \in \mathcal{A}_{s-\sigma}$. To see this, it is enough to follow verbatim the proof of Lemma 11 noting that Lemma 14 below remains valid for $-i\hbar\{a, b\}$ instead of $[a, b]_\hbar$ and then using the formal expansion

$$a \circ \phi^G_t = \sum_{j=0}^{\infty} \frac{t}{j!} \mathrm{Ad}^j_G(a),$$

where $\mathrm{Ad}^j_G(a) = \{G, \mathrm{Ad}^{j-1}_G(a)\}$ and $\mathrm{Ad}^0_G(a) = a$ instead of the analogous quantities for $\Psi^{G,\hbar}_t a$.

A.1. *Preliminary lemmas.* Before proceeding to the proof, we start with some preliminary results.

Lemma 13. *For every $a, b \in \mathcal{A}_s$, the following holds*:

$$\|ab\|_s \le \|a\|_s \|b\|_s.$$

Proof. To see this, write

$$\begin{aligned}\|ab\|_s &= \int_{\mathbb{R}^{2d}} |\widehat{ab}(w)| e^{s|w|}\, dw \\ &= \int_{\mathbb{R}^{2d}} \left| \int_{\mathbb{R}^{2d}} \hat{a}(w - w^*)\hat{b}(w^*)\, dw^* \right| e^{s|w|}\, dw \\ &\le \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} |\hat{a}(w - w^*)| e^{s|w-w^*|} |\hat{b}(w^*)| e^{s|w^*|}\, dw^*\, dw \le \|a\|_s \|b\|_s.\end{aligned}$$ □

We shall also need some estimates on the Moyal product of elements in $\mathcal{A}_s$:

Lemma 14. *Let $a, b \in \mathcal{A}_s$. Then, for every $0 < \sigma_1 + \sigma_2 < s$, we have $[a, b]_\hbar \in \mathcal{A}_{s-\sigma_1-\sigma_2}$ and*

$$\|[a, b]_\hbar\|_{s-\sigma_1-\sigma_2} \le \frac{2\hbar}{e^2 \sigma_1(\sigma_1 + \sigma_2)} \|a\|_s \|b\|_{s-\sigma_2}.$$

Proof. From (37), we have

$$[a, b]_\hbar(z) = 2i \int_{\mathbb{R}^{4d}} \hat{a}(w^*)\hat{b}(z^* - w^*) \sin\!\left(\tfrac{1}{2}\hbar \varsigma(w^*, z^* - w^*)\right) \frac{e^{iz^*\cdot z}}{(2\pi)^{4d}}\, dw^*\, dz^*.$$

Then, using that

$$|\varsigma(w^*, z^* - w^*)| \le 2|w^*|\,|z^* - w^*|, \tag{40}$$

we obtain

$$\begin{aligned}\|[a, b]_\hbar\|_{s-\sigma_1-\sigma_2} &\le \frac{2\hbar}{(2\pi)^{4d}} \int_{\mathbb{R}^{4d}} |\hat{a}(w^*)|\,|w^*|\,|\hat{b}(z^* - w^*)|\,|z^* - w^*| e^{(s-\sigma_1-\sigma_2)(|z^*-w^*|+|w^*|)}\, dw^*\, dz^* \\ &\le \frac{2\hbar}{(2\pi)^{4d}} \Big(\sup_{r\ge 0} r e^{-\sigma_1 r}\Big)\Big(\sup_{r\ge 0} r e^{-(\sigma_1+\sigma_2) r}\Big) \|a\|_s \|b\|_{s-\sigma_2} \\ &\le \frac{2\hbar}{e^2\sigma_1(\sigma_1 + \sigma_2)} \|a\|_s \|b\|_{s-\sigma_2}.\end{aligned}$$ □

Finally, one has:

Lemma 15. *Let $a, b \in \mathcal{A}_s$ and $0 < \sigma < s$. Then there exists a constant $C_\sigma > 0$ depending only on σ such that*

$$\left\| \frac{i}{\hbar}[a, b]_\hbar - \{a, b\} \right\|_{s-\sigma} \le C_\sigma \hbar^2 \|a\|_s \|b\|_{s-\sigma}. \tag{41}$$

Proof. First write

$$\begin{aligned}&[a, b]_\hbar(z) + i\hbar\{a, b\}(z) \\ &\quad = 2i \int_{\mathbb{R}^{4d}} \hat{a}(w^*)\hat{b}(z^* - w^*)\Big(\sin\!\left(\tfrac{1}{2}\hbar\varsigma(w^*, z^* - w^*)\right) - \tfrac{1}{2}\hbar\varsigma(w^*, z^* - w^*)\Big) \frac{e^{iz^*\cdot z}}{(2\pi)^{4d}}\, dw^*\, dz^*.\end{aligned}$$

Using (40) and $\sin(x) = x - \frac{1}{2}x^2 \int_0^1 \sin(tx)(1-t)\,dt$, we obtain

$$\begin{aligned}\|[a,b]_\hbar + i\hbar\{a,b\}\|_{s-\sigma} &\le \frac{\hbar^3}{(2\pi)^{4d}} \int_{\mathbb{R}^{4d}} |\hat{a}(w^*)|\,|w^*|^3\,|\hat{b}(z^*-w^*)|\,|z^*-w^*|^3 e^{(s-\sigma)(|z^*-w^*|+|w^*|)}\,dw^*\,dz^* \\ &\le C_\sigma\,\hbar^3\|a\|_s\|b\|_{s-\sigma}. \qquad \square\end{aligned}$$

A.2. ***Proof of the analytic Egorov lemma.*** We are now in position to prove Lemma 11. Let us start with points (1) and (2). By definition (38), we have

$$\|\Psi_t^{G,\hbar} a - a\|_{s-\sigma} \le \sum_{j=1}^{\infty} \frac{1}{j!}\left(\frac{|t|}{\hbar}\right)^j \|\operatorname{Ad}_G^{\sharp_\hbar, j}(a)\|_{s-\sigma}.$$

Using Lemma 14, we also find that, for every $j \ge 1$,

$$\begin{aligned}\|\operatorname{Ad}_G^{\sharp_\hbar, j}(a)\|_{s-\sigma} &\le \frac{2\hbar j}{e^2\sigma^2}\|\operatorname{Ad}_G^{\sharp_\hbar, j-1}(a)\|_{s-(j-1)\sigma/j}\|G\|_s \\ &\le \frac{2^2\hbar^2 j^3}{e^4\sigma^4(j-1)}\|\operatorname{Ad}_G^{\sharp_\hbar, j-2}(a)\|_{s-(j-2)\sigma/j}\|G\|_s^2 \\ &\le \cdots \le \frac{2^j\hbar^j j^{2j}}{e^{2j}\sigma^{2j} j!}\|a\|_s\|G\|_s^j.\end{aligned}$$

Then, using Stirling's formula and as $2|t|\|G\|_s/\sigma^2 < 1$, one gets

$$\|\Psi_t^{G,\hbar} a - a\|_{s-\sigma} \le \sum_{j=1}^{\infty} \frac{j^{2j}|t|^j\|G\|_s^j}{(j!)^2(e\sigma)^{2j}}\|a\|_s \le C_\sigma|t|\|G\|_s\|a\|_s \tag{42}$$

for some constant $C_\sigma > 0$ depending only on σ. In order to prove point (3), we now write

$$\|\Psi_t^{G,\hbar} a - a + t\{G,a\}\|_{s-\sigma} \le |t|\left\|\frac{i}{\hbar}[G,a]_\hbar - \{G,a\}\right\|_{s-\sigma} + \sum_{j=2}^{\infty}\frac{1}{j!}\left(\frac{|t|}{\hbar}\right)^j\|\operatorname{Ad}_G^{\sharp_\hbar, j}(a)\|_{s-\sigma}.$$

We can now reproduce the above argument and combining this bound with Lemma 15, we can deduce point (3) of Lemma 11.

Acknowledgements

We are grateful to S. Nonnenmacher for suggesting to consider how the methods from [Macià and Rivière 2016; 2019] can be applied in the nonselfadjoint context. We warmly thank J. Sjöstrand and the anonymous referee for their useful comments on the article. Part of this work was achieved when Arnaiz was visiting the Laboratoire Paul Painlevé in Lille. Arnaiz was supported by La Caixa, Severo Ochoa ICMAT, International Ph.D. Programme, 2014, and MTM2017-85934-C3-3-P (MINECO, Spain). Rivière was partially supported by the ANR project GERASIC (ANR-13-BS01-0007-01) and the Labex CEMPI (ANR-11-LABX-0007-01).

References

[Anantharaman 2010] N. Anantharaman, "Spectral deviations for the damped wave equation", *Geom. Funct. Anal.* **20**:3 (2010), 593–626. MR Zbl

[Anantharaman and Léautaud 2014] N. Anantharaman and M. Léautaud, "Sharp polynomial decay rates for the damped wave equation on the torus", *Anal. PDE* **7**:1 (2014), 159–214. MR Zbl

[Arnaiz and Macià 2018] V. Arnaiz and F. Macià, "Concentration of quasimodes for perturbed harmonic oscillators", in preparation, 2018.

[Asch and Lebeau 2003] M. Asch and G. Lebeau, "The spectrum of the damped wave operator for a bounded domain in $\mathbb{R}^2$", *Experiment. Math.* **12**:2 (2003), 227–241. MR Zbl

[Burq and Gérard 2018] N. Burq and P. Gérard, "Stabilisation of wave equations on the torus with rough dampings", preprint, 2018. arXiv

[Burq and Hitrik 2007] N. Burq and M. Hitrik, "Energy decay for damped wave equations on partially rectangular domains", *Math. Res. Lett.* **14**:1 (2007), 35–47. MR Zbl

[Christianson 2007] H. Christianson, "Semiclassical non-concentration near hyperbolic orbits", *J. Funct. Anal.* **246**:2 (2007), 145–195. MR Zbl

[Christianson et al. 2014] H. Christianson, E. Schenck, A. Vasy, and J. Wunsch, "From resolvent estimates to damped waves", *J. Anal. Math.* **122** (2014), 143–162. MR Zbl

[Dimassi and Sjöstrand 1999] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semi-classical limit*, London Mathematical Society Lecture Note Series **268**, Cambridge University Press, 1999. MR Zbl

[Engel and Nagel 2000] K.-J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, Graduate Texts in Mathematics **194**, Springer, 2000. MR Zbl

[Helffer and Sjoestrand 2010] B. Helffer and J. Sjoestrand, "From resolvent bounds to semigroup bounds", preprint, 2010. arXiv

[Hitrik 2002] M. Hitrik, "Eigenfrequencies for damped wave equations on Zoll manifolds", *Asymptot. Anal.* **31**:3-4 (2002), 265–277. MR Zbl

[Hitrik and Sjöstrand 2004] M. Hitrik and J. Sjöstrand, "Non-selfadjoint perturbations of selfadjoint operators in 2 dimensions, I", *Ann. Henri Poincaré* **5**:1 (2004), 1–73. MR Zbl

[Hitrik and Sjöstrand 2005] M. Hitrik and J. Sjöstrand, "Nonselfadjoint perturbations of selfadjoint operators in two dimensions, II: Vanishing averages", *Comm. Partial Differential Equations* **30**:7-9 (2005), 1065–1106. MR Zbl

[Hitrik and Sjöstrand 2008a] M. Hitrik and J. Sjöstrand, "Non-selfadjoint perturbations of selfadjoint operators in two dimensions, IIIa: One branching point", *Canad. J. Math.* **60**:3 (2008), 572–657. MR Zbl

[Hitrik and Sjöstrand 2008b] M. Hitrik and J. Sjöstrand, "Rational invariant tori, phase space tunneling, and spectra for non-selfadjoint operators in dimension 2", *Ann. Sci. Éc. Norm. Supér.* (4) **41**:4 (2008), 511–571. MR Zbl

[Hitrik and Sjöstrand 2012] M. Hitrik and J. Sjöstrand, "Diophantine tori and Weyl laws for non-selfadjoint operators in dimension two", *Comm. Math. Phys.* **314**:2 (2012), 373–417. MR Zbl

[Hitrik and Sjöstrand 2018] M. Hitrik and J. Sjöstrand, "Rational invariant tori and band edge spectra for non-selfadjoint operators", *J. Eur. Math. Soc. (JEMS)* **20**:2 (2018), 391–457. MR Zbl

[Hitrik et al. 2007] M. Hitrik, J. Sjöstrand, and S. V. Ngọc, "Diophantine tori and spectral asymptotics for nonselfadjoint operators", *Amer. J. Math.* **129**:1 (2007), 105–182. MR Zbl

[Jin 2017] L. Jin, "Damped wave equations on compact hyperbolic surfaces", preprint, 2017. arXiv

[Lebeau 1996] G. Lebeau, "Équation des ondes amorties", pp. 73–109 in *Algebraic and geometric methods in mathematical physics* (Kaciveli, 1993), edited by A. Boutet de Monvel and V. Marchenko, Math. Phys. Stud. **19**, Kluwer Acad. Publ., Dordrecht, 1996. MR Zbl

[de la Llave 2001] R. de la Llave, "A tutorial on KAM theory", pp. 175–292 in *Smooth ergodic theory and its applications* (Seattle, WA, 1999), edited by A. Katok et al., Proc. Sympos. Pure Math. **69**, Amer. Math. Soc., Providence, RI, 2001. MR Zbl

[Macià and Rivière 2016] F. Macià and G. Rivière, "Concentration and non-concentration for the Schrödinger evolution on Zoll manifolds", *Comm. Math. Phys.* **345**:3 (2016), 1019–1054. MR Zbl

[Macià and Rivière 2019] F. Macià and G. Rivière, "Observability and quantum limits for the Schrödinger equation on $\mathbb{S}^d$", pp. 139–153 in *Probabilistic methods in geometry, topology and spectral theory*, Contemp. Math. **739**, Amer. Math. Soc., Providence, RI, 2019. MR

[Markus 1988] A. S. Markus, *Introduction to the spectral theory of polynomial operator pencils*, Translations of Mathematical Monographs **71**, Amer. Math. Soc., Providence, RI, 1988. MR Zbl

[Markus and Macaev 1979] A. S. Markus and V. I. Macaev, "Asymptotic behavior of the spectrum of close-to-normal operators", *Funktsional. Anal. i Prilozhen.* **13**:3 (1979), 93–94. In Russian; translated in *Funct. Anal. Appl.* **13**:3 (1979), 233–235. MR Zbl

[Nonnenmacher 2011] S. Nonnenmacher, "Spectral theory of damped quantum chaotic systems", preprint, 2011. arXiv

[Rauch and Taylor 1975] J. Rauch and M. Taylor, "Decay of solutions to nondissipative hyperbolic systems on compact manifolds", *Comm. Pure Appl. Math.* **28**:4 (1975), 501–523. MR Zbl

[Rivière 2014] G. Rivière, "Eigenmodes of the damped wave equation and small hyperbolic subsets", *Ann. Inst. Fourier* (*Grenoble*) **64**:3 (2014), 1229–1267. MR Zbl

[Schenck 2010] E. Schenck, "Energy decay for the damped wave equation under a pressure condition", *Comm. Math. Phys.* **300**:2 (2010), 375–410. MR Zbl

[Sjöstrand 2000] J. Sjöstrand, "Asymptotic distribution of eigenfrequencies for damped wave equations", *Publ. Res. Inst. Math. Sci.* **36**:5 (2000), 573–611. MR Zbl

[Zworski 2012] M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics **138**, Amer. Math. Soc., 2012. MR Zbl

Received 27 Jul 2019. Revised 21 Oct 2019. Accepted 19 Nov 2019.

VÍCTOR ARNAIZ: victor.arnaiz@math.u-psud.fr, victor.arnaiz@icmat.es
Université Paris-Saclay, CNRS, Laboratoire de mathématiques d'Orsay, Orsay, France

and

Instituto de Ciencias Matemáticas (ICMAT), Campus de Cantoblanco, UAM, Madrid, Spain

GABRIEL RIVIÈRE: gabriel.riviere@univ-nantes.fr
Laboratoire de mathématiques Jean Leray (UMR CNRS 6629), Université de Nantes, Nantes, France

PURE and APPLIED
ANALYSIS

Vol. 2, No. 2, 2020
https://doi.org/10.2140/paa.2020.2.447

ON THE SOLUTION OF LAPLACE'S EQUATION IN THE VICINITY OF TRIPLE JUNCTIONS

JEREMY HOSKINS AND MANAS RACHH

We characterize the behavior of solutions to systems of boundary integral equations associated with Laplace transmission problems in composite media consisting of regions with polygonal boundaries. In particular we consider triple junctions, i.e., points at which three distinct media meet. We show that, under suitable conditions, solutions to the boundary integral equations in the vicinity of a triple junction are well-approximated by linear combinations of functions of the form t^β, where t is the distance of the point from the junction and the powers β depend only on the material properties of the media and the angles at which their boundaries meet. Moreover, we use this analysis to design efficient discretizations of boundary integral equations for Laplace transmission problems in regions with triple junctions and demonstrate the accuracy and efficiency of this algorithm with a number of examples.

1. Introduction

Composite media, i.e., media consisting of multiple materials in close proximity or contact, are both ubiquitous in nature and fascinating in applications since their macroscopic properties can be substantially different than those of their components. One property of particular interest is the electrostatic response of composite media, typically the electric potential in the medium which is produced by an externally applied time-independent electric field. In such situations one often assumes that the associated electric potential satisfies Laplace's equation in the interior of each medium and that along each edge where two media meet one prescribes the jump in the normal derivative of the potential. Typically the potentials in these jump relations appear multiplied by coefficients depending on the electric permittivity. This leads to a collection of coupled partial differential equations (PDEs). In addition to classical electrostatics problems, the same equations also arise in, among other things, percolation theory, homogenization theory, and the study of field enhancements in vacuum insulators; see, for example, [Lee 2008; Fredkin and Mayergoyz 2003; Milton 2002; Tully et al. 2007; Tuncer et al. 2002; Fel et al. 2000; Ovchinnikov 2004].

Using classical potential theory this set of partial differential equations (PDEs) can be reduced to a system of second-kind boundary integral equations (BIEs). In particular, the solution to the PDE in each region is represented as a linear combination of a single-layer and a double-layer potential on the boundary of each subregion. If the edges of the media are smooth then the corresponding kernels in the integral equation are as well. Near corners, however, the solutions to both the differential equations and the integral equations can develop singularities.

Analytically, the behavior of solutions to both the PDEs and BIEs has been the subject of extensive analysis; see, for example [Craster and Obnosov 2004; Keller 1987; Helsing 1991; Chung et al. 2005;

MSC2010: 31A10, 35Q60, 45L05, 65E05, 65R20.
Keywords: boundary integral equations, multiple junction interfaces, corners, singular solutions, potential theory.

Schächter 1998; Berggren et al. 2001; Techaumnat et al. 2002; Afanas'ev et al. 2004; Greengard and Lee 2012; Claeys et al. 2015]. In particular, the existence and uniqueness of solutions in an $\mathbb{L}^2$-sense is well known, under certain natural assumptions on the material properties [Claeys et al. 2017; McLean 2000]. Moreover, the asymptotic form of the singularities in the vicinity of a junction has been determined for the solutions of both the PDE and its corresponding BIE [Chung et al. 2005; Schächter 1998; Craster and Obnosov 2004; Milton et al. 1981; Helsing 2011].

Computationally the singular nature of the solutions poses significant challenges for many existing numerical methods for solving both the PDEs and BIEs. Typical approaches involve introducing many additional degrees of freedom near the junctions which can impede the speed of the solver and impose prohibitive limits on the size and complexity of geometries which can be considered. *Recursive compressed inverse preconditioning* (*RCIP*) is one way of circumventing the difficulty introduced by the presence of junctions in the BIE formulation [Helsing 2013]. In this approach, the extra degrees of freedom introduced by the refinement near the junctions are eliminated from the linear system. Moreover, the compression and refinement are performed concomitantly for multiple junctions in parallel. This approach gives an algorithm which scales linearly in the number of degrees of freedom added to resolve the singularities near the junction. The resulting linear system has essentially the same number of degrees of freedom as it would if the junctions were absent.

In this paper we restrict our attention to the case of triple junctions, extending the existing analysis by showing that under suitable restrictions the solution to the BIEs can be well-approximated in the vicinity of a triple junction by a linear combination of t^{β_j}, where t is the distance from the triple junction and the β_j's are a countable collection of real numbers defined implicitly by an equation depending only on the angles at which the interfaces meet and the material properties of the corresponding media. This analysis enables the construction of an efficient computational algorithm for solving Laplace's equation in regions with multiple junctions. In particular, using this representation we construct an accurate and efficient quadrature scheme for the BIE which requires no refinement near the junction. The properties of this discretization are illustrated with a number of numerical examples.

This paper is organized as follows. In Section 2 we state the boundary value problem for the Laplace triple junction transmission problem, summarize relevant properties of layer potentials, and describe the reduction of the boundary value problem to a system of boundary integral equations. In Section 3 we present the main theoretical results of this work, the proofs of which are given in Appendices A and B. In Section 4 we discuss two conjectures extending the results of Section 3 based on extensive numerical evidence. In Section 5, we describe a Nyström discretization which exploits explicit knowledge of the structure of solutions to the integral equations in the vicinity of triple junctions, and in Section 6 we demonstrate its effectiveness of numerical solvers. Finally, in Section 7 we summarize the results and outline directions for future research.

2. Boundary value problem

Consider a composite medium consisting of a set of n polygonal domains $\Omega_1, \dots, \Omega_n$ (see Figure 1) with boundaries consisting of m edges $\Gamma_1, \dots, \Gamma_m$ and k vertices $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$. For a given edge Γ_i let L_i

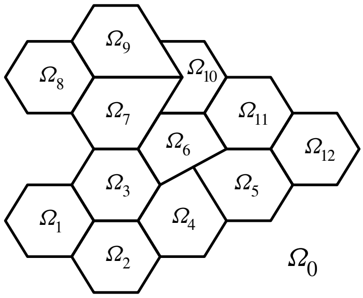

Figure 1. Example of a composite region.

denote its length, $\boldsymbol{n}_i$ its normal, $\ell(i), r(i)$ the polygons to the left and right, respectively, and let γ_i be an arc length parametrization of Γ_i. Finally we denote the union of the regions $\Omega_1, \ldots, \Omega_n$ by Ω and denote the complement of Ω by Ω_0.

Given positive constants $\mu_1, \ldots, \mu_n$ and $\nu_1, \ldots, \nu_n$ we consider the boundary value problem

$$\begin{aligned} \Delta u_i &= 0 \qquad x \in \Omega_i,\ i = 0, 1, 2, \ldots, n, \\ \mu_{\ell(i)} u_{\ell(i)} - \mu_{r(i)} u_{r(i)} &= f_i, \quad x \in \Gamma_i,\ i = 1, \ldots, m, \\ \nu_{\ell(i)} \frac{\partial u_{\ell(i)}}{\partial n_i} - \nu_{r(i)} \frac{\partial u_{r(i)}}{\partial n_i} &= g_i, \quad x \in \Gamma_i,\ i = 1, \ldots, m, \\ \lim_{|r| \to \infty} (r \log(r) u_0'(r) - u_0(r)) &= 0, \end{aligned} \tag{1}$$

where f_i and g_i are analytic functions on Γ_i, $i = 1, \ldots, m$, and $\ell(i), r(i)$ denote the regions on the left and right with respect to the normal of edge Γ_i.

Remark 2.1. In this work we assume that all the normals $\boldsymbol{n}_1, \ldots, \boldsymbol{n}_m$ to $\Gamma_1, \ldots, \Gamma_m$ are positively oriented with respect to the parametrization $\gamma_i(t)$ of the edge Γ_i. Specifically, if Γ_i is a line segment between vertices $\boldsymbol{v}_\ell$, $\boldsymbol{v}_r$, and $\gamma_i(t) : [0, L_i] \to \Gamma_i$ is a parametrization of Γ_i, given by

$$\gamma_i(t) = \boldsymbol{v}_\ell + t \frac{\boldsymbol{v}_r - \boldsymbol{v}_\ell}{\|\boldsymbol{v}_r - \boldsymbol{v}_\ell\|}, \tag{2}$$

then the normal on edge Γ_i is given by

$$\boldsymbol{n}_i = \frac{(\boldsymbol{v}_r - \boldsymbol{v}_\ell)^\perp}{\|\boldsymbol{v}_r - \boldsymbol{v}_\ell\|}, \tag{3}$$

where for a point $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$, we have $\boldsymbol{x}^\perp = (x_2, -x_1)$.

Remark 2.2. The existence and uniqueness of solutions to (1) is a classical result [McLean 2000].

Remark 2.3. In this paper we assume that no more than three edges meet at each vertex. Similar analysis holds for domains with higher-order junctions and will be published at a later date.

Remark 2.4. Here we assume that $\mu_1, \ldots, \mu_n$, and $\nu_1, \ldots, \nu_n$ are positive constants. In principle the analysis presented here extends to the case where the constants are negative or complex provided the constants $(\mu_j \nu_i + \mu_i \nu_j)/(\mu_j \nu_i - \mu_i \nu_j)$ across each edge are outside the closure of the essential spectrum of the double-layer potential defined on the boundary, and the underlying differential equation admits a unique solution. Note that for nonnegative coefficients this is always true, since these constants are all of magnitude greater than 1, and the spectral radius of the double-layer potential is bounded by 1.

2A. *Layer potentials.* Before reducing the boundary value problem (1) to a boundary integral equation we first introduce the layer potential operators and summarize their relevant properties.

Definition 2.5. Given a density σ defined on Γ_i, $i = 1, \dots, m$, the single-layer potential is defined by

$$\mathcal{S}_{\Gamma_i}[\sigma](\boldsymbol{y}) = -\frac{1}{2\pi} \int_{\Gamma_i} \log \|\boldsymbol{x} - \boldsymbol{y}\| \sigma(\boldsymbol{x})\, dS_{\boldsymbol{x}}, \tag{4}$$

and the double-layer potential is defined via the formula

$$\mathcal{D}_{\Gamma_i}[\sigma](\boldsymbol{y}) = \frac{1}{2\pi} \int_{\Gamma_i} \frac{\boldsymbol{n}(\boldsymbol{x}) \cdot (\boldsymbol{y} - \boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{y}\|^2} \sigma(\boldsymbol{x})\, dS_{\boldsymbol{x}}. \tag{5}$$

Remark 2.6. In light of the previous definition, evidently the adjoint of the double-layer potential is given by the formula

$$\mathcal{D}^*_{\Gamma_i}[\sigma](\boldsymbol{y}) = \frac{1}{2\pi} \int_{\Gamma_i} \frac{\boldsymbol{n}(\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y})}{\|\boldsymbol{x} - \boldsymbol{y}\|^2} \sigma(\boldsymbol{x})\, dS_{\boldsymbol{x}}. \tag{6}$$

Definition 2.7. For $\boldsymbol{x} \in \Gamma$ we define the kernel $K(\boldsymbol{x}, \boldsymbol{y})$ by

$$K(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2\pi} \frac{\boldsymbol{n}(\boldsymbol{x}) \cdot (\boldsymbol{y} - \boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{y}\|^2}. \tag{7}$$

The following theorems describe the limiting values of the single- and double-layer potential on the boundary Γ_i.

Theorem 2.8. *Suppose that $\boldsymbol{x}_0$ is a point in the interior of the segment Γ_i. Suppose the point $\boldsymbol{x}$ approaches a point $\boldsymbol{x}_0$ along a path such that*

$$-1 + \alpha < \frac{\boldsymbol{x} - \boldsymbol{x}_0}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \cdot \gamma_i'(t_0) < 1 - \alpha \tag{8}$$

for some $\alpha > 0$. If $(\boldsymbol{x} - \boldsymbol{x}_0) \cdot \boldsymbol{n}_i < 0$, we will refer to this limit as $\boldsymbol{x} \to \boldsymbol{x}_0^-$, and if $(\boldsymbol{x} - \boldsymbol{x}_0) \cdot \boldsymbol{n}_i > 0$, we will refer to this limit as $\boldsymbol{x} \to \boldsymbol{x}_0^+$.

Then

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0^\pm} \mathcal{S}_{\Gamma_i}[\sigma](\boldsymbol{x}) = \mathcal{S}_{\Gamma_i}[\sigma](\boldsymbol{x}_0), \tag{9}$$

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0^\pm} \mathcal{D}_{\Gamma_i}[\rho](\boldsymbol{x}) = \text{p.v.}\, \mathcal{D}_{\Gamma_i}[\rho](\boldsymbol{x}_0) \mp \frac{\rho(\boldsymbol{x}_0)}{2}, \tag{10}$$

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0^\pm} \boldsymbol{n}_i \cdot \nabla \mathcal{S}_{\Gamma_i}[\rho](\boldsymbol{x}) = \text{p.v.}\, \mathcal{D}^*_{\Gamma_i}[\rho](\boldsymbol{x}_0) \pm \frac{\rho(\boldsymbol{x}_0)}{2}, \tag{11}$$

where p.v. refers to the fact that the principal value of the integral should be taken.

Moreover, both the limits

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0^{\pm}} \boldsymbol{n}_i \cdot \nabla \mathcal{D}_{\Gamma_i}[\rho](\boldsymbol{x}) \tag{12}$$

exist and are equal.

Remark 2.9. In the following we will suppress the p.v. from expressions involving layer potentials evaluated at a point on the boundary. Unless otherwise stated, in such cases the principal value should always be taken.

2B. *Integral representation.* In classical potential theory the boundary value problem (1) is reduced to a boundary integral equation for a new collection of unknowns $\rho_i, \sigma_i \in \mathbb{L}^2(\Gamma_i)$, $i = 1, \dots, m$, related to $u_i : \Omega_i \to \mathbb{R}$, $i = 1, \dots, n$, in the following manner:

$$u_i(\boldsymbol{x}) = \frac{1}{\mu_i} \sum_{j=1}^{m} \mathcal{S}_{\Gamma_j}[\rho_j](\boldsymbol{x}) + \frac{1}{\nu_i} \sum_{j=1}^{m} \mathcal{D}_{\Gamma_j}[\sigma_j](\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega_i. \tag{13}$$

We note that by construction u_i is harmonic in Ω_i, $i = 0, 1, \dots, n$. Enforcing the jump conditions across the edges and applying Theorem 2.8 yields the following system of integral equations for the unknown densities ρ_i and σ_i for $i = 1, \dots, m$:

$$-\tfrac{1}{2}\sigma_i + \frac{\mu_{r(i)}\nu_{\ell(i)} - \mu_{\ell(i)}\nu_{r(i)}}{\mu_{r(i)}\nu_{\ell(i)} + \mu_{\ell(i)}\nu_{r(i)}} \sum_{\ell=1}^{m} \mathcal{D}_{\Gamma_\ell}[\sigma_\ell] = \frac{\nu_{\ell(i)}\nu_{r(i)} f_i}{\mu_{r(i)}\nu_{\ell(i)} + \mu_{\ell(i)}\nu_{r(i)}}, \tag{14}$$

$$-\tfrac{1}{2}\rho_i + \frac{\mu_{r(i)}\nu_{\ell(i)} - \mu_{\ell(i)}\nu_{r(i)}}{\mu_{r(i)}\nu_{\ell(i)} + \mu_{\ell(i)}\nu_{r(i)}} \sum_{\ell=1}^{m} \mathcal{D}^*_{\Gamma_\ell}[\rho_\ell] = -\frac{\mu_{\ell(i)}\mu_{r(i)} g_i}{\mu_{r(i)}\nu_{\ell(i)} + \mu_{\ell(i)}\nu_{r(i)}}. \tag{15}$$

We note that the preceding representation has several advantages. Firstly, the kernels of integral equations (14) and (15) are smooth except at the vertices. In particular, the weakly singular terms arising from the single-layer potential and the hypersingular terms arising from the derivative of the double-layer potential are absent. Secondly, the equations for the single-layer density ρ and the double-layer density σ are completely decoupled and can be analyzed separately. Moreover, (15) is the adjoint of (14) and hence the structure of solutions to (15) can be inferred from the behavior of solutions to (14).

Remark 2.10. The above representation also appears in [Helsing 2011] and is related to the work in [Greengard and Lee 2012]. It has been shown in [Claeys et al. 2017] that the boundary integral equations (14) and (15) are well-posed for $f_i, g_i \in \mathbb{L}^2[\Gamma_i]$.

2C. *The single-vertex problem.* The following lemma reduces the problem of analyzing the behavior of the densities ρ and σ in the vicinity of a triple junction with locally analytic data to the analysis of an integral equation on a set of three intersecting line segments.

Lemma 2.11. *Let σ, ρ satisfy the boundary integral equation* (14) *and* (15), *respectively. Consider three edges Γ_i, Γ_j, and Γ_k meeting at a vertex v_p. If $\boldsymbol{x}_p$ denotes the coordinates of the vertex v_m then there*

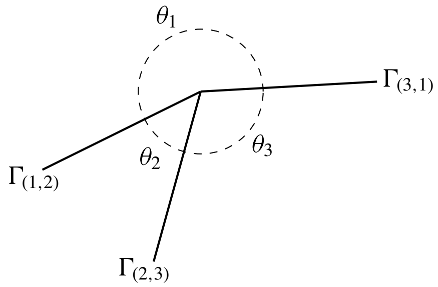

Figure 2. Geometry near a triple junction.

exists an $r > 0$ *such that*

$$\int_{\Gamma\setminus B_r(\boldsymbol{x}_p)} K(\boldsymbol{x},\boldsymbol{y})\sigma(\boldsymbol{x})\,dS_{\boldsymbol{x}} \quad \textit{and} \quad \int_{\Gamma\setminus B_r(\boldsymbol{x}_p)} K(\boldsymbol{y},\boldsymbol{x})\rho(\boldsymbol{x})\,dS_{\boldsymbol{x}} \tag{16}$$

are analytic functions of $\boldsymbol{y}$ *for all* $\boldsymbol{y} \in B_r(\boldsymbol{x}_p)$. *Here* $B_r(\boldsymbol{x}_p)$ *denotes the ball of radius* r *centered at* $\boldsymbol{x}_p$.

Remark 2.12. We note that by choosing r sufficiently small we can assume that the intersection of all three-edges with $B_r(\boldsymbol{x}_p)$ are of length r. Moreover, since Laplace's equation is invariant under scalings, the subproblem associated with the corner can be mapped to an integral equation on three intersecting edges of unit length.

In light of the preceding remark, in the remainder of this paper we restrict our attention to the geometry shown in Figure 2.

The following notation will be used in our analysis of triple junctions.

Remark 2.13. Suppose that $\Gamma_{(\ell,m)}$ and $\Gamma_{(\ell',m')}$ are two (possibly identical) edges of a triple junction in which all edges are of length 1. For (ℓ, m) and (ℓ', m') in $\{(1, 2), (2, 3), (3, 1)\}$ and $t \in (0, 1)$ let

$$\mathcal{D}_{(\ell,m);(\ell',m')}[\sigma](t) = \text{p.v.}\,\mathcal{D}_{\Gamma_{(\ell,m)}}[\sigma]\big|_{\Gamma_{(\ell',m')}}, \tag{17}$$

$$\mathcal{D}^*_{(\ell,m);(\ell',m')}[\rho](t) = \text{p.v.}\,\mathcal{D}^*_{\Gamma_{(\ell,m)}}[\rho]\big|_{\Gamma_{(\ell',m')}} \tag{18}$$

for any $\sigma, \rho \in \mathbb{L}^2(\Gamma_{(3,1)} \cup \Gamma_{(1,2)} \cup \Gamma_{(2,3)})$. Note that if $(\ell, m) = (\ell', m')$ then both quantities are identically zero for any σ and ρ. If $(\ell, m) \neq (\ell', m')$ then the principal value is not required.

Finally, in the following we will also denote the restrictions of σ and ρ to an edge $\Gamma_{(\ell,m)}$ by $\sigma_{(\ell,m)}$ and $\rho_{(\ell,m)}$, respectively.

3. Main results

In this section we state several theorems which characterize the behavior of the solutions σ, ρ to (14) and (15) for the single-vertex problem with piecewise smooth boundary data f and g. Before doing so we first introduce some convenient notation. To that end, let $\Gamma_{(1,2)}$, $\Gamma_{(2,3)}$ and $\Gamma_{(3,1)}$ be three edges of unit length meeting at a vertex as in Figure 2. Let θ_1, θ_2, and θ_3 be the angles at which they meet and suppose that $0 < \theta_1, \theta_2, \theta_3 < 2\pi$ are real numbers summing to 2π. Let Ω_1 denote the region bordered by

$\Gamma_{(3,1)}$ and $\Gamma_{(1,2)}$, Ω_2 denote the region bordered by $\Gamma_{(1,2)}$ and $\Gamma_{(2,3)}$, and Ω_3 denote the region bordered by $\Gamma_{(2,3)}$ and $\Gamma_{(3,1)}$. Finally, let μ_i and ν_i be the parameters corresponding to Ω_i, $i = 1, 2, 3$, and define the constants $d_{(1,2)}$, $d_{(2,3)}$ and $d_{(3,1)}$ by

$$d_{(1,2)} = \frac{\mu_1\nu_2 - \mu_2\nu_1}{\mu_1\nu_2 + \mu_2\nu_1}, \quad d_{(2,3)} = \frac{\mu_2\nu_3 - \mu_3\nu_2}{\mu_2\nu_3 + \mu_3\nu_2}, \quad d_{(3,1)} = \frac{\mu_3\nu_1 - \mu_1\nu_3}{\mu_3\nu_1 + \mu_1\nu_3}. \tag{19}$$

Remark 3.1. We note the following properties of $d_{(3,1)}$, $d_{(1,2)}$, $d_{(2,3)}$ which, for notational convenience, we will denote by a, b, and c, respectively. Firstly, since μ_i, ν_i are positive real numbers, it follows that $a, b, c \in (-1, 1)$. Secondly, a simple calculation shows that $c = -(a+b)/(1+ab)$. Thus, at each triple junction, there are two parameters (a, b) which encapsulate the relevant information regarding material properties at that junction. For the rest of the paper, in a slight abuse of notation, we will refer to (a, b) as the material parameters.

Next we define several quantities which will be used in the statement of the main results. Let $\mathcal{J}$ denote the set of indices $\{(1, 2), (2, 3), (3, 1)\}$ and $X = \mathbb{L}^2(\Gamma_{(1,2)}) \otimes \mathbb{L}^2(\Gamma_{(2,3)}) \otimes \mathbb{L}^2(\Gamma_{(3,1)})$. Let $\mathcal{K}_{\text{dir}} : X \to X$ and $\mathcal{K}_{\text{neu}} : X \to X$ denote the bounded operators in (14) and (15) respectively. For any operator $A : X \to X$, $\boldsymbol{h} \in X$, and $(i, j) \in \mathcal{J}$, we denote the restriction of $A[\boldsymbol{h}]$ to the edge $\Gamma_{(i,j)}$ by $A[\boldsymbol{h}]_{(i,j)}$. For example, given $\boldsymbol{h}(t) = [h_{(1,2)}(t), h_{(2,3)}(t), h_{(3,1)}(t)]^T \in X$, and $(i, j) \in \mathcal{J}$,

$$\mathcal{K}_{\text{dir}}[\boldsymbol{h}]_{(i,j)} = -\tfrac{1}{2}h_{(i,j)} + d_{(i,j)} \sum_{(\ell,m)\in\mathcal{J}} \mathcal{D}_{(\ell,m);(i,j)}[h_{(\ell,m)}], \tag{20}$$

where the operators $\mathcal{D}_{(\ell,m);(i,j)}$ are defined in (17).

We are interested in the following two problems:

(1) For what collection of $\boldsymbol{h} \in X$ are $\mathcal{K}_{\text{dir}}[\boldsymbol{h}]$ and $\mathcal{K}_{\text{neu}}[\boldsymbol{h}]$ piecewise smooth functions on each of the edges $\Gamma_{(i,j)}$, $(i, j) \in \mathcal{J}$?

(2) Given $h_{(i,j)} \in \mathcal{P}_N$, a polynomial of degree at most N, construct an explicit basis for $\mathcal{K}_{\text{dir}}^{-1}[\boldsymbol{h}]$ and $\mathcal{K}_{\text{neu}}^{-1}[\boldsymbol{h}]$.

In Section 3A, we address these questions for $\mathcal{K}_{\text{dir}}$, while in Section 3B we present analogous results for $\mathcal{K}_{\text{neu}}$.

3A. *Analysis of $\mathcal{K}_{\text{dir}}$*. Suppose that $\boldsymbol{h}(t) = [h_{(1,2)}(t), h_{(2,3)}(t), h_{(3,1)}(t)]^T = \boldsymbol{v}t^\beta$, where t denotes the distance along the edge $\Gamma_{(i,j)}$ from the triple junction, and $\boldsymbol{v} \in \mathbb{R}^3$ and $\beta \in \mathbb{R}$ are constants.

In the following theorem, we derive necessary conditions on β, $\boldsymbol{v}$ such that $\mathcal{K}_{\text{dir}}[\boldsymbol{h}]_{(i,j)}$ is a smooth function on each edge $\Gamma_{(i,j)}$, $(i, j) \in \mathcal{J}$.

Theorem 3.2. *Let $\mathcal{A}_{\text{dir}}(a, b, \beta) \in \mathbb{R}^{3\times 3}$ denote the matrix given by*

$$\mathcal{A}_{\text{dir}}(a, b, \beta) = \begin{pmatrix} \sin(\pi\beta) & b\sin\beta(\pi-\theta_2) & -b\sin\beta(\pi-\theta_1) \\ (a+b)/(1+ab)\sin\beta(\pi-\theta_2) & \sin(\pi\beta) & -(a+b)/(1+ab)\sin\beta(\pi-\theta_3) \\ a\sin\pi\beta(1-\theta_1) & -a\sin\pi\beta(1-\theta_3) & \sin(\pi\beta) \end{pmatrix}. \tag{21}$$

Suppose that β *is a positive real number such that* $\det \mathcal{A}_{\text{dir}}(d_{(3,1)}, d_{(1,2)}, \beta) = 0$ *and that* $\boldsymbol{v}$ *is a null vector of* $\mathcal{A}_{\text{dir}}(d_{(3,1)}, d_{(1,2)}, \beta)$. *Let* $\boldsymbol{h}(t) = \boldsymbol{v}t^{\beta}$, $0 < t < 1$. *Then* $\mathcal{K}_{\text{dir}}[\boldsymbol{h}]_{(i,j)}$ *is an analytic function of* t, *for* $0 < t < 1$, *on each of the edges* $\Gamma_{(i,j)}$, $(i, j) \in \mathcal{J}$.

The above theorem guarantees that for appropriately chosen densities $\boldsymbol{h} \in X$, the potential $\mathcal{K}_{\text{dir}}[\boldsymbol{h}]$ is an analytic function on each of the edges.

We now consider the construction of a basis for $\mathcal{K}_{\text{dir}}^{-1}[\boldsymbol{h}]$, when $h_{(i,j)} \in \mathcal{P}_N$, $(i, j) \in \mathcal{J}$, for some $N > 0$.

In order to prove this result, we require a collection of β, $\boldsymbol{v}$ satisfying the conditions of Theorem 3.2. The following lemma states the existence of a countable collection of β, $\boldsymbol{v}$ which are analytic on a subset of $(-1, 1)^2$.

Lemma 3.3. *Suppose that* $\theta_1, \theta_2, \theta_3$ *are irrational numbers summing to* 2π, *and* $(a, b) \in (-1, 1)^2$. *Then there exists a countable collection of open subsets of* $(-1, 1)^2$, *denoted by* $S_{i,j}$, *as well as a corresponding set of functions* $\beta_{i,j} : S_{i,j} \to \mathbb{R}$, $i = 0, 1, 2, \ldots,$ $j = 0, 1, 2$, *such that* $\det \mathcal{A}_{\text{dir}}(a, b, \beta_{i,j}) = 0$ *for all* $(a, b) \in S_{i,j}$. *The corresponding null vectors* $\boldsymbol{v}_{i,j} : S_{i,j} \to \mathbb{R}^3$ *of* $\mathcal{A}_{\text{dir}}(a, b, \beta_{i,j})$ *are also analytic functions. Finally, for any* $N > 0$, *we have* $\left|\bigcap_{i=0}^{N} \bigcap_{j=0}^{2} S_{i,j}\right| > 0$.

In the following theorem, we present the main result of this section, which gives a basis for $\mathcal{K}_{\text{dir}}^{-1}[\boldsymbol{h}]$.

Theorem 3.4. *Consider the same geometry as in Figure 2, where* θ_1, θ_2, *and* θ_3 *sum to* 2π *and* θ_1/π, θ_2/π, *and* θ_3/π *are irrational. Let* $\beta_{i,j}, \boldsymbol{v}_{i,j}, S_{i,j}$, $i = 0, 1, 2, \ldots,$ $j = 0, 1, 2$, *be as defined in Lemma 3.3, and for any positive integer* N, *let* S_N *denote the region of common analyticity of* $\beta_{i,j}, \boldsymbol{v}_{i,j}$, *i.e.,* $S_N = \bigcap_{i=0}^{N} \bigcap_{j=0}^{2} S_{i,j}$. *Finally, suppose that* $h^k_{(i,j)}$, $(i, j) \in \mathcal{J}$, $k = 0, 1, 2, \ldots N$, *are real constants, and define* $h_{(i,j)}$ *by*

$$h_{(i,j)}(t) = \sum_{k=0}^{N} h^k_{(i,j)} t^k, \tag{22}$$

$0 < t < 1$. *Then there exists an open region* $\widetilde{S}_N \subset S_N \subset (-1, 1)^2$ *with* $|\widetilde{S}_N| > 0$ *such that the following holds. For all* $(a, b) \in \widetilde{S}_N$, *there exist constants* $p_{i,j}$, $i = 0, 1, \ldots N$, $j = 0, 1, 2$, *such that*

$$\sigma = \begin{bmatrix} \sigma_{1,2}(t) \\ \sigma_{2,3}(t) \\ \sigma_{3,1}(t) \end{bmatrix} = \sum_{i=0}^{N} \sum_{j=0}^{2} p_{i,j} \boldsymbol{v}_{i,j} t^{\beta_{i,j}} \tag{23}$$

satisfies

$$\max_{(i,j)\in\mathcal{J}} |h_{(i,j)} - \mathcal{K}_{\text{dir}}[\sigma]_{(i,j)}| \le C t^{N+1} \tag{24}$$

for $0 < t < 1$, *where* C *is a constant.*

3B. *Analysis of* $\boldsymbol{\mathcal{K}_{\text{neu}}}$**.** Suppose that $\boldsymbol{h}(t) = [h_{(1,2)}(t), h_{(2,3)}(t), h_{(3,1)}(t)]^T = \boldsymbol{w}t^{\beta-1}$, where t denotes the distance on the edge $\Gamma_{(i,j)}$ from the triple junction, and $\boldsymbol{w} \in \mathbb{R}^3$ and β are constants. In the following theorem, we discuss necessary conditions on β, $\boldsymbol{w}$ guaranteeing that $\mathcal{K}_{\text{neu}}[\boldsymbol{h}]_{(i,j)}$ is a smooth function on each edge $\Gamma_{(i,j)}$, $(i, j) \in \mathcal{J}$.

Theorem 3.5. *Let $\mathcal{A}_{\text{neu}}(a, b, \beta) \in \mathbb{R}^{3\times 3}$ denote the matrix given by*

$$\mathcal{A}_{\text{neu}}(a,b,\beta) = \begin{pmatrix} \sin(\pi\beta) & -b\sin\beta(\pi-\theta_2) & b\sin\beta(\pi-\theta_1) \\ -(a+b)/(1+ab)\sin\beta(\pi-\theta_2) & \sin(\pi\beta) & (a+b)/(1+ab)\sin\beta(\pi-\theta_3) \\ -a\sin\beta(\pi-\theta_1) & a\sin\pi\beta(1-\theta_3) & \sin(\pi\beta) \end{pmatrix}. \quad (25)$$

Suppose that β is a positive real number such that $\det \mathcal{A}_{\text{neu}}(d_{(3,1)}, d_{(1,2)}, \beta) = 0$ and let $\boldsymbol{w}$ denote a corresponding null vector of $\mathcal{A}_{\text{neu}}(d_{(3,1)}, d_{(1,2)}, \beta)$. Let $h = \boldsymbol{w}t^{\beta-1}$, $0 < t < 1$. Then $\mathcal{K}_{\text{neu}}[h]_{(i,j)}$ is an analytic function of t, for $0 < t < 1$, on each of the edges $\Gamma_{(i,j)}$, $(i, j) \in \mathcal{J}$.

Before proceeding a few remarks are in order.

Remark 3.6. We note that $\det \mathcal{A}_{\text{dir}}(a, b, \beta) = \det \mathcal{A}_{\text{neu}}(a, b, \beta)$. Thus, the existence of β, $\boldsymbol{w}$ which satisfy the conditions of Theorem 3.5 is guaranteed by Lemma 3.3.

Remark 3.7. For a given β, if there exists a $\boldsymbol{v} \in \mathbb{R}^3$ such that $\mathcal{K}_{\text{dir}}[\boldsymbol{v}t^{\beta}]$ is piecewise smooth then there also exists a vector $\boldsymbol{w} \in \mathbb{R}^3$ such that $\mathcal{K}_{\text{neu}}[\boldsymbol{w}t^{\beta-1}]$ is also a smooth function. However, the requirement that $\boldsymbol{w}t^{\beta-1} \in X$ implies that, for $\mathcal{K}_{\text{neu}}$, only β's which satisfy $\beta > \frac{1}{2}$ are admissible.

For $\mathcal{K}_{\text{dir}}$, note that $\beta_{0,j} = 0$ for $j = 0, 1, 2$ (see the proof of Lemma 3.3 contained in Appendix A.1). These densities are essential for the proof of Theorem 3.4, since these are the only basis functions for which the projection of their image under $\mathcal{K}_{\text{dir}}$ onto the constant functions are nonzero.

However, since $\beta_{0,j} \not> \frac{1}{2}$, the densities $\boldsymbol{w}_{0,j}t^{\beta_{0,j}-1}$ are excluded from the representation for the solution to the equation $\mathcal{K}_{\text{neu}}[\sigma] = \boldsymbol{h}$. Note that, unlike $\mathcal{K}_{\text{dir}}[\boldsymbol{v}_{i,j}t^{\beta_{i,j}}]$, $\mathcal{K}_{\text{neu}}[\boldsymbol{w}_{i,j}t^{\beta_{i,j}-1}]$, $i = 1, 2, \dots,$ $j = 0, 1, 2$, have a nonzero projection onto the constants (see Lemma B.2).

The following theorem is a converse of Theorem 3.5 under suitable restrictions.

Theorem 3.8. *Consider the same geometry as in Figure 2, where θ_1, θ_2, and θ_3 are irrational numbers summing to 2. Let $\beta_{i,j}$, $\boldsymbol{w}_{i,j}$, $S_{i,j}$, $i = 0, 1, 2, \dots,$ $j = 0, 1, 2$, be as defined in Lemma 3.3. Let $T_{i,j}$ denote the open subset of $(-1, 1)^2$ on which $\beta_{i,j}$ and $\boldsymbol{w}_{i,j}$ are analytic and $\beta_{i,j} > \frac{1}{2}$. For any positive integer N, let S_N^{neu} denote the region of common analyticity of $\beta_{i,j}$, $\boldsymbol{w}_{i,j}$; i.e., $S_N^{\text{neu}} = \bigcap_{i=1}^{N+1}\bigcap_{j=0}^{2} T_{i,j}$. Finally, suppose that $h^k_{(i,j)}$, $(i, j) \in \mathcal{J}$, $k = 0, 1, 2, \dots N$, are real constants, and define $h_{(i,j)}$ by*

$$h_{(i,j)}(t) = \sum_{k=0}^{N} h^k_{(i,j)} t^k, \quad (26)$$

$0 < t < 1$.

Then there exists an open region $\widetilde{S}_N^{\text{neu}} \subset S_N^{\text{neu}} \subset (-1, 1)^2$ with $|\widetilde{S}_N^{\text{neu}}| > 0$ such that the following holds. For all $(a, b) \in \widetilde{S}_N^{\text{neu}}$, there exist constants $p_{i,j}$, $i = 1, 2, \dots N + 1$, $j = 0, 1, 2$, such that

$$\sigma = \begin{bmatrix} \sigma_{1,2}(t) \\ \sigma_{2,3}(t) \\ \sigma_{3,1}(t) \end{bmatrix} = \sum_{i=1}^{N+1}\sum_{j=0}^{2} p_{i,j}\boldsymbol{w}_{i,j}t^{\beta_{i,j}-1} \quad (27)$$

satisfies

$$\max_{(i,j)\in\mathcal{J}} |h_{(i,j)} - \mathcal{K}_{\text{neu}}[\sigma]_{(i,j)}| \le Ct^{N+1} \quad (28)$$

for $0 < t < 1$, where C is a constant.

4. Conjectures

There are four independent parameters that completely describe the triple junction problem, any two out of the three angles $\{\theta_1, \theta_2, \theta_3\}$, and any two of the parameters $\{d_{(1,2)}, d_{(2,3)}, d_{(3,1)}\} = \{b, c, a\}$. Let $Y \subset \mathbb{R}^4$ denote the subset of $\mathbb{R}^4$ associated with the four free parameters that completely describe any triple junction given by

$$Y = \{(\theta_1, \theta_2, a, b) : 0 < \theta_1, \theta_2 < 2\pi,\ \theta_1 + \theta_2 < 2\pi,\ -1 < a, b < 1\}. \tag{29}$$

When θ_1, θ_2, are irrational multiples of π, and (a, b) are in the neighborhoods of $a = 0$, $b = 0$, and $c = 0$, Theorems 3.4 and 3.8 construct an explicit basis of nonsmooth functions for the solutions of $\mathcal{K}_{\mathrm{dir}}[\sigma] = \boldsymbol{h}$ and $\mathcal{K}_{\mathrm{neu}}[\sigma] = \boldsymbol{h}$ and show that this basis maps onto the space of boundary data given by piecewise polynomials on each of the edges meeting at the triple junction. However, extensive numerical studies suggest that both of these results can be improved significantly. In particular, we believe that this analysis extends to all $(\theta_1, \theta_2, a, b) \in Y$, except for a set of measure zero. Moreover, on the measure-zero set where this basis is not sufficient, we expect the solution to have additional logarithmic singularities; including functions of the form $t^\beta \log(t)\boldsymbol{v}$ should be sufficient to fix the deficiency of the basis. We expect the analysis to be similar in spirit to the analysis carried out for the solution of Dirichlet and Neumann problems for Laplace's equations on vicinity of corners; see [Serkh and Rokhlin 2016; Serkh 2019].

In this section, we present a few open questions for further extending Theorems 3.4 and 3.8, and present numerical evidence to support these conjectures.

4A. *Existence of $\beta_{i,j}$*. The solutions $\beta_{i,j}$, $i = 0, 1, 2, \dots,$ $j = 0, 1, 2$, are constructed as the implicit solutions of $\det \mathcal{A}_{\mathrm{dir}}(a, b, \beta) = 0$ (recall that $\det \mathcal{A}_{\mathrm{dir}}(a, b, \beta) = \det \mathcal{A}_{\mathrm{neu}}(a, b, \beta)$). Note that $\det \mathcal{A}_{\mathrm{dir}}(a, b, \beta) = \sin(\pi\beta)\cdot\alpha(a, b, c; \beta)$, where α is as defined in (52). From this, it follows that $\beta_{i,0} = i$ always satisfies $\det \mathcal{A}_{\mathrm{dir}}(a, b, \beta) = 0$ for all θ_1, θ_2, and that $\beta_{0,j} = 0$ results in three linearly independent basis functions of the form $t^\beta \boldsymbol{v}$ since $\mathcal{A}_{\mathrm{dir}}(a, b, 0) = 0$.

The remaining $\beta_{i,j}$, $i = 1, 2, \dots,$ $j = 1, 2$, are constructed in the following manner. $\alpha(a, b, c; \beta)$ simplifies significantly along $a = 0$, $b = 0$, and $c = 0$, and the existence of $\beta_{i,j}$ which satisfy $\det \mathcal{A}_{\mathrm{dir}}(a, b, \beta) = 0$ is guaranteed based on the explicit construction detailed in [Hoskins 2018]. The construction then uses the implicit function theorem to extend the existence of $\beta_{i,j}$ to a subset of $(a, b) \in (-1, 1)^2$. The implicit function theorem is a local result and only guarantees existence in local neighborhoods of the initial points. However, extensive numerical evidence suggests that the $\beta_{i,j}$ are well-defined and analytic for all $(a, b) \in (-1, 1)^2$ and all θ_1, θ_2. In Figure 3, we plot a few of these functions to illustrate this result.

Conjecture 4.1. *There exists a countable collection of $\beta_{i,j}$, $i = 1, 2, \dots,$ $j = 1, 2$, which satisfy $\alpha(a, b, c; \beta_{i,j}) = 0$. Moreover, these $\beta_{i,j}$ are analytic functions of θ_1, θ_2, a, and b for all $(\theta_1, \theta_2, a, b) \in Y$.*

An alternate strategy for proving this result is by making the following observation. For fixed θ_1, θ_2, consider the curve $\gamma_m : (m, m+1) \to \mathbb{R}^3$ defined by

$$\gamma_m(\beta) := \frac{1}{\sin(\pi\beta)}(\sin\beta(\pi - \theta_2), \sin\beta(\pi - \theta_3), \sin\beta(\pi - \theta_1)), \tag{30}$$

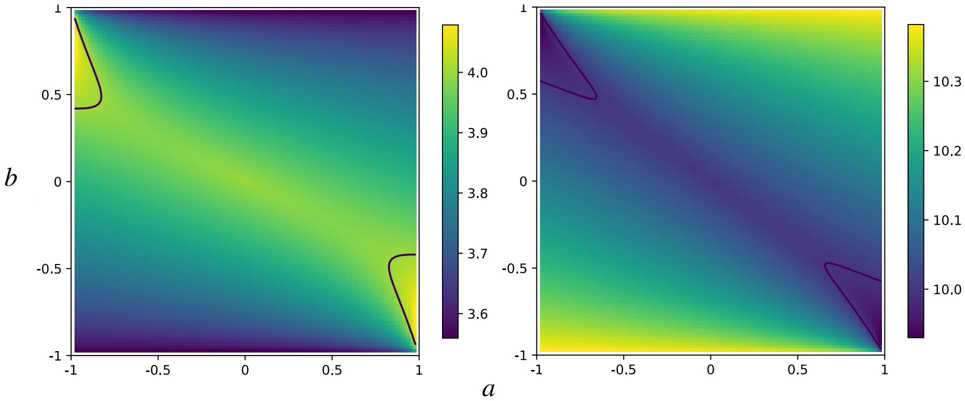

Figure 3. Plots for $\beta(a, b)$ which satisfy $\det \mathcal{A}_{\text{dir}}(a, b, \beta) = 0$ at a triple junction with angles $\theta_1 = \pi/\sqrt{2}$, $\theta_2 = \pi/\sqrt{3}$, with $\beta(0, 0) = 4$ for the figure on the left, and $\beta(0, 0) = 10$ for figure on the right. In both of the figures, the solid black lines indicate sections of the conjectured measure-zero set S defined in Conjecture 4.2.

where m is an integer. This defines a curve in $\mathbb{R}^3$ for which $|\gamma_m| \to \infty$ for each m. Then consider the family of hyperboloids parametrized by (a, b) given by

$$H(x, y, z; a, b) := -\frac{b(a+b)}{1+ab}x^2 - \frac{a(a+b)}{1+ab}y^2 + abz^2 + 1 = 0. \tag{31}$$

It follows immediately that the solutions to $\alpha(a, b, c; \beta) = 0$ can be characterized geometrically as points in the intersection of the hyperboloid $H(x, y, z; a, b)$ with the curve γ_m.

4B. ***Completeness of the singular basis.*** Having identified the $\beta_{i,j}$ and the corresponding null vectors $\boldsymbol{v}_{i,j}$ for $\mathcal{A}_{\text{dir}}$ and $\boldsymbol{w}_{i,j}$ for $\mathcal{A}_{\text{neu}}$, the second part of the proof shows that every set of boundary data which is a polynomial of degree less than or equal to N on each of the edges has a solution to the integral equations (14) and (15) in the $\boldsymbol{v}_{i,j} t^{\beta_{i,j}}$ basis for $\mathcal{K}_{\text{dir}}$ and $\boldsymbol{w}_{i,j} t^{\beta_{i,j}-1}$ for $\mathcal{K}_{\text{neu}}$ which agrees with the boundary data with error $O(t^{N+1})$.

This part of the proof relies on constructing an explicit mapping from the coefficients of the density σ in the $\boldsymbol{v}_{i,j} t^{\beta_{i,j}}$ to the coefficients of Taylor expansions for $\mathcal{K}_{\text{dir}}[\sigma]$. Then, along $a = 0$, $b = 0$, or $c = 0$, based on the results in [Hoskins 2018], we show that this mapping is invertible along these edges. It then follows from the continuity of determinants that the mapping is invertible for open neighborhoods of the line segments $a = 0$, $b = 0$, $c = 0$. This implies that in the basis $\boldsymbol{v}_{i,j} t^{\beta_{i,j}}$ there exists a σ such that $|\mathcal{K}_{\text{dir}}[\sigma] - \boldsymbol{h}| \le O(t^{N+1})$ for all boundary data f in the space of polynomials with degree less than or equal to N.

While we prove this result for an open neighborhood (a, b) of the line segments $a = 0$, $b = 0$, $c = 0$, when the angles θ_1, θ_2 are irrational multiples of π, we expect the bases to have this property for all $(\theta_1, \theta_2, a, b) \in Y$ except for a measure-zero set. Moreover, this measure-zero set is the set of $(\theta_1, \theta_2, a, b)$ for which the multiplicity of $\beta_{i,j}$ as a repeated root of $\det \mathcal{A}_{\text{dir}}(a, b, \beta_{i,j}) = 0$ is not the same as the dimension of the null space of $\mathcal{A}_{\text{dir}}(a, b, \beta_{i,j})$.

Conjecture 4.2. *Suppose that Conjecture 4.1 holds; i.e., $\beta_{i,j} : Y \to \mathbb{R}$ are analytic functions. Suppose further that $h^k_{(i,j)}$, $(i, j) \in \mathcal{J}$, $k = 0, 1, 2, \dots N$, are real constants, and suppose that*

$$h_{(i,j)}(t) = \sum_{k=0}^{N} h^k_{(i,j)} t^k, \tag{32}$$

$0 < t < 1$. *Then there exists a measure-zero set S such that for all $(\theta_1, \theta_2, a, b) \in Y \setminus S$ the following result holds. There exist constants $p_{i,j}$, $i = 0, 1, \dots N$, $j = 0, 1, 2$, such that*

$$\sigma = \begin{bmatrix} \sigma_{1,2}(t) \\ \sigma_{2,3}(t) \\ \sigma_{3,1}(t) \end{bmatrix} = \sum_{i=0}^{N} \sum_{j=0}^{2} p_{i,j} \boldsymbol{v}_{i,j} t^{\beta_{i,j}} \tag{33}$$

satisfies

$$\max_{(i,j)\in\mathcal{J}} |h_{(i,j)} - \mathcal{K}_{\text{dir}}[\sigma]_{(i,j)}| \le C t^{N+1} \tag{34}$$

for $0 < t < 1$, *where C is a constant.*

In Figure 3, we plot sections of the zero measure set on which Conjecture 4.2 does not hold.

5. Discretization of (14) and (15)

In this section we discuss a numerical method for solving (14) and (15) for the unknown densities σ, ρ which exploits the analysis of their behavior in the vicinity of triple junctions. There are two general approaches for discretizing these integral equations: Galerkin methods, in which the densities ρ and σ are represented directly in terms of appropriate basis functions, and Nyström methods, where the solution is represented in terms of its values at specially chosen discretization nodes. In this paper, we use a Nyström discretization for solving (14), though we note that the expansions in Theorems 3.4 and 3.8 can also be used to construct efficient Galerkin discretizations.

In [Bremer et al. 2010], the authors developed a Nyström discretization for resolving the singular behavior of solutions to integral equations in the vicinity of corners. In this approach, the authors obtain a basis of solutions to the integral equation in the vicinity of the corner by solving a small number of local problems. Based on these families of solutions, discretization nodes capable of interpolating the span of these solutions, coupled with quadratures for handling far-field interactions (inner products of the basis of solutions with smooth functions), and special quadratures for handling near interactions (for resolving the near-singular behavior of the kernel in the vicinity of the corner) are developed. This approach was later specialized for the solution of Laplace's equation on polygonal domains to obtain universal discretization nodes, and quadrature rules [Bremer and Rokhlin 2010].

Recent advances in the analysis of integral equations for Laplace's equation have provided analytic representations of solutions to integral equations in the vicinity of the corners [Serkh and Rokhlin 2016; Serkh 2019], obviating the need for obtaining the span of solutions in the vicinity of corners through numerical means. Based on the approach above, these analytical results have been exploited to construct universal discretization and quadratures for solutions in vicinity of corners [Hoskins et al. 2017]. Below we briefly discuss the construction of the Nyström discretization in [Hoskins et al. 2017]. Let $\mathcal{F}$ denote the family of functions

$$\mathcal{F} = \left\{t^{\beta} \text{ for all } \beta \in \{0\} \cup \left[\tfrac{1}{2}, 50\right],\ 0 < t < 1\right\}. \tag{35}$$

Then there exist $t_j \in [0, 1]$, $w_j > 0$, an orthogonal basis $\phi_j(t)$, $j = 1, 2, \dots, \boldsymbol{k}_{\mathrm{AB}} = 36$, and a $\boldsymbol{k}_{\mathrm{AB}} \times \boldsymbol{k}_{\mathrm{AB}}$ matrix V whose condition number is $O(1)$, with the following features. For any $f \in \mathcal{F}$, there exists c_j such that

$$\left| f(t) - \sum_{j=1}^{\boldsymbol{k}_{\mathrm{AB}}} c_j \phi_j(t) \right|_{\mathbb{L}^2[0,1]} < \varepsilon. \tag{36}$$

Let $f_j = f(t_j)\sqrt{w_j}$ denote the samples of the function at the discretization nodes scaled by the square root of the quadrature weights. The matrix V maps f_j to its coefficients c_j in the ϕ_j basis. Finally the weights w_j are such that

$$\left| \int_0^1 f(t)\,dt - \sum_{j=1}^{\boldsymbol{k}_{\mathrm{AB}}} f_j \sqrt{w_j} \right| = \left| \int_0^1 f(t)\,dt - \sum_{j=1}^{\boldsymbol{k}_{\mathrm{AB}}} f(t_j) w_j \right| \le \varepsilon. \tag{37}$$

Specialized quadrature rules for handling the near-singular interaction between corner panels which meet at the same vertex are also constructed. The Dirichlet problem for Laplace's equation can then be discretized using panels with scaled Gauss–Legendre nodes for panels which are away from corners, and using scaled nodes t_j for panels at corners.

In the vicinity of triple junctions, the behavior of the solution σ of (14) can be represented to high-order as a linear combination of functions in $\mathcal{F}$. Thus the discretization for the Dirichlet problem discussed above can be used to obtain a Nyström discretization for (14). Unfortunately, the same is not true when solving (15), since the singular behavior of ρ is not contained in the span of $\mathcal{F}$. In particular, the leading-order singularity in ρ is of the form t^{β}, where $\beta \in \left(-\tfrac{1}{2}, 0\right)$. The nature of the singularity of ρ is similar to the singular behavior of solutions to integral equations corresponding to the Neumann problem on polygonal domains.

Recall that (15) is the adjoint of (14). Thus, formally, one could use the transpose of the Nyström discretization of (14) to solve (15). Specifically, if $\bar{\rho} = \{\rho_j\}_{j=1}^{N}$ are the unknown values of ρ at the discretization nodes, and $\bar{g} = \{g_j\}_{j=1}^{N}$ denote the samples of the boundary data for (15) at the discretization nodes, then we solve the linear system

$$M^T \bar{\rho} = \bar{g}, \tag{38}$$

where M is the matrix corresponding to Nyström discretization of (14). The solution $\bar{\rho}$ is a high-order accurate weak solution for the density ρ which can be used to evaluate the solution to (15) accurately away from the corner panels of the boundary Γ. This weak solution can be further refined to obtain accurate approximations of the potentials in the vicinity of corner panels through solving a sequence of small linear systems for updating the solution ρ_j in the vicinity of the corner panels. This procedure is discussed in detail in [Hoskins and Rachh 2020].

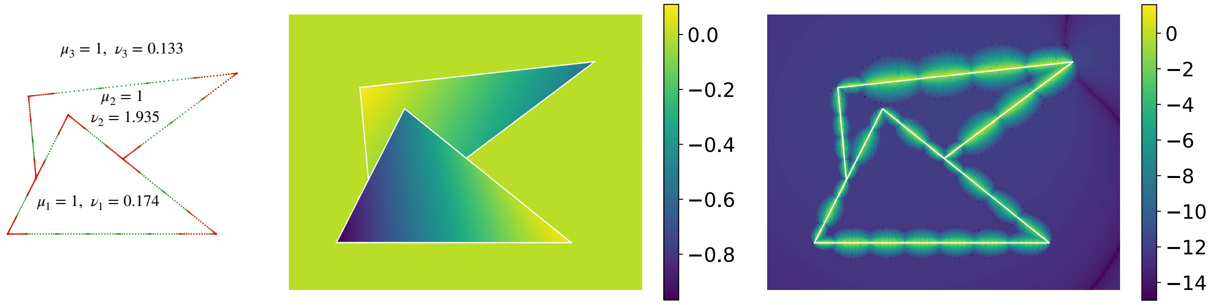

Figure 4. Discretization of geometry along with material parameters μ_i, ν_i (left), the panels at corners/triple junctions are indicated in red; exact solution u_j in the domains (center), and $\log_{10}$ of the absolute error in the solution (right). The geometry consists of 7 vertices, 8 edges, 3 regions, and is discretized with 768 points. In order for the solution of the linear system to converge to a residual of 10^{-16}, GMRES required 35 iterations for (14) and 48 iterations for (15).

6. Numerical examples

We illustrate the performance of the algorithm with several numerical examples. In each of the problems let Ω_0 denote the exterior domain and Ω_i, $i = 1, 2, \ldots N_r$, denote the interior regions. Let $c_{j,k}$, $k = 1, 2, \ldots 10$, denote points outside of the region Ω_j for $j = 1, 2, \ldots N_r$. The results in Sections 6B and 6C have been computed using dense linear algebra routines, while the results in Sections 6A and 6D have been computed using GMRES where the matrix vector product computation has been accelerated using fast multipole methods [Greengard and Rokhlin 1986].

6A. ***Accuracy.*** In order to demonstrate the accuracy of our method we solve the PDE with boundary data corresponding to known harmonic functions using our discretization of the integral equation formulation. We set $u_j(x) = \sum_{k=1}^{10} \log|x - c_{j,k}|$ and set $u \equiv 0$ for $x \in \Omega_0$. We then compute the boundary data

$$f_i = \mu_{\ell(i)} u_{\ell(i)} - \mu_{r(i)} u_{r(i)}, \qquad g_i = \mu_{\ell(i)} \frac{\partial u_{\ell(i)}}{\partial n} - \mu_{r(i)} \frac{\partial u_{r(i)}}{\partial n}, \tag{39}$$

and solve for σ, ρ. Given the discrete solution for σ, ρ, we compare the computed solution and plot the error in the computed at targets in the interior of each of the regions. In Figures 4 and 5, we demonstrate the results for two sample geometries.

Remark 6.1. Note that we do not use special quadratures for handling near boundary targets which is responsible for the loss of accuracy close to the boundary. For panels away from the corner, the potential at near boundary targets can be computed accurately using several standard methods such as quadrature by expansion, or product quadrature; see [Klöckner et al. 2013; Helsing and Ojala 2008b; Barnett et al. 2015]. In order to evaluate the solution at points lying close to a corner panel, a different approach is required. A detailed description of a computationally efficient algorithm for evaluating the solution accurately arbitrarily close to a corner is presented in [Hoskins and Rachh 2020].

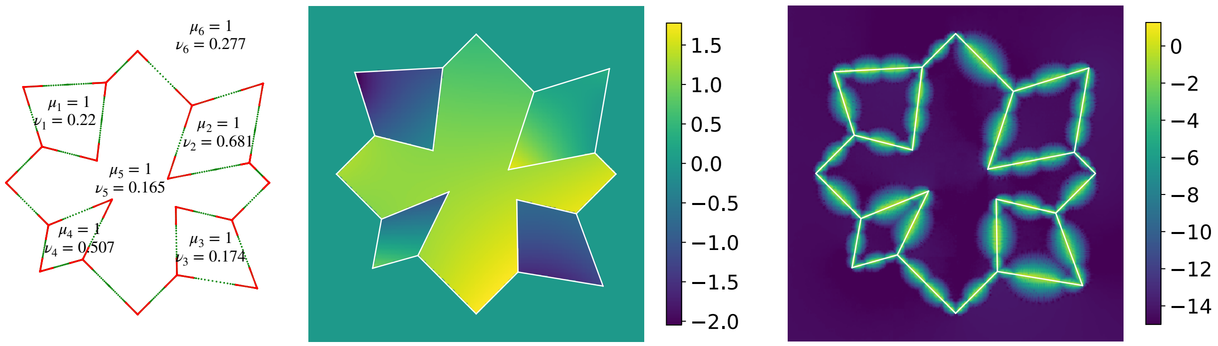

Figure 5. Discretization of geometry along with material parameters μ_i, ν_i (left), the panels at corners/triple junctions are indicated in red; exact solution u_j in the domains (center), and $\log_{10}$ of the absolute error in the solution (right). The geometry consists of 20 vertices, 24 edges, 5 regions, and is discretized with 1952 points. In order for the solution of the linear system to converge to a residual of 10^{-16}, GMRES required 22 iterations for (14) and 28 iterations for (15).

6B. ***Condition number dependence on μ, ν.*** In this section, we discuss the dependence of the condition number of the discretized linear systems as a function of the material parameters of the regions. Recall that the condition number of a linear system A, which we denote by $\kappa(A)$, is the ratio of the largest singular value $s_{\max}$ to the smallest singular value $s_{\min}$, i.e., $\kappa(A) = s_{\max}/s_{\min}$. As discussed in Section 3, for fixed angles the integral equation and the analytical behavior of integral equations (14) and (15) are solely a function of $d_{(1,2)}, d_{(2,3)}, d_{(3,1)}$ defined in (19). Furthermore, $d_{(1,2)}$ can be expressed in terms of $d_{(3,1)}, d_{(2,3)}$ which are contained in the interval $(-1, 1)$. As before, let $a = d_{(3,1)}$ and $b = d_{(2,3)}$. Since the discrete linear system corresponding to (15) is the adjoint of the linear system corresponding to (14), it suffices to study the condition number for either linear system.

In Figure 6, we plot the condition number of the discretization of (14) as we vary $(a, b) \in (-1, 1)^2$ by holding the values of μ in each of the regions to be fixed. In particular, we set $\mu_1 = 0.37$, $\mu_2 = 0.81$, $\mu_3 = 1$, and $\nu_3 = 0.77$. The constants ν_1, ν_2 can then be defined in terms of (a, b) as

$$\nu_1 = \frac{\nu_3 \mu_1}{\mu_3} \frac{1+a}{1-a}, \qquad \mu_2 = \frac{\nu_3 \mu_2}{\mu_3} \frac{1-b}{1+b}. \tag{40}$$

We note that the problem is well-behaved for almost all values of (a, b) and becomes ill-conditioned as we approach the lines $b = -1$ and $a = 1$. This behavior is expected since the underlying physical problem also has rank-deficiency along these limits since these values of the parameters correspond to interior Neumann problems in regions 1 and 2 respectively.

6C. ***Condition number dependence on angles at the triple junction.*** In this section we discuss the dependence of the condition number of the discretized linear systems as a function of the angles at the triple junction. Let $\theta_1, \theta_2, \theta_3$, denote the angles at the triple junction; then $\theta_1 + \theta_2 + \theta_3 = 2\pi$. The three angles at any triple junction can be parametrized by θ_1, θ_2 in the simplex $\{(\theta_1, \theta_2) : \theta_1 > 0, \theta_2 > 0, \theta_1 + \theta_2 < 2\pi\}$.

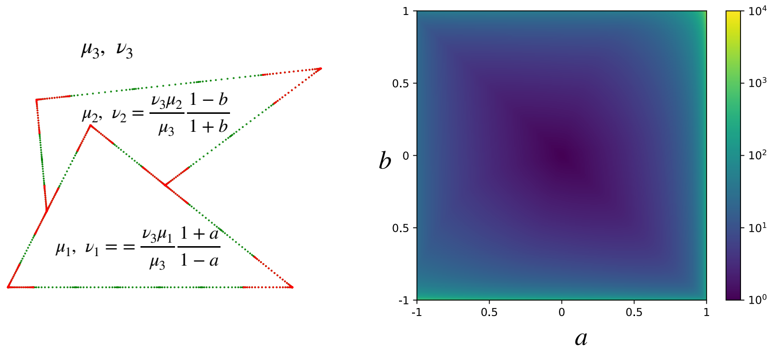

Figure 6. Left: discretization of geometry and material parameters μ, ν as a function of a, b. Right: condition number as a function of (a, b) with $\mu_1 = 0.37$, $\mu_2 = 0.81$, $\mu_3 = 1$, and $\nu_3 = 0.77$.

Suppose that we split this simplex into four regions as shown in Figure 7. By symmetry it suffices to vary the angles $(\theta_1, \theta_2) \in (0, \pi)^2$.

The physical problem as either of the angles approach 0 or 2π becomes increasingly ill-conditioned due to close-to-touching interactions on the entire edge (not just near the corner). In order to avoid these issues and to automate geometry generation as we vary the angles θ_1, θ_2, we use two different types of geometries for regions I and IV, which are shown in Figure 7.

Resolving the close-to-touching interactions has numerical consequences as well; due to the increased number of quadrature nodes required as the angles tend to 0 in the universal quadrature rules. In order for the universal quadrature rules to remain efficient, they are generated for the range $(\theta_1, \theta_2) \in \left(\frac{\pi}{12}, 2\pi - \frac{\pi}{12}\right)$. Regions with narrower angles should be handled on a case-by-case basis and regions with careful discretization of the boundary should be coupled with special purpose quadrature rules which account for the specific singular behavior of the solutions in the vicinity of triple junctions. In Figure 7, the top right missing corner corresponds to $\theta_3 \in \left(0, \frac{\pi}{12}\right)$.

Referring to Figure 7, we observe that the condition number of the discrete linear systems varies mildly as we vary the angles θ_1, θ_2, with a maximum condition number of 2.8. The discontinuity in the plot is explained by the different choice of geometries for regions I, IV.

6D. ***Application*: *polarization computation.*** In this section, we demonstrate the efficiency of our approach for computing polarization tensors for a perturbed hexagonal lattice with cavities. The polarization computation corresponds to the following particular setup of the triple junction problem, $\mu_i = 1$, $f_i = 0$, $\nu_i = \varepsilon_i$, where ε_i denotes the permittivity of the medium, and $g_1(\boldsymbol{x}) = (\varepsilon_{\ell(i)} - \varepsilon_{r(i)}) n_1(\boldsymbol{x})$ or $g_2(\boldsymbol{x}) = (\varepsilon_{\ell(i)} - \varepsilon_{r(i)}) n_2(\boldsymbol{x})$, where $\boldsymbol{x} = (x_1, x_2) \in \Gamma_i$, $\boldsymbol{n}(\boldsymbol{x}) = (n_1(\boldsymbol{x}), n_2(\boldsymbol{x}))$, and $\varepsilon_{\ell(i)}, \varepsilon_{r(i)}$ are the conductivities of the regions on either side of the edge Γ_i. If u_1 is the solution corresponding to g_1 and u_2 is the solution corresponding to g_2, then the polarization tensor P is the 2×2 matrix given by

$$P = \begin{bmatrix} \int_\Gamma x_1 \cdot (\partial u_1 / \partial n)\, ds & \int_\Gamma x_2 \cdot (\partial u_1 / \partial n)\, ds \\ \int_\Gamma x_1 \cdot (\partial u_2 / \partial n)\, ds & \int_\Gamma x_2 \cdot (\partial u_2 / \partial n)\, ds \end{bmatrix}. \tag{41}$$

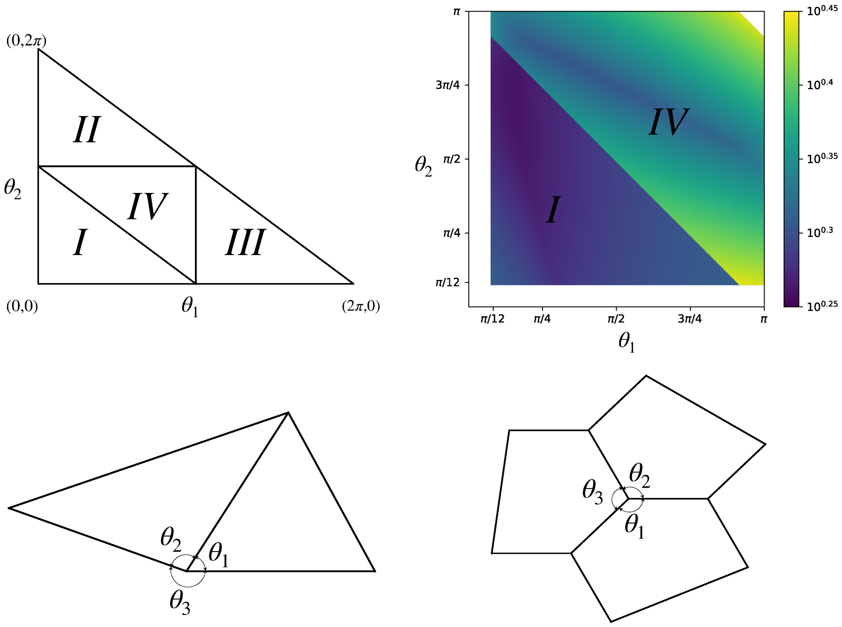

Figure 7. Regions I–IV in (θ_1, θ_2) simplex (top left); condition number of discretized linear system corresponding to (14) as a function of (θ_1, θ_2) (top right); sample domain for (θ_1, θ_2) in region I (bottom left); sample domain for (θ_1, θ_2) in region IV (bottom right).

Note that in this particular setup, we only need to solve the problem corresponding to the operator $\mathcal{K}_{\text{neu}}$, as the solution σ for $\mathcal{K}_{\text{dir}}[\sigma] = 0$ is $\sigma = 0$. Let ρ_1, ρ_2 denote the solutions of (15) corresponding to boundary data g_1 and g_2 respectively. Using properties of the single-layer potential, the integrals of the polarization tensor can be expressed in terms of ρ as

$$P = \begin{bmatrix} \int_\Gamma x_1 \cdot \rho_1 \, ds & \int_\Gamma x_2 \cdot \rho_1 \, ds \\ \int_\Gamma x_1 \cdot \rho_2 \, ds & \int_\Gamma x_2 \cdot \rho_2 \, ds \end{bmatrix}. \tag{42}$$

We compare the efficiency of our approach to RCIP, which to the best of our knowledge is the state-of-the-art method for such problems. The geometry is generated using a regular hexagonal lattice inside the unit square whose vertices are perturbed in a random direction by a tenth of the side length, and the permittivity ε is region i is given by 10^{c_i}, where c_i is a uniform random number between $[-1, 1]$. The choice of parameters for the problem setup is identical to the setup in Section 11 in [Helsing and Ojala 2008a].

We discretize the geometry with 3 panels on each edge of roughly equal size, and the reference solution is computed using 5 panels on each edge. The geometry contains 10688 vertices, 15855 edges, and 5189 regions. There are 1395240 degrees of freedom for the coarse discretization (approximately 88 degrees of freedom per edge) and 1902600 degrees of freedom for the reference solution. These discretizations required 131 iterations for GMRES to converge to a relative residual of 10^{-16}, and the absolute error in

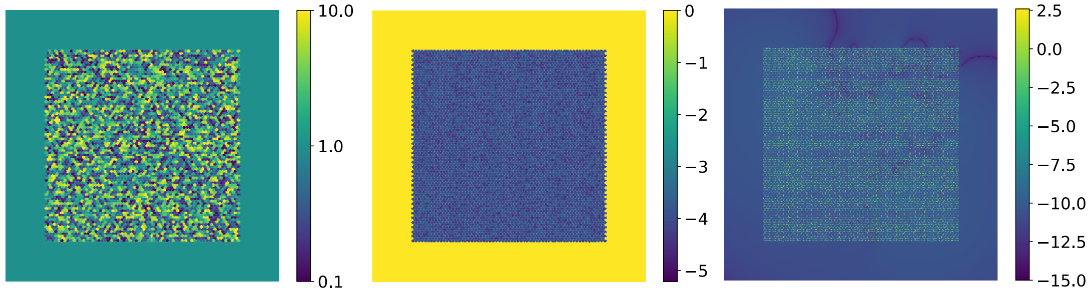

Figure 8. Material parameters ν_i for each of the regions (left), exact solution u_j in the domains (center), and $\log_{10}$ of the absolute error in the solution (right). The geometry consists of 10688 vertices, 15855 edges, 5189 regions, and is discretized with 1395240 points. In order for the solution of the linear system to converge to a residual of 10^{-16}, GMRES required 138 iterations for (14), and 130 iterations for (15) in the accuracy tests, and 131 and 130 iterations (for g_1, and g_2 respectively) for (15) in computing the polarization tensors.

the polarization tensor when compared to the reference solution is 5.1×10^{-12}. In comparison, RCIP using approximately 71 degrees of freedom per edge in a hexagonal lattice with 5293 inclusions obtained an accuracy of 2×10^{-14} in computing the 2, 2 entry of the polarization matrix and required 105 GMRES iterations to converge.

The polarization tensor for this configuration, correct to 13 significant digits, is given by

$$P = \begin{bmatrix} -0.038291586646 & -0.004056508957 \\ -0.004056508957 & 0.045585776453 \end{bmatrix}. \tag{43}$$

In Figure 8, we plot the material parameters ν_i, an analytical solution generated using a process similar to the one described in Section 6A, and the error in the computed solution using our discretization.

Remark 6.2. We note that the performance of our approach is close to current state of the art methods such as RCIP [Helsing 2013]. In our examples, further improvements in speed can be achieved using additional compression techniques to reduce the degrees of freedom in the resulting linear system [Bremer et al. 2015; Greengard et al. 2009].

7. Concluding remarks and future work

In this paper we analyze the systems of boundary integral equations which arise when solving the Laplace transmission problem in composite media consisting of regions with polygonal boundaries. Our discussion is focused on the particular case of composite media with triple junctions (points at which three distinct media meet), though our analysis extends to higher-order junctions in a natural way.

We show that under some restrictions the solutions to the boundary integral equations corresponding to a triple junction are well-approximated by a linear combination of powers t^{β_j}, where t denotes the distance from the corner along the edge, and the β_j, $j = 1, 2, \ldots$, form a countable collection of real numbers obtained by solving a certain equation depending only on the material properties of the media and the angles at which the interfaces meet.

In addition to the theoretical interest of the result, our analysis also enables an easy construction of near-optimal discretizations for triple junctions. In particular, RCIP, which is the leading method for solving electrostatic problems on multiple junction interfaces, requires approximately 71 discretization nodes per edge to compute solutions to near machine precision accuracy, whereas our proposed discretization achieves an accuracy of 5×10^{-12} using roughly 88 discretization nodes per edge. Finally, we illustrate the properties of this discretization with a number of numerical examples.

The results of this paper admit a number of natural extensions and generalizations. Firstly, the analysis outlined in this paper extends almost immediately to junctions involving greater numbers of media. However, the construction of an efficient Nyström discretization of higher-order junctions requires special care since the solutions to corresponding integral equations are not $\mathbb{L}^2$ functions on the boundary; in fact the solutions are known to be $\mathbb{L}^1$ functions on the boundary [Helsing 2011]. Secondly, with a small modification a similar analysis should be possible for boundary integral equations arising from triple junction problems for other partial differential equations such as the Helmholtz equation, Maxwell's equations, and the biharmonic equation. This line of inquiry is being vigorously pursued and will be reported at a later date.

Finally, a similar approach will also work for generating discretizations of triple junctions in three dimensions. This is particularly valuable since geometric singularities in three-dimensions can often result in prohibitively large linear systems. Accurate discretization with few degrees of freedom would greatly improve the size and complexity of systems which could be simulated.

Appendix A. Analysis of $\mathcal{K}_{\mathbf{dir}}$

First we present the proof of Theorem 3.2. In order to do so, we require the following technical lemma which describes the double-layer potential defined on a straight line segment with density s^β at an arbitrary point near the boundary. Here s is the distance along the segment.

Lemma A.1. *Suppose that* Γ *is an edge of unit length oriented along an angle* θ, *parametrized by* $s(\cos(\theta), \sin(\theta))$, $0 < s < 1$. *Suppose that* $\boldsymbol{x} = t(\cos(\theta+\theta_0), \sin(\theta+\theta_0))$ (*see Figure 9*) *where* $0 < t < 1$, *and* $\boldsymbol{x} \notin \Gamma$. *Suppose that* $\sigma(s) = s^\beta$ *for* $0 < s < 1$, *where* $\beta \geq 0$. *If* β *is not an integer*, *then*

$$\mathcal{D}_\Gamma[\sigma](\boldsymbol{x}) = \frac{\sin(\beta(\pi-\theta_0))}{2\sin(\pi\beta)} t^\beta + \frac{1}{2\pi} \sum_{k=1}^{\infty} \frac{\sin(k\theta_0)}{\beta-k} t^k. \tag{44}$$

If $\beta = m$ *is an integer*, *then*

$$\mathcal{D}_\Gamma[\sigma](\boldsymbol{x}) = \frac{(\pi-\theta_0)\cos(m\theta_0)}{2\pi} t^m - \frac{\sin(m\theta_0)}{2\pi} t^m \log(t) + \frac{1}{2\pi} \sum_{\substack{k=1\\k\neq m}}^{\infty} \frac{\sin(k\theta_0)}{m-k} t^k. \tag{45}$$

In the following lemma, we compute the potential $\mathcal{K}_{\text{dir}}[\boldsymbol{v}t^\beta]$, in the vicinity of a triple junction with angles $\theta_1, \theta_2, \theta_3$, and material parameters $\boldsymbol{d} = (d_{(1,2)}, d_{(2,3)}, d_{(3,1)})$, where $\boldsymbol{v} \in \mathbb{R}^3$ and β are constants (see Figure 2).

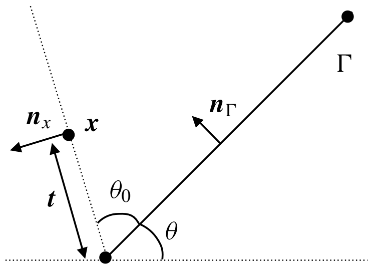

Figure 9. Illustrative figure for geometry in Lemma A.1.

Lemma A.2. *Consider the geometry setup of the single vertex problem presented in Section 3. For a constant vector* $\boldsymbol{v} \in \mathbb{R}^3$, *suppose that the density on the edges is of the form*

$$\sigma = \begin{bmatrix} \sigma_{1,2} \\ \sigma_{2,3} \\ \sigma_{3,1} \end{bmatrix} = \boldsymbol{v} t^{\beta}. \tag{46}$$

If β *is not an integer, then*

$$\mathcal{K}_{\text{dir}}[\sigma] = -\frac{1}{2\sin(\pi\beta)} \mathcal{A}_{\text{dir}}(d_{3,1}, d_{1,2}, \beta)\boldsymbol{v} t^{\beta} + \sum_{k=1}^{\infty} \frac{1}{\beta - k} \boldsymbol{C}(\boldsymbol{d}, k)\boldsymbol{v} t^{k}, \tag{47}$$

where $\mathcal{A}_{\text{dir}}$ *is defined in (21) and*

$$\boldsymbol{C}(\boldsymbol{d}, k) = \frac{1}{2\pi} \begin{bmatrix} 0 & -d_{(1,2)} \sin(k\theta_2) & d_{(1,2)} \sin(k\theta_1) \\ d_{(2,3)} \sin(k\theta_2) & 0 & -d_{(2,3)} \sin(k\theta_3) \\ -d_{(3,1)} \sin(k\theta_1) & d_{(3,1)} \sin(k\theta_3) & 0 \end{bmatrix}. \tag{48}$$

If $\beta = m$ *is an integer, then*

$$\mathcal{K}_{\text{dir}}[\sigma] = -\frac{(-1)^m}{2\pi} \mathcal{A}_{\text{dir}}(d_{3,1}, d_{1,2}, m)\boldsymbol{v} t^{m} \log(t) + \sum_{\substack{k=1 \\ k \neq m}}^{\infty} \frac{1}{m - k} \boldsymbol{C}(\boldsymbol{d}, k)\boldsymbol{v} t^{k} + \boldsymbol{C}_{\text{diag}}(\boldsymbol{d}, m)\boldsymbol{v} t^{m}, \tag{49}$$

where

$$\boldsymbol{C}_{\text{diag}}(\boldsymbol{d}, m) = -\frac{1}{2\pi} \begin{bmatrix} \pi & d_{(1,2)}(\pi - \theta_2)\cos(m\theta_2) & -d_{(1,2)}(\pi - \theta_1)\cos(m\theta_1) \\ -d_{(2,3)}(\pi - \theta_2)\cos(m\theta_2) & \pi & d_{(2,3)}(\pi - \theta_3)\cos(m\theta_3) \\ d_{(3,1)}(\pi - \theta_1)\cos(m\theta_1) & -d_{(3,1)}(\pi - \theta_3)\cos(m\theta_3) & \pi \end{bmatrix}. \tag{50}$$

Proof. The result follows from repeated application of Lemma A.1 for computing $\mathcal{D}_{(l,m):(i,j)}\sigma_{(i,j)}$. □

The proof of Theorem 3.2 then follows immediately from Lemma A.2.

We now turn our attention to the proof of Lemma 3.3, which provides a construction of β, $\boldsymbol{v}$ satisfying the conditions of Theorem 3.2. In order to do that, we first observe that if one of a, b, or c is 0, then the expression of $\det \mathcal{A}_{\text{dir}}$ simplifies significantly, and there exists an explicit construction of β satisfying $\det \mathcal{A}_{\text{dir}}(a, b, \beta) = 0$. Recall that we interchangeably use the following variables for the material properties:

$(a, b, c) = (d_{3,1}, d_{1,2}, d_{2,3})$. Having established the existence of analytic β, $\boldsymbol{v}$ on a 1-dimensional manifold which is a subset $(a, b) \in (-1, 1)^2$, we now analytically continue these values of β, $\boldsymbol{v}$ to carve out the open region S on which β, $\boldsymbol{v}$ can be analytically extended. This proof is discussed in Appendix A.1.

A1. ***Existence of $\boldsymbol{\beta}$, $\boldsymbol{v}$ satisfying Theorem 3.2.*** The determinant of the matrix $\mathcal{A}_{\text{dir}}(a, b, \beta)$ is given by

$$\det \mathcal{A}_{\text{dir}}(a, b, \beta) = \sin(\pi\beta)\, \alpha(a, b, c; \beta), \tag{51}$$

where $c = -(a+b)/(1+ab)$, and

$$\alpha(a, b, c; \beta) = \sin^2(\pi\beta) + bc\sin^2(\beta(\pi-\theta_2)) + ac\sin^2(\beta(\pi-\theta_3)) + ab\sin^2(\beta(\pi-\theta_1)). \tag{52}$$

Given the formula above, for all $(a, b) \in (-1, 1)^2$ when $\beta = m \geq 0$ is an integer, $\det \mathcal{A}_{\text{dir}}(a, b, \beta) = 0$. When $m \neq 0$, the matrix $\mathcal{A}_{\text{dir}}$ has rank 2, since the matrix is similar to an antisymmetric matrix and is not identically zero. The null vector $\boldsymbol{v}$ of $\mathcal{A}_{\text{dir}}(a, b, m)$ is given by $\boldsymbol{v}_m = [\sin(m\theta_3), \sin(m\theta_1), \sin(m\theta_2)]^T$; i.e., the pair $(m, \boldsymbol{v}_m)$ always satisfies (21). When $\beta = 0$, $\mathcal{A}_{\text{dir}}(a, b, \beta) = 0$ and hence for any $\boldsymbol{v} \in \mathbb{R}^3$, the pair β, $\boldsymbol{v}$ satisfies (21). Based on this observation we set

$$\begin{aligned}
&\beta_{m,0} = m, && \boldsymbol{v}_{m,0} = [\sin(m\theta_3), \sin(m\theta_1), \sin(m\theta_2)]^T, && S_{m,0} = (-1, 1)^2,\\
&\beta_{0,0} = 0, && \boldsymbol{v}_{0,0} = [1, 0, 0]^T, && S_{0,0} = (-1, 1)^2,\\
&\beta_{1,0} = 0, && \boldsymbol{v}_{1,0} = [0, 1, 0]^T, && S_{0,1} = (-1, 1)^2,\\
&\beta_{2,0} = 0, && \boldsymbol{v}_{2,0} = [0, 0, 1]^T, && S_{0,2} = (-1, 1)^2.
\end{aligned} \tag{53}$$

We now turn our attention to constructing the remaining $\beta_{i,j}$, the corresponding vectors $\boldsymbol{v}_{i,j}$, and their regions of analyticity $S_{i,j}$, $i = 1, 2, \ldots,$ $j = 1, 2$. From (51), the remaining values of $\beta_{i,j}$ as a function of the material parameters (a, b) are defined implicitly via the roots of the equation $\alpha(a, b, c(a, b); \beta_{i,j}(a, b)) = 0$, where $c = -(a+b)/(1+ab)$ and α is defined in (52).

It turns out that the implicit solutions $\beta(a, b)$ of $\alpha(a, b, c(a, b); \beta(a, b)) = 0$, are known when $a = 0$, $b = 0$, or $c = 0$. This gives us an initial value for defining $\beta_{i,j}$ in order to apply the implicit function theorem, and extend it to a region containing the segments $a = 0$, $b = 0$, or $c = 0$. Given this strategy, let $R_1, \ldots, R_6 \subset (-1, 1) \times (-1, 1)$ be defined as follows (see Figure 10):

$$R_1 = \{(x, 0) : x > 0\}, \tag{54}$$

$$R_2 = \{(-x, 0) : x > 0\}, \tag{55}$$

$$R_3 = \{(0, x) : x > 0\}, \tag{56}$$

$$R_4 = \{(0, -x) : x > 0\}, \tag{57}$$

$$R_5 = \{(-x, x) : x > 0\}, \tag{58}$$

$$R_6 = \{(x, -x) : x > 0\}. \tag{59}$$

In the following, we will consider only the segment R_1, and construct an open region $S^1_{i,j} \subset (-1, 1)^2$ which contains R_1 on which we define a family of functions $\beta_{i,j}(a, b) : S^1_{i,j} \to \mathbb{R}$, $j = 1, 2$, which satisfy the conditions of Lemma 3.3. Analogous results hold for the open sets containing the remaining

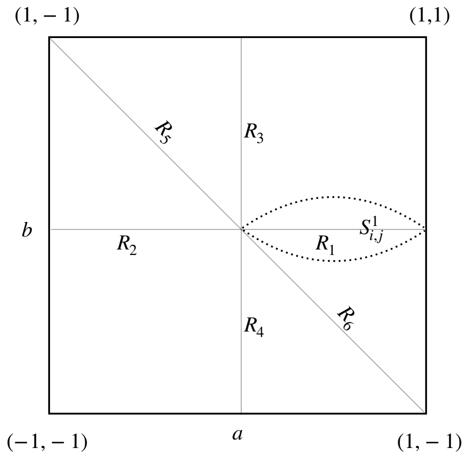

Figure 10. Illustration of the edge segments R_i, $i = 1, 2, \dots 6$, and a typical region of analyticity of $\beta_{i,j}$ denoted by $S^1_{i,j}$.

segments $R_2, R_3, \dots R_6$ with almost identical proofs. The region of analyticity for $\beta_{i,j}$ is then given by $S_{i,j} = \bigcup_{k=1}^{6} S^k_{i,j}$.

Definition A.3. For $(a, 0) \in R_1$ and $i = 1, 2 \dots$ let $\beta_{i,1}(a, 0)$ be the solution to the equation

$$\sin(\pi\beta_{i,1}) = -a \sin(\beta_{i,1}(\pi - \theta_3)) \tag{60}$$

such that

$$\lim_{a\to 0} \beta_{i,1}(a, 0) = i. \tag{61}$$

Similarly, for $i = 1, 2, \dots$ let $\beta_{i,2}(a, 0)$ be the solution to the equation

$$\sin(\pi\beta_{i,2}) = a \sin(\beta_{i,2}(\pi - \theta_3)) \tag{62}$$

such that

$$\lim_{a\to 0} \beta_{i,2}(a, 0) = i. \tag{63}$$

The existence of $\beta_{i,j}$ for $i = 1, 2, \dots$ and $j = 1, 2$ satisfying these conditions is guaranteed by the following Lemma A.4, proved in [Hoskins 2018].

Lemma A.4. *Suppose that $\delta \in \mathbb{R}$, $0 < |\delta| < 1$, and $\theta \in (0, 2\pi)$ and θ/π is irrational. Consider the equations*

$$\sin(\pi z) = \pm\delta \sin(z(\pi - \theta)).$$

Then there exist a countable collection of functions $z_i^{\pm}(\delta)$, $i = 1, 2, \dots,$ such that

(1) $\sin^2(\pi z_i^{\pm}(\delta)) = \delta^2 \sin^2(z_i^{\pm}(\delta)(\pi - \theta))$ *for all* $\delta \in [0, 1]$, *and* $i = 1, 2, \dots,$

(2) *the functions* $z_i^{\pm}$ *are analytic in* $(0, 1)$,

(3) $\lim_{\delta\to 0} z_i^{\pm}(\delta) = i$,

(4) $z_i^{+}(\delta) > i$ *and* $z_i^{-}(\delta) < i$ *for all* $\delta \in (0, 1)$.

The following lemma extends the domain of definition of the functions $\beta_{i,j}$, $j = 1, 2$, to some open subset $S^1_{i,j}$ containing R_1.

Lemma A.5. *Suppose θ_1, θ_2 and θ_3 are positive numbers summing to 2π, and $\theta_1/\pi, \theta_2/\pi$, and θ_3/π are irrational numbers. Suppose that $\beta_{i,j}$ are defined as above for $i = 1, 2, \ldots$ and $j = 1, 2$. For $a \in (0, 1)$, the function $\beta_{i,j}$ satisfies*

$$\alpha(a, 0, -a; \beta_{i,j}) = 0. \tag{64}$$

Moreover, there exists a unique extension of $\beta_{i,j}$ to an analytic function of (a, b) on an open neighborhood $R_1 \subset S^1_{i,j} \subset (-1, 1)^2$ which satisfies

$$\alpha(a, b, c(a, b); \beta_{i,j}) = 0. \tag{65}$$

Proof. We begin by observing that, for $j = 1, 2$, $\beta_{i,j}$ satisfies

$$\begin{aligned} &\alpha(a, 0, -a; \beta_{i,j}) = -a^2 \sin^2(\beta_{i,j}(\pi - \theta_3)) + \sin^2(\pi \beta_{i,j}) = 0, \\ &\frac{\partial \alpha}{\partial \beta}(a, 0, -a; \beta_{i,j}) = 2\big(-(\pi - \theta_3)a^2 \sin(\beta_{i,j}(\pi - \theta_3)) \cos(\beta_{i,j}(\pi - \theta_3)) + \pi \sin(\pi \beta_{i,j}) \cos(\pi \beta_{i,j})\big). \end{aligned} \tag{66}$$

Upon multiplication by

$$g(a; \beta) = (\pi - \theta_3)a^2 \sin(\beta(\pi - \theta_3)) \cos(\beta(\pi - \theta_3)) + \pi \sin(\pi\beta) \cos(\pi\beta)$$

and using (66) we get

$$g(a; \beta_{i,j}) \frac{\partial \alpha}{\partial \beta}(a, 0, -a; \beta_{i,j}) = -2 \sin^2(\pi \beta_{i,j}) \big(\pi^2 - a^2(\pi - \theta_3)^2 - (\pi^2 - (\pi - \theta_3)^2) \sin^2(\pi \beta_{i,j})\big),$$

which does not vanish for all $a > 0$. Thus by the implicit function theorem, there exists an analytic extension of $\beta_{i,j}$ to a neighborhood $(a, b) \in R_1 \subset S^1_{i,j} \subset (-1, 1)^2$ which satisfies $\alpha(a, b, c(a, b); \beta_{i,j}) = 0$. □

The following theorem establishes the analyticity of the null vectors of $\mathcal{A}_{\text{dir}}(a, b; \beta)$ in a neighborhood of R_1 when $\beta = \beta_{i,j}$.

Theorem A.6. *For each $j = 1, 2$, and $i = 1, 2, \ldots$, the matrix $\mathcal{A}_{\text{dir}}(a, b, \beta_{i,j})$ defined in (21) has a null vector $\boldsymbol{v}_{i,j}$ whose entries are analytic functions of (a, b) on $S^1_{i,j}$.*

Proof. Since $\beta_{i,j}$ is such that the matrix $\mathcal{A}_{\text{dir}}(a, b, \beta_{i,j})$ is singular, it has a null vector $\boldsymbol{v}_{i,j}$. Moreover, as long as $(a, b) \neq (0, 0)$ and $\beta_{i,j}$ is not an integer, the matrix $\mathcal{A}_{\text{dir}}$ has rank at least 2. Thus 0 is an eigenvalue of $\mathcal{A}_{\text{dir}}(a, b, \beta_{i,j}(a, b))$ with multiplicity 1 for all $(a, b) \in S^1_{i,j}$. Since the entries of the matrix $\mathcal{A}_{\text{dir}}$ are analytic functions of (a, b), we conclude that the entries of $\boldsymbol{v}_{i,j}$ are analytic on S^1_i. □

Finally, each $S^k_{i,j}$ is an open subset containing the segments R^k, $k = 1, 2, \ldots 6$. Then $S_{i,j} = \bigcup_{k=1}^6 S^k_{i,j}$ is an open subset of $(-1, 1)^2$ containing $\bigcup_{k=1}^6 R_k$. Thus, for any finite N, $\big|\bigcap_{i=0}^N \bigcap_{j=0}^2 S_{i,j}\big| > 0$.

A2. ***Completeness of density basis.*** Recall that for any β, $\boldsymbol{v}$ which satisfy the conditions of Theorem 3.2, and $\sigma = \boldsymbol{v}t^\beta$, the potential $\mathcal{K}_{\text{dir}}[\sigma]$ corresponding to any of these densities is an analytic function. In order to show that the potential corresponding to a particular collection of β, $\boldsymbol{v}$ span all polynomials of a fixed degree on all the three edges meeting at the triple junction, we explicitly write down the linear map from the coefficients of the density in the $\boldsymbol{v}t^\beta$ basis to the coefficients of Taylor series of the potentials on each of the edges using Lemma A.2. We then observe that this mapping is invertible along the line segments corresponding to $a = 0$, $b = 0$, or $c = 0$, and since the mapping is an analytic function of the parameters (a, b), it must also be invertible in an open region containing the segments $a = 0$, $b = 0$ or $c = 0$. This part of the proof is discussed in this section.

For any integer $N > 0$, let $\widehat{S}_N$, denote the common region of analyticity of $\beta_{i,j}$, $\boldsymbol{v}_{i,j}$, $j = 0, 1, 2$, $i = 0, 1, 2 \dots N$; i.e., $S_N = \bigcup_{k=1}^{6} S_N^k$, where $S_N^k = \bigcap_{i=0}^{N} \bigcap_{j=0}^{2} S_{i,j}^k$. By construction, $R_j \subset S_N^j$ for all N. We now prove the result Theorem 3.4 in one of the components of S_N, say S_N^1. The proof for the other components follows in a similar manner.

Let $\boldsymbol{p}_i = [p_{i,0}, p_{i,1}, p_{i,2}]^T$, and suppose that

$$\sigma(t) = \sum_{i=0}^{N} \sum_{j=0}^{2} p_{i,j} \boldsymbol{v}_{i,j} |t|^{\beta_{i,j}}. \tag{67}$$

Then, using Lemma A.2, since $\beta_{i,j}$, $\boldsymbol{v}_{i,j}$ are such that $\mathcal{A}_{\text{dir}}(a, b, \beta_{i,j}) \cdot \boldsymbol{v}_{i,j} = 0$, the potential corresponding to this density on the boundary $(\Gamma_{(1,2)}, \Gamma_{(2,3)}, \Gamma_{(3,1)})$ is given by

$$\begin{bmatrix} u_{(1,2)}(t) \\ u_{(2,3)}(t) \\ u_{(3,1)}(t) \end{bmatrix} = \sum_{i=0}^{N} \biggl(\sum_{j=0}^{N} B_{i,j} \cdot \boldsymbol{p}_j \biggr) |t|^i + O(|t|^{N+1}), \tag{68}$$

where $B_{i,j}$ are the 3×3 matrices given by

$$B_{i,j} = \begin{cases} \left[\dfrac{1}{\beta_{j,0} - i} \boldsymbol{C}(\boldsymbol{d}, i) \boldsymbol{v}_{j,0} \;\middle|\; \dfrac{1}{\beta_{j,1} - i} \boldsymbol{C}(\boldsymbol{d}, i) \boldsymbol{v}_{j,1} \;\middle|\; \dfrac{1}{\beta_{j,2} - i} \boldsymbol{C}(\boldsymbol{d}, i) \boldsymbol{v}_{j,2} \right] & \text{if } i \neq j, \\ \left[\boldsymbol{C}_{\text{diag}}(\boldsymbol{d}, i) \boldsymbol{v}_{j,0} \;\middle|\; \dfrac{1}{\beta_{j,1} - i} \boldsymbol{C}(\boldsymbol{d}, i) \boldsymbol{v}_{j,1} \;\middle|\; \dfrac{1}{\beta_{j,2} - i} \boldsymbol{C}(\boldsymbol{d}, i) \boldsymbol{v}_{j,2} \right] & \text{if } i = j \neq 0, \\ \left[\boldsymbol{C}_{\text{diag}}(\boldsymbol{d}, i) \boldsymbol{v}_{j,0} \;\middle|\; \boldsymbol{C}_{\text{diag}}(\boldsymbol{d}, i) \boldsymbol{v}_{j,1} \;\middle|\; \boldsymbol{C}_{\text{diag}}(\boldsymbol{d}, i) \boldsymbol{v}_{j,2} \right] & \text{if } i = j = 0. \end{cases} \tag{69}$$

Let $\boldsymbol{B}$ denote the $3(N+1) \times 3(N+1)$ matrix whose 3×3 blocks are given by $B_{i,j}$, $i, j = 0, 1, 2, \dots, N$.

Recall that on $R_1 \subset S_N^1$, $b = 0$, $\beta_{i,0} = i$, $\beta_{i,1}$, satisfies $\sin(\pi \beta_{i,1}) = -a \sin(\beta_{i,1}(\pi - \theta_3))$, $\beta_{i,2}$ satisfies $\sin(\pi \beta_{i,2}) = a \sin(\beta_{i,2}(\pi - \theta_3))$, $i = 0, 1, 2 \dots$, and the corresponding vectors $\boldsymbol{v}_{i,j}$, $i = 0, 1, 2 \dots$, $j = 0, 1, 2$, are given by

$$\boldsymbol{v}_{i,0} = \frac{1}{\eta_i} \begin{bmatrix} \sin(i\theta_3) \\ \sin(i\theta_1) \\ \sin(i\theta_2) \end{bmatrix}, \quad \boldsymbol{v}_{i,1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_{i,2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \tag{70}$$

where

$$\eta_i = \sqrt{\sin^2(i\theta_1) + \sin^2(i\theta_2) + \sin^2(i\theta_3)}. \tag{71}$$

Furthermore, the matrices $\boldsymbol{C}$ and $\boldsymbol{C}_{\text{diag}}$ defined in (48), and (50) respectively also simplify to

$$\boldsymbol{C} = \frac{a}{2\pi}\begin{bmatrix} 0 & 0 & 0 \\ -\sin(m\theta_2) & 0 & -\sin(m\theta_3) \\ -\sin(m\theta_1) & \sin(m\theta_3) & 0 \end{bmatrix}, \tag{72}$$

$$\boldsymbol{C}_{\text{diag}} = -\frac{1}{2\pi}\begin{bmatrix} \pi & 0 & 0 \\ a(\pi-\theta_2)\cos(m\theta_2) & \pi & -a(\pi-\theta_3)\cos(m\theta_3) \\ a(\pi-\theta_1)\cos(m\theta_1) & -a(\pi-\theta_3)\cos(m\theta_3) & \pi \end{bmatrix}. \tag{73}$$

Let $u_{(1,2),i}, u_{(2,3),i}, u_{(3,1),i}$ denote the coefficient of $|t|^i$ in the Taylor expansions of $u_{(1,2)}, u_{(2,3)}, u_{(3,1)}$ respectively. Let P denote the permutation matrix whose action is given by

$$P\begin{bmatrix} p_{0,0} \\ p_{0,1} \\ p_{0,2} \\ \hline p_{1,0} \\ p_{1,1} \\ p_{1,2} \\ \hline \vdots \\ \hline \vdots \\ \hline \vdots \\ \hline p_{N,0} \\ p_{N,1} \\ p_{N,2} \end{bmatrix} = \begin{bmatrix} p_{0,0} \\ p_{1,0} \\ \vdots \\ p_{N,0} \\ \hline p_{0,1} \\ p_{1,1} \\ \vdots \\ p_{N,1} \\ \hline p_{0,2} \\ p_{1,2} \\ \vdots \\ p_{N,2} \end{bmatrix}. \tag{74}$$

Then along R_1, the matrix PBP^T is demonstrated in Figure 11.

The matrices D_1, D_2 are diagonal and are given by

$$D_1 = \begin{bmatrix} \sin(\theta_3) & & & & \\ & \sin(2\theta_3) & & & \\ & & \ddots & & \\ & & & \sin((N-1)\theta_3) & \\ & & & & \sin(N\theta_3) \end{bmatrix}, \quad D_2 = -\frac{1}{2}\begin{bmatrix} \eta_1 & & & & \\ & \eta_2 & & & \\ & & \ddots & & \\ & & & \eta_{N-1} & \\ & & & & \eta_N \end{bmatrix}. \tag{75}$$

The matrices C_1, C_2 are Cauchy matrices whose entries are given by

$$C_{1,i,j} = \frac{1}{\beta_{i,1} - j}, \quad C_{2,i,j} = \frac{1}{\beta_{i,2} - j}. \tag{76}$$

Since we have assumed $\theta_1/\pi, \theta_2/\pi, \theta_3/\pi$, to be irrational, we note that $\eta_i > 0$ and that $\sin(m\theta_3) \neq 0$ for all $m \neq 0$. Thus, the diagonal matrices D_1, D_2 are invertible. Furthermore on $(a, 0)$, neither of $\beta_{i,1}$ or $\beta_{i,2}$ take on integer values by Lemma A.4. Thus, the Cauchy matrices C_1, C_2 are invertible.

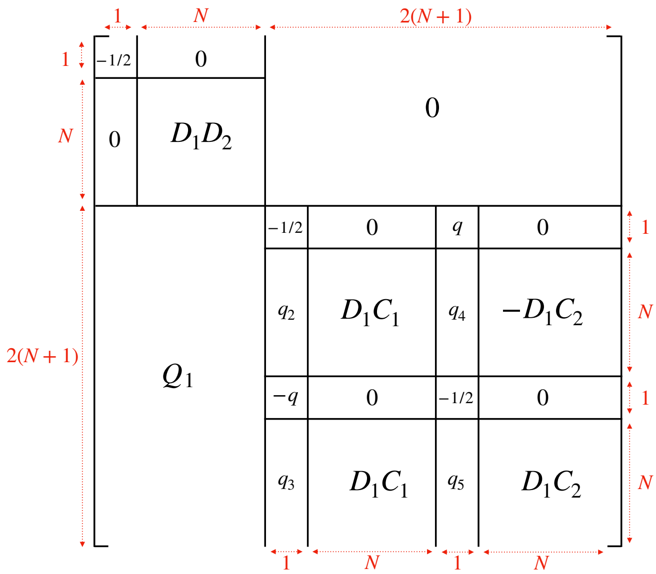

Figure 11. Structure of the matrix PBP^T.

Let T denote the bottom-right $2(N+1)\times 2(N+1)$ block. Then from the structure of $P\boldsymbol{B}P^T$ and the fact that the diagonal matrix D_1D_2 is invertible, it is clear that B is invertible if and only if T is invertible.

Remark A.7. The matrix T is the mapping from the coefficients of the singular basis of solutions for the transmission problem with angle $\pi\theta_3$ and material parameter a to the corresponding coefficients of the Taylor expansion of the potential on the edges $(2,3)$, $(3,1)$. The invertibility of T follows from the analysis in [Hoskins 2018]. We present the proof here in terms of the notation used in this paper.

Upon applying an appropriate permutation matrix P_2 to T from the right and the left, we note that

$$P_2TP_2^T = \begin{bmatrix} -\frac{1}{2} & q & 0 & 0 \\ -q & -\frac{1}{2} & 0 & 0 \\ q_2 & q_4 & D_1C_1 & -D_1C_2 \\ q_3 & q_5 & D_1C_1 & D_1C_2 \end{bmatrix}. \tag{77}$$

The matrix $P_2TP_2^T$ is invertible if and only if its bottom-right $2N\times 2N$ corner is invertible. Let I_N denote the $N\times N$ identity matrix; then the bottom right corner of $P_2TP_2^T$ factorizes as

$$\begin{bmatrix} D_1 & 0 \\ 0 & D_1 \end{bmatrix}\begin{bmatrix} I_N & -I_N \\ I_N & I_N \end{bmatrix}\begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix}, \tag{78}$$

which is clearly invertible since the matrices D_1, C_1, C_2 are invertible.

Finally, using all of these results, it follows that the matrix $\boldsymbol{B}$ is invertible for all $(a,0)=R_1$. Since all of the quantities involved are analytic, on every compact subset of S_N^1, we conclude that the matrix $\boldsymbol{B}$ is invertible in an open neighborhood $R_1\subset \widetilde{S}_N^1\subset S_N^1$. By construction $|\widetilde{S}_N^1|>0$.

Appendix B. Analysis of $\mathcal{K}_{\mathbf{neu}}$

All the proofs for the analysis of $\mathcal{K}_{\text{neu}}$ are similar to the corresponding proofs of $\mathcal{K}_{\text{dir}}$. We only present the analogs of Lemmas A.1 and A.2.

In the following lemma we present the directional derivative of a single-layer potential defined on straight line segment with density s^{β} at an arbitrary point near the boundary. Here s is the distance along the segment, at an arbitrary point near the boundary.

Lemma B.1. *Suppose that* Γ *is an edge of unit length oriented along an angle* $\pi\theta$, *parametrized by* $s(\cos(\theta), \sin(\theta))$, $0 < s < 1$. *Suppose* $\boldsymbol{x} = t(\cos(\theta+\theta_0), \sin(\theta+\theta_0))$ *and* $\boldsymbol{n} = (-\sin(\theta+\theta_0), \cos(\theta+\theta_0))$ (*see Figure 9*) *where* $0 < t < 1$, *and* $\boldsymbol{x} \notin \Gamma$. *Suppose that* $\sigma(s) = s^{\beta-1}$ *for* $0 < s < 1$, *where* $\beta \geq \frac{1}{2}$. *If* β *is not an integer, then*

$$\nabla\mathcal{S}[\sigma](\boldsymbol{x}) \cdot \boldsymbol{n} = -\frac{\sin(\beta(\pi - \theta_0))}{2\sin(\pi\beta)} t^{\beta-1} - \frac{1}{2\pi}\sum_{k=1}^{\infty} \frac{\sin(k\theta_0)}{\beta - k} t^{k-1}. \tag{79}$$

If $\beta = m$ *is an integer, then*

$$\nabla\mathcal{S}_{\Gamma}[\sigma](\boldsymbol{x}) \cdot \boldsymbol{n} = -\frac{(\pi - \theta_0)\cos(m\theta_0)}{2\pi} t^{m-1} + \frac{\sin(m\theta_0)}{2\pi} t^{m-1}\log(t) - \frac{1}{2\pi}\sum_{\substack{k=1\\k\neq m}}^{\infty} \frac{\sin(\pi k\theta_0)}{m - k} t^{k-1}. \tag{80}$$

In the following lemma, we compute the potential at a triple junction with angles $\pi\theta_1, \pi\theta_2, \pi\theta_3$, and material parameters $\boldsymbol{d} = (d_{(1,2)}, d_{(2,3)}, d_{(3,1)})$ (see Figure 2).

Lemma B.2. *Consider the geometry setup of the single vertex problem presented in Section 3. For a constant vector* $\boldsymbol{v} \in \mathbb{R}^3$, *suppose that the density on the edges is of the form*

$$\sigma = \begin{bmatrix} \sigma_{1,2} \\ \sigma_{2,3} \\ \sigma_{3,1} \end{bmatrix} = \boldsymbol{w} t^{\beta-1}. \tag{81}$$

If β *is not an integer, then*

$$\mathcal{K}_{\text{dir}}[\sigma] = -\frac{1}{2\sin(\pi\beta)} \mathcal{A}_{\text{neu}}(d_{3,1}, d_{1,2}, \beta)\boldsymbol{w} t^{\beta} - \sum_{k=1}^{\infty} \frac{1}{\beta - k} \boldsymbol{C}(\boldsymbol{d}, k)\boldsymbol{w} t^{k-1}, \tag{82}$$

where $\mathcal{A}_{\text{neu}}$ *is defined in (25), and* $\boldsymbol{C}(\boldsymbol{d}, k)$ *is defined in (48). If* $\beta = m$ *is an integer, then*

$$\mathcal{K}_{\text{neu}}[\sigma] = -\frac{(-1)^m}{2\pi} \mathcal{A}_{\text{neu}}(d_{3,1}, d_{1,2}, m)\boldsymbol{w} t^m \log(t) - \sum_{\substack{k=1\\k\neq m}}^{\infty} \frac{1}{m-k} \boldsymbol{C}(\boldsymbol{d}, k)\boldsymbol{w} t^{k-1} - \boldsymbol{C}_{\text{diag}}(\boldsymbol{d}, m)\boldsymbol{w} t^{m-1}, \tag{83}$$

where $\boldsymbol{C}_{\text{diag}}$ *is defined in (50).*

Proof. The result follows from repeated application of Lemma B.1 for computing $\mathcal{D}^*_{(l,m):(i,j)}\sigma_{(i,j)}$. □

The proof of Theorem 3.5 then follows immediately from Lemma B.2.

In the following lemma, we prove that $\beta_{i,j}, S_{i,j}$, $i = 1, 2, \ldots$, $j = 0, 1, 2$, defined in Appendix A.1 satisfy $\beta_{i,j}(a, b) > \frac{1}{2}$ for all (a, b) in an open subset $T_{i,j} \subset S_{i,j}$.

Lemma B.3. *Suppose that* $\beta_{i,j}$, $S_{i,j}$, $i = 1, 2, \ldots,$ $j = 0, 1, 2,$ *are as defined in Appendix A.1. Then there exists an open subset* $T_{i,j} \subset S_{i,j}$ *such that* $\beta_{i,j}(a, b) > \frac{1}{2}$ *for all* $(a, b) \in T_{i,j}$. *Moreover for any* $N > 0$, $\bigcap_{i=1}^{N+1} \bigcap_{j=0}^{2} |T_{i,j}| > 0$.

Proof. Since $\beta_{i,0} = i$, the statement is trivially true with $T_{i,j} = (-1, 1)^2$. Since $\beta_{i,j} = z_i^{\pm}(\delta, \theta)$ on $a = 0$, $b = 0$, or $c = 0$, for appropriate parameters δ, θ, we conclude that $\beta_{i,j} > \frac{1}{2}$, on $a = 0$, $b = 0$, or $c = 0$, for $i = 1, 2, \ldots,$ $j = 1, 2$. Since $\beta_{i,j}$ are analytic on $S_{i,j}$, there exists an open subset containing the segments $a = 0$, $b = 0$, or $c = 0$, which we denote by $T_{i,j}$, such that $\beta_{i,j}(a, b) > \frac{1}{2}$ for all $(a, b) \in T_{i,j}$. Since each $T_{i,j}$ is an open subset of $(-1, 1)^2$ containing $\bigcup_{k=1}^{6} R_k$, we conclude that $\left|\bigcap_{i=1}^{N+1} \bigcap_{j=0}^{2} T_{i,j}\right| > 0$. □

Acknowledgments

The authors would like to thank Alex Barnett, Charles Epstein, Leslie Greengard, Vladimir Rokhlin, and Kirill Serkh for many useful discussions.

References

[Afanas'ev et al. 2004] V. P. Afanas'ev, A. A. Kostin, and V. A. Kuptsov, "On computation of electrostatic field strength at triple junctions", pp. 675–677 in *Proceedings of XXIst International Symposium on Discharges and Electrical Insulation in Vacuum* (Yalta, 2004), vol. 2, 2004.

[Barnett et al. 2015] A. Barnett, B. Wu, and S. Veerapaneni, "Spectrally accurate quadratures for evaluation of layer potentials close to the boundary for the 2D Stokes and Laplace equations", *SIAM J. Sci. Comput.* **37**:4 (2015), B519–B542. MR Zbl

[Berggren et al. 2001] S. A. Berggren, D. Lukkassen, A. Meidell, and L. Simula, "A new method for numerical solution of checkerboard fields", *J. Appl. Math.* **1**:4 (2001), 157–173. MR Zbl

[Bremer and Rokhlin 2010] J. Bremer and V. Rokhlin, "Efficient discretization of Laplace boundary integral equations on polygonal domains", *J. Comput. Phys.* **229**:7 (2010), 2507–2525. MR Zbl

[Bremer et al. 2010] J. Bremer, V. Rokhlin, and I. Sammis, "Universal quadratures for boundary integral equations on two-dimensional domains with corners", *J. Comput. Phys.* **229**:22 (2010), 8259–8280. MR Zbl

[Bremer et al. 2015] J. Bremer, A. Gillman, and P.-G. Martinsson, "A high-order accurate accelerated direct solver for acoustic scattering from surfaces", *BIT* **55**:2 (2015), 367–397. MR Zbl

[Chung et al. 2005] M. S. Chung, T. S. Choi, and B.-G. Yoon, "Theoretical analysis of the field enhancement in a two-dimensional triple junction", *Appl. Surface Sci.* **251**:1-4 (2005), 177–181.

[Claeys et al. 2015] X. Claeys, R. Hiptmair, and E. Spindler, "A second-kind Galerkin boundary element method for scattering at composite objects", *BIT* **55**:1 (2015), 33–57. MR Zbl

[Claeys et al. 2017] X. Claeys, R. Hiptmair, and E. Spindler, "Second kind boundary integral equation for multi-subdomain diffusion problems", *Adv. Comput. Math.* **43**:5 (2017), 1075–1101. MR Zbl

[Craster and Obnosov 2004] R. V. Craster and Y. V. Obnosov, "A three-phase tessellation: solution and effective properties", *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **460**:2044 (2004), 1017–1037. MR Zbl

[Fel et al. 2000] L. G. Fel, V. S. Machavariani, and D. J. Bergman, "Isotropic conductivity of two-dimensional three-component symmetric composites", *J. Phys. A* **33**:38 (2000), 6669–6681. MR Zbl

[Fredkin and Mayergoyz 2003] D. R. Fredkin and I. D. Mayergoyz, "Resonant behavior of dielectric objects (electrostatic resonances)", *Phys. Rev. Lett.* **91**:25 (2003), art. id. 253902.

[Greengard and Lee 2012] L. Greengard and J.-Y. Lee, "Stable and accurate integral equation methods for scattering problems with multiple material interfaces in two dimensions", *J. Comput. Phys.* **231**:6 (2012), 2389–2395. MR Zbl

[Greengard and Rokhlin 1986] L. Greengard and V. Rokhlin, "A fast algorithm for particle simulations", technical report YALEU/DCS/RR-459, Yale University, 1986, available at https://cpsc.yale.edu/sites/default/files/files/tr459.pdf.

[Greengard et al. 2009] L. Greengard, D. Gueyffier, P.-G. Martinsson, and V. Rokhlin, "Fast direct solvers for integral equations in complex three-dimensional domains", *Acta Numer.* **18** (2009), 243–275. MR Zbl

[Helsing 1991] J. Helsing, "Transport properties of two-dimensional tilings with corners", *Phys. Rev. B* **44** (1991), 11677–11682.

[Helsing 2011] J. Helsing, "The effective conductivity of random checkerboards", *J. Comput. Phys.* **230**:4 (2011), 1171–1181. MR Zbl

[Helsing 2013] J. Helsing, "Solving integral equations on piecewise smooth boundaries using the RCIP method: a tutorial", *Abstr. Appl. Anal.* **2013**:Special Issue (2013), art. id. 938167. MR Zbl

[Helsing and Ojala 2008a] J. Helsing and R. Ojala, "Corner singularities for elliptic problems: integral equations, graded meshes, quadrature, and compressed inverse preconditioning", *J. Comput. Phys.* **227**:20 (2008), 8820–8840. MR Zbl

[Helsing and Ojala 2008b] J. Helsing and R. Ojala, "On the evaluation of layer potentials close to their sources", *J. Comput. Phys.* **227**:5 (2008), 2899–2921. MR Zbl

[Hoskins 2018] J. Hoskins, "On the numerical solution of transmission problems for the Laplace equation on polygonal domains", technical report YALEU/DCS/TR-1543, Yale University, 2018, available at https://cpsc.yale.edu/sites/default/files/files/tr1543.pdf.

[Hoskins and Rachh 2020] J. Hoskins and M. Rachh, "On the discretization of Laplace's equation with Neumann boundary conditions on polygonal domains", preprint, 2020. arXiv

[Hoskins et al. 2017] J. Hoskins, V. Rokhlin, and K. Serkh, "On the numerical solution of elliptic partial differential equations on polygonal domains", technical report YALEU/DCS/TR-1538, Yale University, 2017, available at https://cpsc.yale.edu/sites/default/files/files/tr1538.pdf.

[Keller 1987] J. B. Keller, "Effective conductivity of periodic composites composed of two very unequal conductors", *J. Math. Phys.* **28**:10 (1987), 2516–2520. MR Zbl

[Klöckner et al. 2013] A. Klöckner, A. Barnett, L. Greengard, and M. O'Neil, "Quadrature by expansion: a new method for the evaluation of layer potentials", *J. Comput. Phys.* **252** (2013), 332–349. MR

[Lee 2008] M. J. Lee, "Pseudo-random-number generators and the square site percolation threshold", *Phys. Rev. E* **78**:3 (2008), art. id. 031131.

[McLean 2000] W. McLean, *Strongly elliptic systems and boundary integral equations*, Cambridge University Press, 2000. MR Zbl

[Milton 2002] G. W. Milton, *The theory of composites*, Cambridge Monographs on Applied and Computational Mathematics **6**, Cambridge University Press, 2002. MR Zbl

[Milton et al. 1981] G. W. Milton, R. C. McPhedran, and D. R. McKenzie, "Transport properties of arrays of intersecting cylinders", *Appl. Phys.* **25**:1 (1981), 23–30.

[Ovchinnikov 2004] S. Ovchinnikov, "Conductivity of a periodic two-component system of rhombic type", *J. Exper. Theoret. Phys.* **98**:1 (2004), 162–169.

[Schächter 1998] L. Schächter, "Analytic expression for triple-point electron emission from an ideal edge", *Appl. Phys. Lett.* **72**:4 (1998), 421–423.

[Serkh 2019] K. Serkh, "On the solution of elliptic partial differential equations on regions with corners, II: Detailed analysis", *Appl. Comput. Harmon. Anal.* **46**:2 (2019), 250–287. MR Zbl

[Serkh and Rokhlin 2016] K. Serkh and V. Rokhlin, "On the solution of elliptic partial differential equations on regions with corners", *J. Comput. Phys.* **305** (2016), 150–171. MR Zbl

[Techaumnat et al. 2002] B. Techaumnat, S. Hamada, and T. Takuma, "Effect of conductivity in triple-junction problems", *J. Electrostatics* **56**:1 (2002), 67–76.

[Tully et al. 2007] L. K. Tully, A. D. White, D. A. Goerz, J. B. Javedani, and T. L. Houck, "Electrostatic modeling of vacuum insulator triple junctions", pp. 732–732 in *2007 IEEE 34th International Conference on Plasma Science* (Albuquerque, NM, 2007), 2007.

[Tuncer et al. 2002] E. Tuncer, Y. V. Serdyuk, and S. M. Gubanski, "Dielectric mixtures: electrical properties and modeling", *IEEE Trans. Dielec. Elec. Insul.* **9**:5 (2002), 809–828.

Received 1 Aug 2019. Revised 12 Feb 2020. Accepted 16 Mar 2020.

JEREMY HOSKINS: jeremy.hoskins@yale.edu
Applied Mathematics Program, Yale University, New Haven, CT, United States

MANAS RACHH: mrachh@flatironinstitute.org
Center for Computational Mathematics, Flatiron Institute, New York, NY, United States

PURE and APPLIED
ANALYSIS

Vol. 2, No. 2, 2020
https://doi.org/10.2140/paa.2020.2.477

ENHANCED CONVERGENCE RATES AND ASYMPTOTICS FOR A DISPERSIVE BOUSSINESQ-TYPE SYSTEM WITH LARGE ILL-PREPARED DATA

FRÉDÉRIC CHARVE

We obtain, for a stratified, rotating, incompressible Navier–Stokes system, generalized asymptotics as the Rossby number ε goes to zero (without assumptions on the diffusion coefficients). For ill-prepared, less regular initial data with large blowing-up norm in terms of ε, we show global well-posedness and improved convergence rates (as a power of ε) towards the solution of the limit system, called the 3-dimensional quasigeostrophic system. Aiming for significant improvements required us to avoid as much as possible resorting to classical energy estimates involving oscillations. Our approach relies on the use of structures and symmetries of the limit system, and of highly improved Strichartz-type estimates.

1. Introduction

1.1. ***Geophysical fluids.*** The primitive system (also called primitive equations; see for example [Chemin 1997; Babin et al. 2001]) is a rotating Boussinesq-type system used to describe geophysical fluids located at the surface of the Earth (in a large physical extent) under the assumption that the vertical motion is much smaller than the horizontal one. Two phenomena exert a crucial influence on geophysical fluids: the Coriolis force induced by the rotation of the Earth around its axis and the vertical stratification of the density induced by gravity. The former induces a vertical rigidity in the fluid velocity as described by the Taylor–Proudman theorem, and the latter induces a horizontal rigidity to the fluid density: heavier masses lie under lighter ones.

In order to measure the importance of these two concurrent phenomena, physicists defined two numbers: the Rossby number Ro and the Froude number Fr. We refer to the introduction of [Charve 2005; 2018a; 2018b] for more details and to [Bourgeois and Beale 1994; Cushman-Roisin 1994; Bougeault and Sadourny 1998; Pedlosky 1979] for an in-depth presentation.

The smaller these numbers, the more important these two phenomena become and we will consider the primitive equations in the whole space, under the Boussinesq approximation and when both phenomena are of the same scale, i.e., $\mathrm{Ro} = \varepsilon$ and $\mathrm{Fr} = \varepsilon F$ with $F > 0$. In what follows ε will be called the Rossby number and F the Froude number. The system is then written as follows (we refer to [Chemin 1997;

MSC2010: 35B40, 35Q35, 35Q86, 76U05.
Keywords: stratified rotating Navier–Stokes system, 3-dimensional quasigeostrophic system, Strichartz estimates.

Babin et al. 2001] for the model):

$$\begin{cases} \partial_t U_\varepsilon + v_\varepsilon \cdot \nabla U_\varepsilon - LU_\varepsilon + \frac{1}{\varepsilon}\mathcal{A}U_\varepsilon = \frac{1}{\varepsilon}(-\nabla\Phi_\varepsilon, 0), \\ \operatorname{div} v_\varepsilon = 0, \\ U_{\varepsilon|t=0} = U_{0,\varepsilon}. \end{cases} \tag{PE$_\varepsilon$}$$

The unknowns are $U_\varepsilon = (v_\varepsilon, \theta_\varepsilon) = (v_\varepsilon^1, v_\varepsilon^2, v_\varepsilon^3, \theta_\varepsilon)$ (where v_ε denotes the velocity of the fluid and θ_ε the scalar potential temperature) and Φ_ε which is called the geopotential and gathers the pressure term and centrifugal force. The diffusion operator L is defined by

$$LU_\varepsilon \stackrel{\text{def}}{=} (\nu\Delta v_\varepsilon, \nu'\Delta\theta_\varepsilon),$$

where $\nu, \nu' > 0$ are the kinematic viscosity and the thermal diffusivity. The matrix $\mathcal{A}$ is defined by

$$\mathcal{A} \stackrel{\text{def}}{=} \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & F^{-1} \\ 0 & 0 & -F^{-1} & 0 \end{pmatrix}.$$

We will also make precise later the properties satisfied by the sequence of initial data $U_{0,\varepsilon}$ (as ε goes to zero). Let us now state some remarks about this system (we refer to the introductions of [Charve 2005; 2018a; 2018b; Charve and Ngo 2011] for more details):

- This system generalizes the well-known rotating fluids system (see [Chemin et al. 2000; 2002; 2006]). The penalized terms (which are divided by the small parameter ε), namely $\mathcal{A}U_\varepsilon$ and the geopotential, will impose a special structure to the limit when ε goes to zero.

- As $\mathcal{A}$ is skew-symmetric, thanks to the incompressibility, any energy method (that is based on L^2 or $H^s/\dot{H}^s$-inner products) will not "see" these penalized terms and will work as for the classical incompressible Navier–Stokes system ($\mathcal{A}U \cdot U = 0$ and $(\nabla\Phi_\varepsilon, v_\varepsilon)_{H^s/\dot{H}^s} = 0$). Therefore the Leray and Fujita–Kato theorems provide global-in-time weak solutions if $U_{0,\varepsilon} \in L^2$ and local-in-time unique strong solutions if $U_{0,\varepsilon} \in \dot{H}^{\frac{1}{2}}$ (global for small initial data).

- There are two distinct regimes: $F \neq 1$ or $F = 1$. The first one features very important dispersive properties. In the second case, the operators are simpler but we cannot rely on Strichartz estimates and the methods are completely different (see [Chemin 1997; Charve 2018b]). *In the present article we focus on the case $F \neq 1$.*

1.2. ***Strong solutions.*** As explained before, thanks to the skew-symmetry of matrix $\mathcal{A}$, any computation involving L^2 or Sobolev inner products will be the same as for the Navier–Stokes system. So given the regularity of the initial data (even if some norms can blow up in ε), we can adapt the Leray and Fujita–Kato theorems as well as the classical weak-strong uniqueness results: for a fixed $\varepsilon > 0$, if $U_{0,\varepsilon} \in \dot{H}^{\frac{1}{2}}(\mathbb{R}^3)$, we denote by U_ε the unique strong solution of system (PE$_\varepsilon$), defined on $[0, T]$ for all $0 < T < T_\varepsilon^*$. In addition, if the lifespan T_ε^* is finite then we have (blow up criterion)

$$\int_0^{T_\varepsilon^*} \|\nabla U_\varepsilon(\tau)\|^2_{\dot{H}^{1/2}(\mathbb{R}^3)}\, d\tau = \infty. \tag{1-1}$$

Moreover, if $U_{0,\varepsilon} \in \dot{H}^s$ then we also can propagate the regularity as done for the Navier–Stokes system.

1.3. ***The limit system, the QG/osc decomposition.*** We are interested in the asymptotics, as the small parameter ε goes to zero. Let us recall that the limit system is a transport-diffusion system coupled with a Biot–Savart inversion law and is called the 3-dimensional quasigeostrophic system:

$$\begin{cases} \partial_t \widetilde{\Omega}_{\mathrm{QG}} + \tilde{v}_{\mathrm{QG}} \cdot \nabla \widetilde{\Omega}_{\mathrm{QG}} - \Gamma \widetilde{\Omega}_{\mathrm{QG}} = 0, \\ \widetilde{U}_{\mathrm{QG}} = (\tilde{v}_{\mathrm{QG}}, \tilde{\theta}_{\mathrm{QG}}) = (-\partial_2, \partial_1, 0, -F\partial_3)\Delta_F^{-1}\widetilde{\Omega}_{\mathrm{QG}}, \end{cases} \tag{QG}$$

where the operator Γ is defined by

$$\Gamma \stackrel{\text{def}}{=} \Delta\Delta_F^{-1}(\nu\partial_1^2 + \nu\partial_2^2 + \nu' F^2 \partial_3^2),$$

with $\Delta_F = \partial_1^2 + \partial_2^2 + F^2\partial_3^2$. Moreover we also have the relation

$$\widetilde{\Omega}_{\mathrm{QG}} = \partial_1 \widetilde{U}^2_{\mathrm{QG}} - \partial_2 \widetilde{U}^1_{\mathrm{QG}} - F\partial_3 \widetilde{U}^4_{\mathrm{QG}} = \partial_1 \tilde{v}^2_{\mathrm{QG}} - \partial_2 \tilde{v}^1_{\mathrm{QG}} - F\partial_3 \tilde{\theta}_{\mathrm{QG}}.$$

Remark 1. The operator Δ_F is a simple anisotropic Laplacian but Γ is in general a tricky nonlocal diffusion operator of order 2 (except in the case $F = 1$ where $\Delta_F = \Delta$ and $\Gamma = \nu\partial_1^2 + \nu\partial_2^2 + \nu'\partial_3^2$, or in the case $\nu = \nu'$ where $\Gamma = \nu\Delta$). We refer to [Charve 2016; 2018a] for an in-depth study of Γ in the general case (neither its Fourier kernel nor singular integral kernel have a constant sign and no classical result can be used).

This limit system is first formally derived, and then rigorously justified (see [Chemin 1997; Charve 2005]). Led by the limit system we introduce the following decomposition: for any 4-dimensional vector field $U = (v, \theta)$ we define its potential vorticity $\Omega(U)$ by

$$\Omega(U) \stackrel{\text{def}}{=} \partial_1 v^2 - \partial_2 v^1 - F\partial_3\theta$$

and its quasigeostrophic and oscillating (or oscillatory) parts by

$$U_{\mathrm{QG}} = \mathcal{Q}(U) \stackrel{\text{def}}{=} \begin{pmatrix} -\partial_2 \\ \partial_1 \\ 0 \\ -F\partial_3 \end{pmatrix} \Delta_F^{-1}\Omega(U) \quad \text{and} \quad U_{\mathrm{osc}} = \mathcal{P}(U) \stackrel{\text{def}}{=} U - U_{\mathrm{QG}}. \tag{1-2}$$

As emphasized in [Charve 2005; Charve and Ngo 2011] this is an orthogonal decomposition of 4-dimensional vector fields (similar to the Leray orthogonal decomposition into divergence-free and gradient vector fields) and if $\mathcal{Q}$ and $\mathcal{P}$ are the associated orthogonal projectors on the quasigeostrophic or oscillating fields, they satisfy (see [Chemin 1997; Charve 2004; 2005; 2018a; 2018b]):

Proposition 2. *For any function* $U = (v, \theta) \in \dot{H}^s$ (*for some* s) *we have*:

(1) *$\mathcal{P}$ and $\mathcal{Q}$ are pseudodifferential operators of order* 0.

(2) *For any* $s \in \mathbb{R}$, *we have* $(\mathcal{P}(U) \mid \mathcal{Q}(U))_{H^s/\dot{H}^s} = (\mathcal{A}U \mid \mathcal{P}(U))_{H^s/\dot{H}^s} = 0$ (*when defined*).

(3) $\mathcal{P}(U) = U \iff \mathcal{Q}(U) = 0 \iff \Omega(U) = 0$.

(4) $\mathcal{Q}(U) = U \iff \mathcal{P}(U) = 0 \iff$ *there exists a scalar function* Φ *such that* $U = (-\partial_2, \partial_1, 0, -F\partial_3)\Phi$. *Such a vector field is said to be quasigeostrophic* (*or QG*) *and is also divergence-free.*

(5) *If* $U = (v, \theta)$ *is a quasigeostrophic vector field, then* $v \cdot \nabla\Omega(U) = \Omega(v \cdot \nabla U)$ *and* $\Gamma U = \mathcal{Q}(LU)$.

(6) *Denoting by* $\mathbb{P}$ *the Leray orthogonal projector on divergence-free vector fields*, $\mathbb{P}\mathcal{P} = \mathcal{P}\mathbb{P}$ *and* $\mathbb{P}\mathcal{Q} = \mathcal{Q}\mathbb{P} = \mathcal{Q}$.

Thanks to this, system (QG) can for example be rewritten in one of the equivalent following velocity formulations:

$$\begin{cases} \partial_t \widetilde{U}_{\text{QG}} + \mathcal{Q}(\widetilde{v}_{\text{QG}} \cdot \nabla \widetilde{U}_{\text{QG}}) - \Gamma \widetilde{U}_{\text{QG}} = 0, \\ \widetilde{U}_{\text{QG}} = \mathcal{Q}(\widetilde{U}_{\text{QG}}), \quad \text{(or equivalently } \mathcal{P}(\widetilde{U}_{\text{QG}}) = 0\text{)}, \\ \widetilde{U}_{\text{QG}|t=0} = \widetilde{U}_{0,\text{QG}}, \end{cases} \tag{QG}$$

or

$$\begin{cases} \partial_t \widetilde{U}_{\text{QG}} + \widetilde{v}_{\text{QG}} \cdot \nabla \widetilde{U}_{\text{QG}} - L \widetilde{U}_{\text{QG}} = \mathcal{P}\widetilde{\Phi}_{\text{QG}}, \\ \widetilde{U}_{\text{QG}} = \mathcal{Q}(\widetilde{U}_{\text{QG}}), \\ \widetilde{U}_{\text{QG}|t=0} = \widetilde{U}_{0,\text{QG}}. \end{cases} \tag{QG}$$

Remark 3. We recall that Theorem 2 from [Charve 2004] claims that if $\widetilde{U}_{0,\text{QG}} \in H^1$ then system (QG) has a unique global solution $\widetilde{U}_{\text{QG}} \in \dot{E}^0 \cap \dot{E}^1$ (see below for the space notation). We refer to [Charve 2004; 2008] and to the next sections for more details.

Remark 4. It is natural to investigate the link between the quasigeostrophic/oscillating parts decomposition of the initial data and the asymptotics when ε goes to zero. This leads to the notion of well-prepared/ill-prepared initial data depending on whether or not the initial data is already close to the quasigeostrophic structure, i.e., when the initial oscillating part is small/large (or going to zero/blowing up as ε goes to zero). In the present article we consider large and ill-prepared initial data with very large oscillating parts depending on ε.

Going back to system (PE$_\varepsilon$), we introduce $\Omega_\varepsilon = \Omega(U_\varepsilon)$, $U_{\varepsilon,\text{QG}} = \mathcal{Q}(U_\varepsilon)$ and $U_{\varepsilon,\text{osc}} = \mathcal{P}(U_\varepsilon)$. We showed in [Charve 2005] that for an initial data in L^2 (independent of ε), the oscillating part $U_{\varepsilon,\text{osc}}$ of a weak global Leray solution U_ε goes to zero in $L^2_{\text{loc}}(\mathbb{R}_+, L^q(\mathbb{R}^3))$ ($q \in]2, 6[$), and the quasigeostrophic part $U_{\varepsilon,\text{QG}}$ goes to a solution of system (QG) (with the QG-part of U_0 as initial data). This requires the study of system (A-79), and its associated matrix in the Fourier space: As explained in detail in Proposition 43, when $\nu \neq \nu'$ there are four distinct eigenvalues (it is necessary to perform frequency truncations to obtain their expression). The first one is explicit but discarded as its associated eigenvector is not divergence-free, the second one is real (and mainly linked to the quasigeostrophic part). The last two are nonreal and mainly linked to the oscillating part.

Let us denote by $\mathbb{P}_i$ ($i \in \{2, 3, 4\}$) the associated projectors. When $\nu = \nu'$, many simplifications arise (see Remark 44). Unfortunately none of these simplifications are true anymore in general (when $\nu \neq \nu'$)

but we are able to bound their operator norms and prove that the $\mathbb{P}_2$-part of an oscillating divergence-free vector field is small (we refer to [Charve 2005; 2006]; see also Proposition 43).

Moreover we are able to obtain Strichartz estimates for the last two projections $\mathbb{P}_{3+4}$. In [Charve 2005] we obtained the following Strichartz estimate upon which the main result depended:

$$\|\mathbb{P}_{3+4}\mathcal{P}_{r,R} f\|_{L^4 L^\infty} \le C_{r,R}\varepsilon^{\frac{1}{4}}(\|\mathcal{P}_{r,R} f_0\|_{L^2} + \|\mathcal{P}_{r,R} F\|_{L^2}).$$

In [Charve 2004] we focused on strong solutions. We first proved that if the initial QG-part $U_{0,\mathrm{QG}}$ is H^1 then the limit system has a unique global solution $\widetilde{U}_{\mathrm{QG}}$. We proved that if $U_{0,\mathrm{osc}} \in \dot{H}^{\frac{1}{2}}$ then U_ε is global if ε is small enough. For this we filtered some waves: we constructed a solution W_ε^T of (A-79) with a particular external force term (constructed from $\widetilde{U}_{\mathrm{QG}}$) and proved that $U_\varepsilon - \widetilde{U}_{\mathrm{QG}} - W_\varepsilon^T$ goes to zero thanks to a generalization of the previous Strichartz estimates (allowing different regularities for the external force term):

$$\|\mathbb{P}_{3+4}\mathcal{P}_{r,R} f\|_{L^2 L^\infty} \le C_{r,R}\varepsilon^{\frac{1}{4}}(\|\mathcal{P}_{r,R} f_0\|_{L^2} + \|\mathcal{P}_{r,R} F^b\|_{L^1 L^2} + \|\mathcal{P}_{r,R} F^l\|_{L^2 L^2}).$$

In [Charve 2006] we generalized the previous result for initial data depending on ε and with large oscillating part (bounded by $|\ln|\ln\varepsilon||$ in the general case and $|\ln\varepsilon|$ when $\nu = \nu'$) considering frequency truncations $\mathcal{P}_{r_\varepsilon,R_\varepsilon}$ with radii depending on ε allowing us to exhibit explicit convergence rates. In this work we distinguished the case $\nu = \nu'$ for which we were able to produce Strichartz estimates *without* frequency truncations in *inhomogeneous* spaces:

$$\|W_\varepsilon\|_{L^2 B^s_{\infty,q}} \le C\varepsilon^{\frac{1}{8}}(\|f_0\|_{B^{s+3/4}_{2,q}} + \|G\|_{L^1(B^{s+3/4}_{2,q})}).$$

In the second part of [Charve 2006], inspired by the work of Dutrifoy [2004] on vortex patches in the inviscid case and by the work of Hmidi [2005] for Navier–Stokes vortex patches, we investigated the case of initial potential vorticity which is a regularized patch, and very large initial oscillating part (regular but bounded by a negative power of ε) when $\nu = \nu'$. This work was recently generalized in the case $\nu \neq \nu'$ in [Charve 2016; 2018a] where we deeply studied the limit quasigeostrophic operator Γ which is nonlocal and nonradial. In this setting, the fact that $\nu \neq \nu'$ highly complicates every computation.

Let us also mention that in [Charve 2008] we obtained global existence when the initial QG-part is only $H^{\frac{1}{2}+\eta}$. This required real interpolation methods (inspired by [Gallagher and Planchon 2002]) in order to obtain economic estimates for the limit system (see (1-12)). In [Charve and Ngo 2011] with V. S. Ngo we studied the asymptotics in the case of evanescent viscosities (as a power of ε) and for simplified oscillating initial data (as the initial QG part is zero, the limit is also zero).

Let us now give a survey of other results on this system. In the nondispersive setting $F = 1$ there are few works: let us mention the seminal work [Chemin 1997] (that we recently generalized in [Charve 2018b]) and [Iftimie 1999a] in the inviscid case.

In [Koba et al. 2012] the authors distinguish the rotation and stratification effects, in the case $\nu = \nu'$ for initial data in $\dot{H}^{\frac{1}{2}} \cap \dot{H}^1$ and for a special condition $\partial_2 u_0^1 - \partial_1 u_0^2 = 0$ (the initial potential vorticity only depends on the temperature), and they obtain existence of a unique global solution to (PE$_\varepsilon$) in $\mathcal{C}(\mathbb{R}_0, \dot{H}^1)$

for strong enough rotation and stratification. If the initial data is small in $\dot{H}^{\frac{1}{2}}$ they manage to obtain that $\nabla U_\varepsilon \in L^2 \dot{H}^{\frac{1}{2}}$.

Lee and Takada [2017] studied global well-posedness in the case of stratification only (no rotational effects) when $\nu = \nu'$ and for large initial oscillating part (independent of ε). They first give global existence of a unique mild solution in $L^4(\mathbb{R}_+, \dot{W}^{\frac{1}{2},3}(\mathbb{R}^3))$ for large initial oscillating part in $\dot{H}^s$ ($s \in \left]\frac{1}{2}, \frac{5}{8}\right]$, there is a kind of smallness condition; see Remark 17) and small QG-part in $\dot{H}^{\frac{1}{2}}$. Then they show global well-posedness in the case $s = \frac{1}{2}$ and for any initial oscillating part and small QG-part, of a unique mild solution in $\mathcal{C}(\mathbb{R}_+, \dot{H}^{\frac{1}{2}}) \cap L^4(\mathbb{R}_+, \dot{W}^{\frac{1}{2},3}(\mathbb{R}^3))$.

These results are adapted to the primitive system in [Iwabuchi et al. 2017]. Iwabuchi, Mahalov and Takada focused on the case $\nu = \nu'$ and obtained (through stationary phase methods) the following Strichartz estimates, which we state with our notation:

Proposition 5 [Iwabuchi et al. 2017, Theorem 1.1 and Corollary 1.2]. *Assume $F \neq 1$. If $r \in]2,4[$ and $p \in]2,\infty[\cap \left[1/\left(2\left(\frac{1}{2}-\frac{1}{r}\right)\right), 2/\left(3\left(\frac{1}{2}-\frac{1}{r}\right)\right)\right]$, there exists a constant $C = C_{F,\nu,p,r}$ such that, if f solves the homogeneous (A-79),*

$$\|f\|_{L^p(\mathbb{R}_+,L^r)} \leq C\varepsilon^{\frac{1}{p}-\frac{3}{2}\left(\frac{1}{2}-\frac{1}{r}\right)} \|f_0\|_{L^2}.$$

If $s \in \left]\frac{1}{2}, \frac{5}{8}\right]$, there exists a constant $C = C(F, s, \nu)$ such that

$$\|f\|_{L^4(\mathbb{R}_+,\dot{W}^{s,6/(1+2s)})} \leq C\varepsilon^{\frac{1}{2}\left(s-\frac{1}{2}\right)} \|f_0\|_{\dot{H}^s}.$$

From this they are able to obtain through a fixed point argument the following global well-posedness results for initial data (independent of ε) with small quasigeostrophic part (assume $\nu = \nu'$ and $F \neq 1$):

- If $s \in \left]\frac{1}{2}, \frac{5}{8}\right]$, there exist $\delta_1, \delta_2 > 0$ (depending on ν, F, s) such that for any $\varepsilon > 0$ and any initial data $U_0 = U_{0,\mathrm{QG}} + U_{0,\mathrm{osc}}$ with $(U_{0,\mathrm{QG}}, U_{0,\mathrm{osc}}) \in \dot{H}^{\frac{1}{2}} \times \dot{H}^s$ and

$$\begin{cases} \|U_{0,\mathrm{QG}}\|_{\dot{H}^{1/2}} \leq \delta_1, \\ \|U_{0,\mathrm{osc}}\|_{\dot{H}^s} \leq \delta_2 \varepsilon^{-\frac{1}{2}\left(s-\frac{1}{2}\right)}, \end{cases} \tag{1-3}$$

there exists a unique global mild solution in $L^4(\mathbb{R}_+, \dot{W}^{s,\frac{1}{3}}(\mathbb{R}^3))$.

- There exists $\delta > 0$ such that for any initial data $U_0 = U_{0,\mathrm{QG}} + U_{0,\mathrm{osc}} \in \dot{H}^{\frac{1}{2}}$ with $\|U_{0,\mathrm{QG}}\|_{\dot{H}^{1/2}} \leq \delta$, there exists $\varepsilon_0 > 0$ such that for any $0 < \varepsilon < \varepsilon_0$, system (PE$_\varepsilon$) has a unique global mild solution in $\mathcal{C}(\mathbb{R}_+, \dot{H}^{\frac{1}{2}}) \cap L^4(\mathbb{R}_+, \dot{W}^{s,\frac{1}{3}}(\mathbb{R}^3))$.

Let us also mention works in the periodic case where resonances have to be studied (see for example [Gallagher 1998; Ngo 2009; Ngo and Scrobogna 2018; Scrobogna 2018]), in the rotating fluids system case (see [Chemin et al. 2000; 2002; 2006; Giga et al. 2008; Hieber and Shibata 2010; Koh et al. 2014a]) or in the inviscid case (see [Dutrifoy 2004; 2005; Koh et al. 2014b; Takada 2016; Widmayer 2018]).

In the present article we wish to generalize our results from [Charve 2004; 2006; 2008] and motivated by the very interesting results in [Iwabuchi et al. 2017] we want to obtain full asymptotics (as in [Charve and Ngo 2011; Charve 2018a]) for very large ill-prepared initial data (less regular, depending on ε and bounded by a negative power of ε). In our work we will provide global well-posedness results but also precise convergence rates as ε goes to zero. We also generalize [Iwabuchi et al. 2017] in the sense that we consider initial data with *large quasigeostrophic part* (with low-frequency assumptions) and provide solutions in *homogeneous* energy spaces $\dot{E}^s$ both in the particular case $\nu = \nu'$ *and* in the general case $\nu \neq \nu'$. Let us also mention that our methods closely rely on the special structures and properties of the 3-dimensional quasigeostrophic system.

1.3.1. *Statement of the results.* We will consider general ill-prepared initial data $U_{0,\varepsilon} = U_{0,\varepsilon,\mathrm{osc}} + U_{0,\varepsilon,\mathrm{QG}}$, whose QG-part converges to some $\widetilde{U}_{0,\mathrm{QG}}$ (without any smallness condition), and whose oscillating part is very large (see below for details).

The aim of the present article is to generalize Theorem 3 from [Charve 2004], Theorems 1.2 and 1.3 from [Charve 2006] and Theorem 4 from [Charve 2008] with the least possible extra regularity for the initial data and the biggest possible blowing-up initial oscillatory part (as a negative power of ε). The energy methods used in [Charve 2004; 2005; 2006] would only allow at best an initial blow-up of $U_{0,\varepsilon,\mathrm{osc}}$ as $|\ln\varepsilon|^\beta$. Indeed, these methods require the use of energy estimates for the oscillations W_ε and W_ε^T and produce large terms involving $\exp(\|U_{0,\varepsilon,\mathrm{osc}}\|^2)$ that can only be balanced thanks to ε^γ provided by the Strichartz estimates. We need to change our point of view and try to not resort to energy estimates for these oscillations. This will require us to make more flexible dispersive estimates so that the oscillations can be estimated with minimal use of their energy (the only term where it was unavoidable is F_8; see below for details). We will here state only the new results. Let us define (in the whole space $\mathbb{R}^3$) the family of spaces $\dot{E}_T^s$ for $s \in \mathbb{R}$,

$$\dot{E}_T^s = \mathcal{C}_T(\dot{H}^s) \cap L_T^2(\dot{H}^{s+1}),$$

endowed with the following norm (where we define $\nu_0 = \min(\nu, \nu')$; see the Appendix for other notation):

$$\|f\|_{\dot{E}_T^s}^2 \overset{\text{def}}{=} \|f\|_{L_T^\infty \dot{H}^s}^2 + \nu_0 \int_0^T \|f(\tau)\|_{\dot{H}^{s+1}}^2 \, d\tau.$$

When $T = \infty$ we write $\dot{E}^s$ and the corresponding norm is over $\mathbb{R}_+$ in time. Let us now state the main result of this article (we *do not* assume $\nu = \nu'$).

Theorem 6. *Assume $F \neq 1$. For any $\mathbb{C}_0 \geq 1$, $\delta \in \left]0, \frac{1}{10}\right]$, and $\alpha_0 > 0$, there exist five constants $\varepsilon_0, \eta, \mathbb{B}_0, \kappa, \beta > 0$ (depending on $F, \nu, \nu', \mathbb{C}_0, \alpha_0$) such that for all $\varepsilon \in]0, \varepsilon_0]$ and all divergence-free initial data $U_{0,\varepsilon} = U_{0,\varepsilon,\mathrm{QG}} + U_{0,\varepsilon,\mathrm{osc}}$ satisfying*

(1) *$U_{0,\varepsilon,\mathrm{QG}}$ converges towards some quasigeostrophic vector field $\widetilde{U}_{0,\mathrm{QG}} \in H^{\frac{1}{2}+\delta}$ with*

$$\begin{cases} \|U_{0,\varepsilon,\mathrm{QG}} - \widetilde{U}_{0,\mathrm{QG}}\|_{H^{1/2+\delta}} \leq \mathbb{C}_0 \varepsilon^{\alpha_0}, \\ \|\widetilde{U}_{0,\mathrm{QG}}\|_{H^{1/2+\delta}} \leq \mathbb{C}_0, \end{cases} \tag{1-4}$$

(2) $\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{F}_\delta} \leq \mathbb{C}_0 \varepsilon^{-\kappa\delta}$, *where the space $\dot{F}_\delta$ is defined as follows* $\left(q = \frac{2}{1+\delta}\right)$:

$$\dot{F}_\delta = \begin{cases} \dot{H}^{\frac{1}{2}-\delta} \cap \dot{H}^{\frac{1}{2}+\delta} & \text{if } \nu = \nu', \\ \dot{B}_{q,q}^{\frac{1}{2}} \cap \dot{H}^{\frac{1}{2}+\delta} & \text{if } \nu \neq \nu', \end{cases}$$

then system (QG) has a unique global solution $\widetilde{U}_{\mathrm{QG}} \in \dot{E}^0 \cap \dot{E}^{\frac{1}{2}+\delta}$, and system (PE$_\varepsilon$) has a unique global solution $U_\varepsilon \in \dot{E}^s$ for all $s \in \left[\frac{1}{2} - \eta\delta, \frac{1}{2} + \eta\delta\right]$, which converges towards $\widetilde{U}_{\mathrm{QG}}$ with the estimate

$$\|U_\varepsilon - \widetilde{U}_{\mathrm{QG}}\|_{L^2 L^\infty} \leq \mathbb{B}_0 \varepsilon^{\min(\alpha_0, \delta\beta)}.$$

Remark 7. In the general case, κ is small $\left(\text{less than } \frac{1}{4000}\right)$, whereas in the case $\nu = \nu'$, $\kappa < \frac{1}{2}$ (and is as close to $\frac{1}{2}$ as we want). We refer to the next section for a more precise statement of this theorem.

Remark 8. It is interesting to adapt these results to the case with only stratification.

1.4. ***Precise statement of the main results.*** This section is devoted to giving the precise statement of Theorem 6, which will be split into two formulations based on whether we have $\nu = \nu'$ or $\nu \neq \nu'$. This statement requires us to introduce auxiliary systems, which is the subject of the first two subsections, and state additional regularity properties for the solution of the limit system (we refer to the third subsection). Then we will state the results we will prove in this article.

1.4.1. *Auxiliary systems in the general case $\nu \neq \nu'$.*

Remark 9. In what follows, we will systematically write, for $f : \mathbb{R}^3 \to \mathbb{R}^4$, $f \cdot \nabla f = \sum_{i=1}^{3} f_i \partial_i f$.

Following [Charve 2004] we rewrite the primitive system, projecting onto the divergence-free vector fields ($\mathbb{P}$ is the classical Leray projector):

$$\begin{cases} \partial_t U_\varepsilon - L U_\varepsilon + \frac{1}{\varepsilon} \mathbb{P} \mathcal{A} U_\varepsilon = -\mathbb{P}(U_\varepsilon \cdot \nabla U_\varepsilon). \\ U_{\varepsilon|t=0} = U_{0,\varepsilon}. \end{cases} \tag{1-5}$$

Notice that we can rewrite (QG) as follows (we also refer to [Charve 2004] where it was first used):

$$\begin{cases} \partial_t \widetilde{U}_{\mathrm{QG}} - L U_\varepsilon + \frac{1}{\varepsilon} \mathbb{P} \mathcal{A} \widetilde{U}_{\mathrm{QG}} = -\mathbb{P}(\widetilde{U}_{\mathrm{QG}} \cdot \nabla \widetilde{U}_{\mathrm{QG}}) + G, \\ \widetilde{U}_{QG|t=0} = \widetilde{U}_{0,\mathrm{QG}}. \end{cases} \tag{QG}$$

where

$$G = G^b + G^l \overset{\text{def}}{=} \mathbb{P}\mathcal{P}(\widetilde{U}_{\mathrm{QG}} \cdot \nabla \widetilde{U}_{\mathrm{QG}}) - F(\nu - \nu') \Delta \Delta_F^{-2} \begin{pmatrix} -F \partial_2 \partial_3^2 \\ F \partial_1 \partial_3^2 \\ 0 \\ (\partial_1^2 + \partial_2^2) \partial_3 \end{pmatrix} \widetilde{\Omega}_{\mathrm{QG}}. \tag{1-6}$$

Remark 10. It is important to notice that G is the sum of two terms, both divergence-free and whose potential vorticity is zero, which is crucial to fully take advantage of (A-84). We refer to [Charve 2004; 2008] for more details.

As explained in [Charve 2004; 2005; 2006; 2008; 2018a; Charve and Ngo 2011], in the case $F \neq 1$ the oscillatory part enjoys dispersive properties that allow us to obtain Strichartz-type estimates. More precisely the oscillatory part satisfies system (A-79) (we refer to the Appendix for details), and in all the cited articles, we used that the frequency-truncated third and fourth projections of the oscillatory part satisfy Strichartz-type estimates as given by Proposition 46. As in [Charve 2004; 2008; Charve and Ngo 2011], in the present article we will consider some particular oscillatory terms whose existence is devoted to absorbing some constant terms in order to get the desired convergence rate for the asymptotics as ε goes to zero.

More precisely, we introduce the following linear system (we refer to the Appendix for the notation r_ε, R_ε and $\mathcal{P}_{r_\varepsilon,R_\varepsilon}$):

$$\begin{cases} \partial_t W_\varepsilon^T - LW_\varepsilon^T + \frac{1}{\varepsilon}\mathbb{P}\mathcal{A}W_\varepsilon^T = -\mathcal{P}_{r_\varepsilon,R_\varepsilon}\mathbb{P}_{3+4}G, \\ W_\varepsilon^T{}_{|t=0} = \mathcal{P}_{r_\varepsilon,R_\varepsilon}\mathbb{P}_{3+4}U_{0,\varepsilon,\mathrm{osc}}. \end{cases} \tag{1-7}$$

Remark 11. We recall that it would be useless to consider the free system: indeed the system satisfied by $U_\varepsilon - \widetilde{U}_{\mathrm{QG}}$ features G as an external force term which is independent of ε and blocks any convergence. It is then necessary to absorb a large part of this term, which is the reason why we considered such an external force term in system (1-7). In other words, W_ε^T is small due to dispersive properties, but still it allows us to "eat" a large part of G. We refer to [Charve 2004] for more details.

Finally we define $\delta_\varepsilon = U_\varepsilon - \widetilde{U}_{\mathrm{QG}} - W_\varepsilon^T$, which satisfies the following system (see [Charve 2004] for details):

$$\begin{cases} \partial_t \delta_\varepsilon - L\delta_\varepsilon + \frac{1}{\varepsilon}\mathbb{P}\mathcal{A}\delta_\varepsilon = \sum_{i=1}^8 F_i + f^b + f^l, \\ \delta_{\varepsilon|t=0} = (U_{0,\varepsilon,\mathrm{QG}} - \widetilde{U}_{0,\mathrm{QG}}) + (\mathrm{Id} - \mathcal{P}_{r_\varepsilon,R_\varepsilon})U_{0,\varepsilon,\mathrm{osc}} + \mathcal{P}_{r_\varepsilon,R_\varepsilon}\mathbb{P}_2 U_{0,\varepsilon,\mathrm{osc}}, \end{cases} \tag{1-8}$$

where we define

$$\begin{aligned} &F_1 \stackrel{\mathrm{def}}{=} -\mathbb{P}(\delta_\varepsilon\cdot\nabla\delta_\varepsilon), \quad F_2 \stackrel{\mathrm{def}}{=} -\mathbb{P}(\delta_\varepsilon\cdot\nabla\widetilde{U}_{\mathrm{QG}}), \quad F_3 \stackrel{\mathrm{def}}{=} -\mathbb{P}(\widetilde{U}_{\mathrm{QG}}\cdot\nabla\delta_\varepsilon), \quad F_4 \stackrel{\mathrm{def}}{=} -\mathbb{P}(\delta_\varepsilon\cdot\nabla W_\varepsilon^T), \\ &F_5 \stackrel{\mathrm{def}}{=} -\mathbb{P}(W_\varepsilon^T\cdot\nabla\delta_\varepsilon), \quad F_6 \stackrel{\mathrm{def}}{=} -\mathbb{P}(\widetilde{U}_{\mathrm{QG}}\cdot\nabla W_\varepsilon^T), \quad F_7 \stackrel{\mathrm{def}}{=} -\mathbb{P}(W_\varepsilon^T\cdot\nabla\widetilde{U}_{\mathrm{QG}}), \quad F_8 \stackrel{\mathrm{def}}{=} -\mathbb{P}(W_\varepsilon^T\cdot\nabla W_\varepsilon^T), \\ &\qquad\qquad f^b \stackrel{\mathrm{def}}{=} -(\mathrm{Id}-\mathcal{P}_{r_\varepsilon,R_\varepsilon})G^b - \mathcal{P}_{r_\varepsilon,R_\varepsilon}\mathbb{P}_2 G^b, \\ &\qquad\qquad f^l \stackrel{\mathrm{def}}{=} -(\mathrm{Id}-\mathcal{P}_{r_\varepsilon,R_\varepsilon})G^l - \mathcal{P}_{r_\varepsilon,R_\varepsilon}\mathbb{P}_2 G^l. \end{aligned} \tag{1-9}$$

1.4.2. *Auxiliary systems in the special case $\nu = \nu'$.* In this case, many simplifications arise in the computations of the eigenvalues and eigenvectors of system (A-79) (see Remark 44). In this case, as used in the first part of [Charve 2006], we can use the following system instead of (1-7):

$$\begin{cases} \partial_t W_\varepsilon - LW_\varepsilon + \frac{1}{\varepsilon}\mathbb{P}\mathcal{A}W_\varepsilon = -G^b, \\ W_{\varepsilon|t=0} = U_{0,\varepsilon,\mathrm{osc}}. \end{cases} \tag{1-10}$$

We will be able in the present article to provide for this system much more accurate Strichartz estimates without any frequency restrictions (generalizing the ones obtained in [Charve 2006]). If we write $\delta_\varepsilon = U_\varepsilon - \widetilde{U}_{\mathrm{QG}} - W_\varepsilon$, it satisfies the system

$$\begin{cases} \partial_t \delta_\varepsilon - L\delta_\varepsilon + \frac{1}{\varepsilon}\mathbb{P}\mathcal{A}\delta_\varepsilon = \sum_{i=1}^8 F_i, \\ \delta_{\varepsilon|t=0} = U_{0,\varepsilon,\mathrm{QG}} - \widetilde{U}_{0,\mathrm{QG}}. \end{cases} \tag{1-11}$$

Remark 12. We choose here to use the same notation as in the general case; the only difference is that W_ε^T has to be replaced by W_ε.

1.4.3. *The limit system.* Let us recall that Theorem 2 from [Charve 2004] states that when the initial data $\widetilde{U}_{0,\mathrm{QG}}$ is in the *inhomogeneous* Sobolev space H^1, system (QG) has a unique global solution $\widetilde{U}_{\mathrm{QG}} \in \dot{E}^0 \cap \dot{E}^1$; moreover there exists a constant $C = C(\delta) > 0$ such that for all $s \in [0, 1]$ and all $t \in \mathbb{R}_+$

(and denoting as usual $\nu_0 = \min(\nu, \nu') > 0$)

$$\|\widetilde{U}_{\rm QG}\|^2_{L^\infty_t \dot{H}^s} + \nu_0 \int_0^t \|\nabla \widetilde{U}_{\rm QG}(\tau)\|^2_{\dot{H}^s}\, d\tau \le C(\|\widetilde{U}_{0,\rm QG}\|^{1-s}_{L^2} \|\widetilde{U}_{0,\rm QG}\|^s_{\dot{H}^1})^2 \le C\|\widetilde{U}_{0,\rm QG}\|^2_{H^1}.$$

In [Charve 2008] we used real interpolation methods from [Gallagher and Planchon 2002] (we also refer to [Calderón 1990]) to obtain a much more accurate estimate, which allowed us to bound the energy in $\dot{E}^0 \cap \dot{E}^{\frac{1}{2}+\delta}$ only with the $H^{\frac{1}{2}+\delta}$ initial norm instead of the full H^1-norm (we refer to Lemma 2.1 in [Charve 2008]; our aim was to consider less regular initial data): for any $\delta > 0$ there exists a constant $C = C_{\delta,\nu_0} > 0$ such that for all $t \in \mathbb{R}_+$

$$\|\widetilde{U}_{\rm QG}\|^2_{L^\infty_t H^{1/2+\delta}} + \nu_0 \int_0^t \|\nabla \widetilde{U}_{\rm QG}(\tau)\|^2_{H^{1/2+\delta}}\, d\tau \le C_{\delta,\nu_0} \|\widetilde{U}_{0,\rm QG}\|^2_{H^{1/2+\delta}} \max(1, \|\widetilde{U}_{0,\rm QG}\|^{\frac{1}{\delta}}_{H^{1/2+\delta}}). \quad (1\text{-}12)$$

Remark 13. The reader may wonder why the right-hand side is not simply $C_{\delta,\nu_0}\|\widetilde{U}_{0,\rm QG}\|^{2+\frac{1}{\delta}}_{H^{1/2+\delta}}$ as stated in [Charve 2008; Gallagher and Planchon 2002]. This is the right formulation when $\|\widetilde{U}_{0,\rm QG}\|_{H^{1/2+\delta}}$ is large (in [Charve 2008] we implicitly focused on large initial QG part). When it is small, the right-hand side is even simpler: $C_{\delta,\nu_0}\|\widetilde{U}_{0,\rm QG}\|^2_{H^{1/2+\delta}}$. In the proof in [Charve 2008] of (1-12) it is crucial to use Lemma 4.3 from [Gallagher and Planchon 2002], and for this some threshold $j_0 \ge 1$ has to be defined:

- either $\|\widetilde{U}_{0,\rm QG}\|_{H^{1/2+\delta}} > \frac{2}{3}c\nu_0 2^{2\delta}$ (where c is the smallness constant from the Fujita–Kato theorem), and we can define the threshold j_0 as stated in [Charve 2008] so that the right-hand side of (1-12) is

$$C_0(1-2^{-4\delta})^{-2}\left(\frac{3}{2c\nu_0}\right)^{\frac{1}{\delta}} \|\widetilde{U}_{0,\rm QG}\|^{2+\frac{1}{\delta}}_{H^{1/2+\delta}}$$

(C_0 is a universal constant),

- or $\|\widetilde{U}_{0,\rm QG}\|_{H^{1/2+\delta}} \le \frac{2}{3}c\nu_0 2^{2\delta}$ and then we can simply choose the threshold $j_0 = 1$ and obtain (1-12) with right-hand side that can be simplified into $C_0(1-2^{-4\delta})^{-2}\|\widetilde{U}_{0,\rm QG}\|^2_{H^{1/2+\delta}}$.

In other words, the right-hand side of (1-12) is in general

$$C_0(1-2^{-4\delta})^{-2\delta}\|\widetilde{U}_{0,\rm QG}\|^2_{H^{1/2+\delta}} \max\left(1, \frac{1}{4}\left(\frac{3}{2c\nu_0}\|\widetilde{U}_{0,\rm QG}\|_{H^{1/2+\delta}}\right)^{\frac{1}{\delta}}\right).$$

Our first result is devoted to the limit system and generalizes Theorem 2 from [Charve 2004] using the precise estimates obtained in [Charve 2008]:

Theorem 14. *Let $\delta > 0$ and $\widetilde{U}_{0,\rm QG} \in H^{\frac{1}{2}+\delta}$ be a quasigeostrophic vector field (that is, $\widetilde{U}_{0,\rm QG} = \mathcal{Q}\widetilde{U}_{0,\rm QG}$). Then system (QG) has a unique global solution in $E^{\frac{1}{2}+\delta} = \dot{E}^0 \cap \dot{E}^{\frac{1}{2}+\delta}$ and the previous estimates hold true.*

1.4.4. *Statement in the case $\nu = \nu'$.*

Theorem 15. *Assume $F \ne 1$. For any $\mathbb{C}_0 \ge 1$, $\delta \in \left]0, \frac{1}{10}\right]$, $\gamma \in \left]0, \frac{\delta}{2}\right[$, and any $\alpha_0 > 0$, if we define $\eta > 0$ such that*

$$\gamma = (1-2\eta)\tfrac{\delta}{2} \quad \left(\textit{that is, } \eta = \tfrac{1}{2}\left(1 - \tfrac{2\gamma}{2}\right)\right),$$

there exist $\varepsilon_0, \mathbb{B}_0 > 0$ (*both of them depending on* $F, \nu, \mathbb{C}_0, \delta, \gamma, \alpha_0$) *such that for all* $\varepsilon \in]0, \varepsilon_0]$ *and all divergence-free initial data* $U_{0,\varepsilon} = U_{0,\varepsilon,\mathrm{QG}} + U_{0,\varepsilon,\mathrm{osc}}$ *satisfying*

(1) *there exists a quasigeostrophic vector field* $\widetilde{U}_{0,\mathrm{QG}} \in H^{\frac{1}{2}+\delta}$ *such that*

$$\begin{cases} \|U_{0,\varepsilon,\mathrm{QG}} - \widetilde{U}_{0,\mathrm{QG}}\|_{H^{1/2+\delta}} \leq \mathbb{C}_0 \varepsilon^{\alpha_0}, \\ \|\widetilde{U}_{0,\mathrm{QG}}\|_{H^{1/2+\delta}} \leq \mathbb{C}_0, \end{cases} \tag{1-13}$$

(2) $U_{0,\varepsilon,\mathrm{osc}} \in \dot{H}^{\frac{1}{2}} \cap \dot{H}^{\frac{1}{2}+\delta}$ *with* $\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^{1/2} \cap \dot{H}^{1/2+\delta}} \leq \mathbb{C}_0 \varepsilon^{-\gamma}$,

then system (PE$_\varepsilon$) *has a unique global solution* $U_\varepsilon \in \dot{E}^s$ *for all* $s \in \left[\frac{1}{2}, \frac{1}{2} + \eta\delta\right]$, *and if we define*

- $\widetilde{U}_{\mathrm{QG}}$ *as the unique global solution of* (QG) *in* $\dot{E}^0 \cap \dot{E}^{\frac{1}{2}+\delta}$,
- W_ε *as the unique global solution of* (1-10) *in* $\dot{E}^{\frac{1}{2}} \cap \dot{E}^{\frac{1}{2}+\delta}$,
- $\delta_\varepsilon = U_\varepsilon - \widetilde{U}_{\mathrm{QG}} - W_\varepsilon$,

then for all $s \in \left[\frac{1}{2}, \frac{1}{2} + \eta\delta\right]$

$$\|\delta_\varepsilon\|_{\dot{E}^s} \leq \mathbb{B}_0 \varepsilon^{\min(\alpha_0, \frac{\delta\eta}{2})}. \tag{1-14}$$

Moreover if we ask for more low-frequency regularity for the initial oscillating part, *that is*, $U_{0,\varepsilon,\mathrm{osc}} \in \dot{H}^{\frac{1}{2}-\delta} \cap \dot{H}^{\frac{1}{2}+\delta}$ *with* $\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^{1/2-\delta} \cap \dot{H}^{1/2+\delta}} \leq \mathbb{C}_0 \varepsilon^{-\gamma}$ *then* (1-14) *is true for all* $s \in \left[\frac{1}{2} - \eta\delta, \frac{1}{2} + \eta\delta\right]$ *and we also can get rid of the oscillations* W_ε *and obtain that*

$$\|U_\varepsilon - \widetilde{U}_{\mathrm{QG}}\|_{L^2 L^\infty} \leq \mathbb{B}_0 \varepsilon^{\min(\alpha_0, \frac{\delta\eta}{2})}.$$

Remark 16. Compared to Theorem 1.3 from [Charve 2006] we highly reduced the regularity of the initial data, only the quasigeostrophic part lies in a inhomogeneous space, and we allow a far greater blowup in ε for the oscillating part, keeping a satisfying convergence rate as a power of ε (in accordance with physicists) for any size of the initial quasigeostrophic part.

Remark 17. In [Lee and Takada 2017; Iwabuchi et al. 2017] there is a smallness condition for the initial quasigeostrophic part (and also for the oscillating part in some sense). Their result states there exist $\delta_{1,2} > 0$ such that for any initial data satisfying (1-3), there exists a global unique mild solution for any $\varepsilon > 0$. This has to be compared with our formulation, where we prove that for any size $\mathbb{C}_0$ and any initial data with $\|U_{0,\varepsilon,\mathrm{QG}}\| \leq \mathbb{C}_0$ and $\|U_{0,\varepsilon,\mathrm{osc}}\| \leq \mathbb{C}_0 \varepsilon^{-\gamma}$, there exists a unique global solution when $\varepsilon \leq \varepsilon_0$.

Remark 18. Compared to the assumptions in [Iwabuchi et al. 2017, Theorems 1.3 and 1.5], we reach the same regularity for the oscillating part, we ask more regularity for the initial QG-part, and we ask more low-frequency regularity for both of them (we have to assume $U_{0,\varepsilon} \in \dot{H}^{\frac{1}{2}}$ as we need to consider Fujita–Kato strong solutions)

$$\begin{cases} U_{0,\varepsilon,\mathrm{osc}} \in \dot{H}^{\frac{1}{2}} \cap \dot{H}^{\frac{1}{2}+\delta} & (\dot{H}^{\frac{1}{2}+\delta} \text{ in [Iwabuchi et al. 2017]}), \\ U_{0,\varepsilon,\mathrm{QG}} \in H^{\frac{1}{2}+\delta} & (\dot{H}^{\frac{1}{2}} \text{ in [Iwabuchi et al. 2017]}), \end{cases}$$

but *we do not ask any smallness to the initial quasigeostrophic part*, and we provide global strong solutions *in the energy spaces* $\dot{E}^s$ for any $s \in \left[\frac{1}{2} - \eta\delta, \frac{1}{2} + \eta\delta\right]$ (compared to mild solutions in $L^4(\mathbb{R}_+, \dot{W}^{\frac{1}{2},3})$).

Remark 19. At first sight our blow-up rate seems slightly less general than the one from [Iwabuchi et al. 2017] (in that work they require $\varepsilon^{\frac{\delta}{2}}\|U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^{1/2+\delta}}$ be smaller than some $\delta_2 > 0$, and in the present work, we choose any $\mathbb{C}_0$ and require $\varepsilon^{\gamma}\|U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^{1/2}\cap\dot{H}^{1/2+\delta}} \leq \mathbb{C}_0$ for any $\gamma < \frac{\delta}{2}$) but in our result we look for explicit rates of convergence as powers of ε. We refer to Remark 32 for more details.

Remark 20. We refer to Remark 48 for a comparison of the Strichartz estimates we use and the ones from [Iwabuchi et al. 2017].

1.4.5. *Statement in the general case $\nu \neq \nu'$.*

Theorem 21. *Assume $F \neq 1$. Let $\delta \in \left]0, \frac{1}{2}\right]$, $q = \frac{2}{1+\delta}$, $\alpha_0 > 0$, $m \in \left]0, \frac{1}{100}\right]$, and $M, \eta > 0$ such that*

$$0 < 2\eta \leq \frac{M}{m} \leq \frac{1}{2}\frac{1}{5+\delta},$$

and let $\gamma_0 \in \left]0, \frac{M\delta}{4}\right]$. If we define $R_\varepsilon = \varepsilon^{-M}$ and $r_\varepsilon = \varepsilon^m$ then for all $\mathbb{C}_0 \geq 1$ there exist ε_0, $\mathbb{B}_0$ (all of them depending on $F, \nu, \nu', \mathbb{C}_0, \delta, \gamma, \alpha_0$) such that for all initial data $U_{0,\varepsilon} = U_{0,\varepsilon,\text{osc}} + U_{0,\varepsilon,\text{QG}}$ satisfying

(1) *there exists a quasigeostrophic vector field $\widetilde{U}_{0,\text{QG}} \in H^{\frac{1}{2}+\delta}$ such that*

$$\begin{cases} \|U_{0,\varepsilon,\text{QG}} - \widetilde{U}_{0,\text{QG}}\|_{H^{1/2+\delta}} \leq \mathbb{C}_0\varepsilon^{\alpha_0}, \\ \|\widetilde{U}_{0,\text{QG}}\|_{H^{1/2+\delta}} \leq \mathbb{C}_0, \end{cases} \tag{1-15}$$

(2) $U_{0,\varepsilon,\text{osc}} \in \dot{B}^{\frac{1}{2}}_{q,q} \cap \dot{H}^{\frac{1}{2}+\delta}$ *with* $\|U_{0,\varepsilon,\text{osc}}\|_{\dot{B}^{1/2}_{q,q}\cap\dot{H}^{1/2+\delta}} \leq \mathbb{C}_0\varepsilon^{-\gamma}$,

system (PE$_\varepsilon$) has a unique global solution $U_\varepsilon \in \dot{E}^s$ for all $s \in \left[\frac{1}{2} - \eta\delta, \frac{1}{2} + \eta\delta\right]$. Moreover, with the same notation as in Theorem 14 (replacing W_ε by W_ε^T, which involves m, M),

$$\|\delta_\varepsilon\|_{\dot{E}^s} \leq \mathbb{B}_0\varepsilon^{\min(\alpha_0, \frac{M\delta}{4})}, \tag{1-16}$$

and finally, thanks to the Strichartz estimates, we can get rid of the oscillations W_ε^T and obtain

$$\|U_\varepsilon - \widetilde{U}_{\text{QG}}\|_{L^2(\mathbb{R}_+, L^\infty)} \leq \mathbb{B}_0\varepsilon^{\min(\alpha_0, \frac{M\delta}{4})}.$$

Remark 22. This generalizes the first result from [Charve 2006]: in the present work we reduced the assumptions on high and low frequencies for the initial oscillating part, and the choice for r_ε and R_ε now correctly fits the power of ε provided by the Strichartz estimates, which produces a convergence rate as a power of ε without any assumption on the viscosities.

Remark 23. The low-frequencies assumption $U_{0,\varepsilon,\text{osc}} \in \dot{B}^{\frac{1}{2}}_{q,q}$ is mainly needed to produce a positive power of ε when estimating $\|\chi(|D|/R_\varepsilon)\chi(|D_3|/r_\varepsilon)U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^s}$ (the other need is to reach regularities less than $\frac{1}{2}$), and the high-frequencies assumption $U_{0,\varepsilon,\text{osc}} \in \dot{H}^{\frac{1}{2}+\delta}$ helps to estimate $\|(1-\chi(|D|/R_\varepsilon))U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^s}$.

Remark 24. The classical Bernstein estimate ensures that $\dot{B}^{\frac{1}{2}}_{q,q} \hookrightarrow \dot{H}^{\frac{1}{2}-\frac{3}{2}\delta}$ so that $U_{0,\varepsilon,\text{osc}} \in \dot{H}^s$ for all $s \in \left[\frac{1}{2} - \frac{3}{2}\delta, \frac{1}{2} + \delta\right]$.

The rest of this article is structured as follows: we will first prove Theorem 14, then turn to the proof of Theorem 15 in the case $\nu = \nu'$ (much easier computations to obtain the eigenvalues and vectors, but needs more careful use for the Strichartz estimates as W_ε is not frequency-truncated) and we will finish with the proof of Theorem 21 (the eigenvectors are not mutually orthogonal anymore, and care is needed for the frequency-truncated terms). We end the article with an Appendix gathering results on Sobolev and Besov spaces, the process of diagonalization of system (A-79), and the new Strichartz estimates that allow us to reach this level of precision.

2. Proof of the results

2.1. ***The limit system.*** If $\widetilde{U}_{0,\mathrm{QG}}$ is as described in Theorem 14, we regularize it by introducing for $\lambda > 0$ (where χ is the smooth cut-off function introduced in the Appendix)

$$\widetilde{U}^{\lambda}_{0,\mathrm{QG}} \overset{\text{def}}{=} \chi\Big(\frac{|D|}{\lambda}\Big)\widetilde{U}_{0,\mathrm{QG}}.$$

Then $\widetilde{U}^{\lambda}_{0,\mathrm{QG}} \in H^1$ and applying Theorem 2 from [Charve 2004] there exists a unique global solution $\widetilde{U}^{\lambda}_{\mathrm{QG}} \in \dot{E}^0 \cap \dot{E}^1$ to system (QG) and thanks to Lemma 2.1 from [Charve 2008] we apply (1-12) to $\widetilde{U}^{\lambda}_{\mathrm{QG}}$ and for all $t \in \mathbb{R}_+$ (taking $\mathbb{C}_0 = \max(1, \|\widetilde{U}_{0,\mathrm{QG}}\|_{\dot{H}^{1/2+\delta}})$)

$$\begin{aligned}\|\widetilde{U}^{\lambda}_{\mathrm{QG}}\|^2_{L^\infty_t H^{1/2+\delta}} &+ \min(\nu,\nu')\int_0^t \|\nabla \widetilde{U}^{\lambda}_{\mathrm{QG}}(\tau)\|^2_{H^{1/2+\delta}}\,d\tau \\ &\le C_{\delta,\nu_0}\Big\|\chi\Big(\frac{|D|}{\lambda}\Big)\widetilde{U}_{0,\mathrm{QG}}\Big\|^2_{H^{1/2+\delta}} \max\Big(1, \Big\|\chi\Big(\frac{|D|}{\lambda}\Big)\widetilde{U}_{0,\mathrm{QG}}\Big\|^{\frac{1}{\delta}}_{H^{1/2+\delta}}\Big) \\ &\le C_{\delta,\nu_0}\|\widetilde{U}_{0,\mathrm{QG}}\|^2_{H^{1/2+\delta}}\max(1, \|\widetilde{U}_{0,\mathrm{QG}}\|^{\frac{1}{\delta}}_{H^{1/2+\delta}}) \le C_{\delta,\nu_0}\mathbb{C}_0^{2+\frac{1}{\delta}}. \end{aligned} \tag{2-17}$$

Then (taking $\lambda = n$) we prove that $(\widetilde{U}^{n}_{\mathrm{QG}})_{n\in\mathbb{N}^*}$ is a Cauchy sequence in $E^{\frac12+\delta} = \dot{E}^0 \cap \dot{E}^{\frac12+\delta}$. For $n \ge m$, let us define $\tilde{\delta}_{n,m} = \widetilde{U}^{n}_{\mathrm{QG}} - \widetilde{U}^{m}_{\mathrm{QG}}$, which satisfies the system

$$\begin{cases} \partial_t \tilde{\delta}_{n,m} - \Gamma\tilde{\delta}_{n,m} = -\mathcal{Q}(\widetilde{U}^{n}_{\mathrm{QG}}\cdot\nabla\tilde{\delta}_{n,m} + \tilde{\delta}_{n,m}\cdot\nabla\widetilde{U}^{m}_{\mathrm{QG}}), \\ \tilde{\delta}_{n,m|t=0} = \Big(\chi\Big(\frac{|D|}{n}\Big) - \chi\Big(\frac{|D|}{m}\Big)\Big)\widetilde{U}_{0,\mathrm{QG}}. \end{cases} \tag{2-18}$$

For any $s \in \big[0, \frac12+\delta\big]$, taking the $\dot{H}^s$-inner product and then using the classical Sobolev product laws (see Proposition 38), we get $\big((s_1,s_2) \in \big\{\big(1, s-\frac12\big), \big(s, \frac12\big)\big\}\big)$

$$\begin{aligned}\frac12\frac{d}{dt}\|\tilde{\delta}_{n,m}\|^2_{\dot{H}^s} &+ \nu_0\|\nabla\tilde{\delta}_{n,m}\|^2_{\dot{H}^s} \\ &\le C\|\widetilde{U}^{n}_{\mathrm{QG}}\cdot\nabla\tilde{\delta}_{n,m} + \tilde{\delta}_{n,m}\cdot\nabla\widetilde{U}^{m}_{\mathrm{QG}}\|_{\dot{H}^{s-1}}\|\tilde{\delta}_{n,m}\|_{\dot{H}^{s+1}} \\ &\le C(\|\widetilde{U}^{n}_{\mathrm{QG}}\|_{\dot{H}^1}\|\tilde{\delta}_{n,m}\|^{\frac12}_{\dot{H}^s}\|\tilde{\delta}_{n,m}\|^{\frac32}_{\dot{H}^{s+1}} + \|\widetilde{U}^{m}_{\mathrm{QG}}\|_{\dot{H}^{3/2}}\|\tilde{\delta}_{n,m}\|_{\dot{H}^s}\|\tilde{\delta}_{n,m}\|_{\dot{H}^{s+1}}) \\ &\le \frac{\nu_0}{2}\|\nabla\tilde{\delta}_{n,m}\|^2_{\dot{H}^s} + \frac{C}{\nu_0}\|\tilde{\delta}_{n,m}\|^2_{\dot{H}^s}\Big(\|\nabla\widetilde{U}^{m}_{\mathrm{QG}}\|^2_{\dot{H}^{1/2}} + \frac{1}{\nu_0^2}\|\widetilde{U}^{n}_{\mathrm{QG}}\|^2_{\dot{H}^{1/2}}\|\nabla\widetilde{U}^{n}_{\mathrm{QG}}\|^2_{\dot{H}^{1/2}}\Big). \end{aligned} \tag{2-19}$$

Thanks to the Gronwall lemma and using (2-17), we obtain that

$$\|\tilde{\delta}_{n,m}\|^2_{E^{1/2+\delta}} \leq \|\tilde{\delta}_{n,m}(0)\|^2_{H^{1/2+\delta}} e^{(C_{\delta,\nu_0}/\nu_0^2)\mathbb{C}_0^{2+1/\delta}(1+(C_{\delta,\nu_0}/\nu_0^2)\mathbb{C}_0^{2+1/\delta})}.$$

As $\|\tilde{\delta}_{n,m}(0)\|_{H^{1/2+\delta}}$ goes to zero when $m = \min(n,m)$ goes to infinity, the sequence is Cauchy and if we denote by $\widetilde{U}_{\mathrm{QG}}$ its limit in $E^{\frac{1}{2}+\delta}$, we immediately get that it solves system (QG) and satisfies the expected estimates. □

As an immediate consequence we easily bound $G^{b,l}$ (introduced with the auxiliary systems) as follows:

Proposition 25. *There exists a constant $C_F > 0$ such that, for all $\delta \in \left]0, \frac{1}{2}\right]$ and $s \in \left[0, \frac{1}{2}+\delta\right]$,*

$$\begin{aligned}\int_0^\infty \|G^b(\tau)\|_{\dot{H}^s}\,d\tau &\leq \frac{C_F}{\nu_0} C_{\delta,\nu_0}\mathbb{C}_0^{2+\frac{1}{\delta}},\\ \int_0^\infty \|G^l(\tau)\|^2_{\dot{H}^{s-1}}\,d\tau &\leq C_F\frac{|\nu-\nu'|^2}{\nu_0} C_{\delta,\nu_0}\mathbb{C}_0^{2+\frac{1}{\delta}}.\end{aligned} \tag{2-20}$$

Remark 26. In [Charve 2004] the previous terms were estimated for any $s \in [0,1]$ with $\|\widetilde{U}_{0,\mathrm{QG}}\|_{H^1}$.

Proof of Proposition 25. G^l is estimated as in [Charve 2004], and for G^b, as we wish to use only $\frac{1}{2}+\delta$ derivatives on $\widetilde{U}_{0,\mathrm{QG}}$, a much better way than in that work is to write (thanks to the Bony decomposition; see the Appendix for details)

$$\begin{aligned}\|G^b\|_{\dot{H}^s} &\leq C_F\|\widetilde{U}_{\mathrm{QG}}\cdot\nabla\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^s} \leq C_F\|\operatorname{div}(\widetilde{U}_{\mathrm{QG}}\otimes\widetilde{U}_{\mathrm{QG}})\|_{\dot{H}^s}\\ &\leq C_F(2\|T_{\widetilde{U}_{\mathrm{QG}}}\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^{s+1}} + \|R(\widetilde{U}_{\mathrm{QG}},\widetilde{U}_{\mathrm{QG}})\|_{\dot{H}^{s+1}})\\ &\leq C_F(2\|\widetilde{U}_{\mathrm{QG}}\|_{L^\infty} + \|\widetilde{U}_{\mathrm{QG}}\|_{\dot{B}^0_{\infty,\infty}})\|\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^{s+1}}.\end{aligned} \tag{2-21}$$

Then using the injection $\dot{B}^0_{\infty,1} \hookrightarrow L^\infty$ together with the Bernstein lemma and Lemma 27 below (whose proof is close to Lemma 5 from [Charve 2016]), we obtain that

$$2\|\widetilde{U}_{\mathrm{QG}}\|_{L^\infty} + \|\widetilde{U}_{\mathrm{QG}}\|_{\dot{B}^0_{\infty,\infty}} \leq 3\|\widetilde{U}_{\mathrm{QG}}\|_{\dot{B}^{3/2}_{2,1}} \leq C\|\widetilde{U}_{\mathrm{QG}}\|^{\frac{1}{2}}_{\dot{H}^{3/2-\delta}}\|\widetilde{U}_{\mathrm{QG}}\|^{\frac{1}{2}}_{\dot{H}^{3/2+\delta}}, \tag{2-22}$$

and we end up with (using also (1-12))

$$\begin{aligned}\int_0^\infty \|G^b\|_{\dot{H}^s}\,d\tau &\leq C_F\|\nabla\widetilde{U}_{\mathrm{QG}}\|^{\frac{1}{2}}_{L^2\dot{H}^{1/2-\delta}}\|\nabla\widetilde{U}_{\mathrm{QG}}\|^{\frac{1}{2}}_{L^2\dot{H}^{1/2+\delta}}\|\nabla\widetilde{U}_{\mathrm{QG}}\|_{L^2\dot{H}^s}\\ &\leq \frac{C_F}{\nu_0}C_{\delta,\nu_0}\mathbb{C}_0^{2+\frac{1}{\delta}}.\end{aligned} \tag{2-23}$$

This completes the proof of the proposition. □

Lemma 27. *For any $\alpha,\beta > 0$ there exists a constant $C_{\alpha,\beta} > 0$ such that for any $u \in \dot{H}^{s-\alpha}\cap\dot{H}^{s+\beta}$ we have $u \in \dot{B}^s_{2,1}$ and*

$$\|u\|_{\dot{B}^s_{2,1}} \leq C_{\alpha,\beta}\|u\|^{\frac{\beta}{\alpha+\beta}}_{\dot{H}^{s-\alpha}}\|u\|^{\frac{\alpha}{\alpha+\beta}}_{\dot{H}^{s+\beta}}. \tag{2-24}$$

2.2. ***The case $\nu = \nu'$.***

2.2.1. *Estimates for W_ε.* Let us first focus on the linear system (1-10). Let us recall that thanks to Proposition 25 we obtain that (see [Charve 2004] for details) for any $s \in \left[\frac{1}{2}, \frac{1}{2} + \delta\right]$

$$\|W_\varepsilon\|_{\dot{E}^s}^2 \le \left(\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^s}^2 + \frac{1}{2}\int_0^t \|G^b(\tau)\|_{\dot{H}^s}\right) e^{\frac{1}{2}\int_0^t \|G^b(\tau)\|_{\dot{H}^s}} \le \mathbb{D}_0(\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^s}^2 + 1), \tag{2-25}$$

with

$$\mathbb{D}_0 \overset{\text{def}}{=} \frac{C_F}{\nu_0} C_{\delta,\nu_0} \mathbb{C}_0^{2+\frac{1}{\delta}} e^{\frac{C_F}{\nu_0} C_{\delta,\nu_0} \mathbb{C}_0^{2+1/\delta}}.$$

One of the main ingredients is to provide a generalization of the Strichartz estimates obtained in [Charve 2006]. Our new Strichartz estimates are much more flexible and we refer to the Appendix for the most general formulation (see Propositions 47 and 51). We also postpone to the end of the next section the precise statement of the Strichartz estimates we will use.

2.2.2. *Energy estimates.* As explained in Section 1.3, we already have a local strong solution U_ε whose lifespan is denoted by T_ε^*. As seen in the previous section $\widetilde{U}_{\mathrm{QG}}$ and W_ε exist globally, and δ_ε is well-defined in $\dot{E}_T^{\frac{1}{2}} \cap \dot{E}_T^{\frac{1}{2}+\delta}$ for all $T < T_\varepsilon^*$ and we can perform for any $s \in \left[\frac{1}{2}, \frac{1}{2} + \eta\delta\right]$ the inner product in $\dot{H}^s$ of system (1-11) with δ_ε. We have to bound each term from the right-hand side.

Let us begin with the easiest terms, namely F_1, F_2 and F_3: thanks to the classical Sobolev product laws ($(s_1, s_2) = \left(\frac{1}{2}, s\right)$; see Proposition 38), we obtain that

$$|(F_1 \mid \delta_\varepsilon)_{\dot{H}^s}| \le \|\delta_\varepsilon \cdot \nabla\delta_\varepsilon\|_{\dot{H}^{s-1}} \|\delta_\varepsilon\|_{\dot{H}^{s+1}} \le C\|\delta_\varepsilon\|_{\dot{H}^{\frac{1}{2}}} \|\delta_\varepsilon\|_{\dot{H}^{s+1}}^2. \tag{2-26}$$

Similarly we obtain that

$$\begin{aligned} |(F_2 \mid \delta_\varepsilon)_{\dot{H}^s}| &\le C\|\nabla\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^{\frac{1}{2}}} \|\delta_\varepsilon\|_{\dot{H}^s} \|\delta_\varepsilon\|_{\dot{H}^{s+1}} \le \frac{\nu}{16}\|\delta_\varepsilon\|_{\dot{H}^{s+1}}^2 + \frac{C}{\nu}\|\nabla\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^{1/2}}^2 \|\delta_\varepsilon\|_{\dot{H}^s}^2, \\ |(F_3 \mid \delta_\varepsilon)_{\dot{H}^s}| &\le C\|\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^1} \|\delta_\varepsilon\|_{\dot{H}^{s+1/2}} \|\delta_\varepsilon\|_{\dot{H}^{s+1}} \le \frac{\nu}{16}\|\delta_\varepsilon\|_{\dot{H}^{s+1}}^2 + \frac{C}{\nu^3}\|\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^1}^4 \|\delta_\varepsilon\|_{\dot{H}^s}^2. \end{aligned} \tag{2-27}$$

Compared to [Charve 2004; 2006] we cannot use for the other F_i the same methods because they would produce (after using the Gronwall lemma) a coefficient of the form $e^{\|W_\varepsilon\|_{\dot{E}^s}}$ which would ruin our efforts to allow large initial blow up for the oscillating part (which could only be of size $(-\ln\varepsilon)^\beta$). We need to estimate carefully these terms and especially use as much as possible the new Strichartz estimates (giving positive powers of ε thanks to Proposition 47) and the least possible basic energy estimates on W_ε (that produce $\varepsilon^{-\gamma}$ from (2-25)).

The most obvious way would be to use the paraproduct and remainder laws (see the Appendix). For example with F_7, as $s - 1 < 0$, we have

$$\begin{aligned} |(F_7 \mid \delta_\varepsilon)_{\dot{H}^s}| &\le \|W_\varepsilon \cdot \nabla\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^{s-1}} \|\delta_\varepsilon\|_{\dot{H}^{s+1}} \\ &\le C(\|T_{W_\varepsilon}\nabla\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^{s-1}} + \|T_{\nabla\widetilde{U}_{\mathrm{QG}}} W_\varepsilon\|_{\dot{H}^{s-1}} + \|\operatorname{div}(R(W_\varepsilon, \widetilde{U}_{\mathrm{QG}}))\|_{\dot{H}^{s-1}})\|\delta_\varepsilon\|_{\dot{H}^{s+1}} \\ &\le C(\|W_\varepsilon\|_{L^\infty}\|\nabla\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^{s-1}} + \|\nabla\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^{s-1}}\|W_\varepsilon\|_{\dot{B}^0_{\infty,\infty}} + \|W_\varepsilon\|_{\dot{B}^0_{\infty,\infty}}\|\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^s})\|\delta_\varepsilon\|_{\dot{H}^{s+1}} \\ &\le C\|W_\varepsilon\|_{\dot{B}^0_{\infty,1}}\|\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^s}\|\delta_\varepsilon\|_{\dot{H}^{s+1}} \le \frac{\nu}{16}\|\delta_\varepsilon\|_{\dot{H}^{s+1}}^2 + \frac{C}{\nu}\|W_\varepsilon\|_{\dot{B}^0_{\infty,1}}^2 \|\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^s}^2. \end{aligned} \tag{2-28}$$

This result could be also usable for F_5 but to deal with $\|W_\varepsilon\|_{L^p\dot{B}^0_{\infty,1}}$ from Proposition 47 we would have to use Lemma 27 which would force us to have a slightly smaller range for γ. More important, for F_8 this method would force us to require $\gamma < \frac{\delta}{4}$, which is clearly not optimal.

Finally, the most important problem is that the previous estimates cannot be used to estimate F_4 and F_6. Indeed for instance if we wish to estimate F_6 the same way

$$\|F_6\|_{\dot{H}^{s-1}} \leq C(\|T_{\widetilde{U}_{\mathrm{QG}}}\nabla W_\varepsilon\|_{\dot{H}^{s-1}} + \|T_{\nabla W_\varepsilon}\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^{s-1}} + \|\operatorname{div}(R(\widetilde{U}_{\mathrm{QG}}, W_\varepsilon))\|_{\dot{H}^{s-1}}),$$

and the first paraproduct (see the Appendix for the Bony decomposition) leads to an obstruction as the only possibilities to estimate it are (for $\beta > s$)

$$\|T_{\widetilde{U}_{\mathrm{QG}}}\nabla W_\varepsilon\|_{\dot{H}^{s-1}} \leq C \begin{cases} \|\widetilde{U}_{\mathrm{QG}}\|_{L^\infty}\|W_\varepsilon\|_{\dot{H}^s}, \\ \|\widetilde{U}_{\mathrm{QG}}\|_{\dot{H}^{s-\beta}}\|W_\varepsilon\|_{\dot{B}^\beta_{\infty,\infty}}, \end{cases} \tag{2-29}$$

In the first estimate each term is well-defined but the $\dot{H}^s$-norm of W_ε produces negative powers of ε, and in the second one the first term is not defined ($\widetilde{U}_{\mathrm{QG}}$ is not defined for negative regularities). It is possible to deal with this term using the same idea as in [Charve 2004] (with $a, b \geq 1$ so that $\frac{1}{a} + \frac{1}{b} = 1$),

$$\begin{aligned}\int_0^t \|\widetilde{U}_{\mathrm{QG}}\cdot\nabla W_\varepsilon\|^2_{\dot{H}^{s-1}}\,d\tau &\leq C\int_0^t \|\widetilde{U}_{\mathrm{QG}}\cdot\nabla W_\varepsilon\|_{L^2}\|\widetilde{U}_{\mathrm{QG}}\cdot\nabla W_\varepsilon\|_{\dot{H}^{2(s-1)}}\,d\tau \\ &\leq \|\widetilde{U}_{\mathrm{QG}}\|_{L^\infty L^2}\|\nabla W_\varepsilon\|_{L^aL^\infty}\|\widetilde{U}_{\mathrm{QG}}\|_{L^b\dot{H}^{s+1/2}}\|W_\varepsilon\|_{L^\infty\dot{H}^s},\end{aligned} \tag{2-30}$$

and due to the gradient pounding on W_ε, the most interesting use of Proposition 47 consists in choosing a as close as possible to 1, which implies that b is very large. As $s + \frac{1}{2} \geq 1$, this forces us to use (1-12) for regularity index close to 1 (in this case it would be necessary to require that $\widetilde{U}_{0,\mathrm{QG}} \in H^s$ with s close to 1), which was something we wished to avoid as we only consider indices $s \leq \frac{1}{2} + \delta$. Moreover it would also produce a clearly nonoptimal decrease in ε.

Finally both of these two methods fail for F_4: the former for the same reason as for F_6, and the latter as we cannot consider $\|\delta_\varepsilon\|_{L^2}$: there is a lack of derivatives pounding on δ_ε.

To overcome this lack of derivatives, we will distribute them differently among the whole $\dot{H}^s$-inner product. We will do this for all the last five external force terms and the idea will be to do as in [Charve 2016; 2018a] and deal with the nonlocal operator $|D|^s$ applied to a product and dispatch s derivatives on δ_ε and obtain something close to the second line of (2-29). More precisely, we directly deal with the inner product as follows:

$$|(F_4 \mid \delta_\varepsilon)_{\dot{H}^s}| = |(\operatorname{div}(\delta_\varepsilon \otimes W_\varepsilon) \mid \delta_\varepsilon)_{\dot{H}^s}| = |(|D|^s(\delta_\varepsilon\cdot W_\varepsilon) \mid |D|^s\nabla\delta_\varepsilon)_{L_2}|. \tag{2-31}$$

The nonlocal operator $|D|^s$ can be written as a singular principal value integral (we refer to [Stein 1970; Córdoba and Córdoba 2004; Hmidi and Keraani 2007; Abidi and Hmidi 2008; Charve 2016; 2018a]) and when the index s lies in $]0,1[$ (which is the case here as s is close to $\frac{1}{2}$) it is a classical singular integral:

$$|D|^s f(x) = C_s\int_{\mathbb{R}^3}\frac{f(x)-f(y)}{|x-y|^{3+s}}\,dy = C_s\int_{\mathbb{R}^3}\frac{f(x)-f(x-y)}{|y|^{3+s}}\,dy.$$

Let us recall that an equivalent formulation of the Besov norm involves translations as stated in the following result:

Theorem 28 [Bahouri et al. 2011, 2.36]. *Let $s \in]0,1[$ and $p, r \in [1,\infty]$. There exists a constant C such that for any $u \in \dot{B}^s_{p,r}$*

$$C^{-1}\|u\|_{\dot{B}^s_{p,r}} \le \Big\|\frac{\|u(\cdot - y) - u(\cdot)\|_{L^p}}{|y|^s}\Big\|_{L^r(\mathbb{R}^d;\frac{dy}{|y|^d})} \le C\|u\|_{\dot{B}^s_{p,r}}.$$

From this we can prove exactly as in [Charve 2018a] (see Section A.3.1 there) the following result:

Proposition 29. *For any $s \in]0,1[$ and any smooth functions f, g we can write*

$$|D|^s(fg) = (|D|^s f)g + f|D|^s g + M_s(f,g),$$

where the bilinear operator M_s is defined for all $x \in \mathbb{R}^3$ by

$$M_s(f,g)(x) = \int_{\mathbb{R}^3} \frac{(f(x)-f(x-y))(g(x)-g(x-y))}{|y|^{3+s}}\,dy. \tag{2-32}$$

Moreover there exists a constant C_s such that for all f, g and all $p, p_1, p_2, r_1, r_2 \in [1,\infty]$ and $s_1, s_2 > 0$ satisfying

$$\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}, \quad 1 = \frac{1}{r_1} + \frac{1}{r_2}, \quad s_1 + s_2 = s,$$

we have

$$\|M_s(f,g)\|_{L^p} \le C_s\|f\|_{\dot{B}^{s_1}_{p_1,r_1}}\|g\|_{\dot{B}^{s_2}_{p_2,r_2}}. \tag{2-33}$$

Remark 30. The additional term M_s allows us to freely dispatch the derivatives as desired provided that $s_1, s_2 > 0$, which will force us to spend a small extra amount of derivative in order to meet these conditions. So even if it is not possible to use Proposition 29 for $(s_1, s_2) = (s, 0)$, our method will enable us to do nearly as if we could estimate $\|M_s(\delta_\varepsilon, W_\varepsilon)\|_{L^2}$ by $\|\delta_\varepsilon\|_{\dot{H}^{1/2}}\||D|^s W_\varepsilon\|_{L^6}$.

More precisely for a small $\alpha_1 > 0$, instead of (2-31), we will write (also using the Sobolev injections):

$$\begin{aligned}
&|(F_6\,|\,\delta_\varepsilon)_{\dot{H}^s}|\\
&= |(\operatorname{div}(\widetilde{U}_{QG}\otimes W_\varepsilon)\,|\,\delta_\varepsilon)_{\dot{H}^s}| = |(|D|^{s+\alpha_1}(\widetilde{U}_{QG}\cdot W_\varepsilon)\,|\,|D|^{s-\alpha_1}\nabla\delta_\varepsilon)_{L^2}|\\
&\le \|(|D|^{s+\alpha_1}\widetilde{U}_{QG})\cdot W_\varepsilon + \widetilde{U}_{QG}\cdot|D|^{s+\alpha_1}W_\varepsilon + M_{s+\alpha_1}(\widetilde{U}_{QG}, W_\varepsilon)\|_{L^{6/(3+2\alpha_1)}}\cdot\||D|^{s-\alpha_1}\nabla\delta_\varepsilon\|_{L^{6/(3-2\alpha_1)}}\\
&\le C(\||D|^{s+\alpha_1}\widetilde{U}_{QG}\|_{L^2}\|W_\varepsilon\|_{L^{3/\alpha_1}} + \|\widetilde{U}_{QG}\|_{L^3}\||D|^{s+\alpha_1}W_\varepsilon\|_{L^{6/(1+2\alpha_1)}} + \|\widetilde{U}_{QG}\|_{\dot{H}^s}\|W_\varepsilon\|_{\dot{B}^{\alpha_1}_{3/\alpha_1,2}})\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\times\||D|^{s-\alpha_1}\nabla\delta_\varepsilon\|_{\dot{H}^{\alpha_1}}\\
&\le C(\|\widetilde{U}_{QG}\|_{\dot{H}^{s+\alpha_1}}\|W_\varepsilon\|_{L^{3/\alpha_1}} + \|\widetilde{U}_{QG}\|_{\dot{H}^{1/2}}\||D|^{s+\alpha_1}W_\varepsilon\|_{L^{6/(1+2\alpha_1)}} + \|\widetilde{U}_{QG}\|_{\dot{H}^s}\|W_\varepsilon\|_{\dot{B}^{\alpha_1}_{3/\alpha_1,2}})\cdot\|\delta_\varepsilon\|_{\dot{H}^{s+1}}\\
&\le \frac{\nu}{16}\|\delta_\varepsilon\|^2_{\dot{H}^{s+1}} + \frac{C}{\nu}\big(\|\widetilde{U}_{QG}\|^{2(1-\alpha_1)}_{\dot{H}^s}\|\widetilde{U}_{QG}\|^{2\alpha_1}_{\dot{H}^{s+1}}\|W_\varepsilon\|^2_{L^{3/\alpha_1}}\\
&\qquad\qquad + \|\widetilde{U}_{QG}\|^2_{\dot{H}^{1/2}}\||D|^{s+\alpha_1}W_\varepsilon\|^2_{L^{6/(1+2\alpha_1)}} + \|\widetilde{U}_{QG}\|^2_{\dot{H}^s}\|W_\varepsilon\|^2_{\dot{B}^{\alpha_1}_{3/\alpha_1,2}}\big).
\end{aligned} \tag{2-34}$$

Remark 31. Notice that as δ_ε, W_ε, $\widetilde{U}_{\mathrm{QG}}$ are divergence-free, we will systematically (thanks to integration by parts) transfer the divergence as a gradient on the right-hand part of the inner product, and as a consequence the computations are the same respectively for F_4 and F_5 and for F_6 and F_7.

Let us continue with F_4, by the classical Sobolev interpolation and Young estimates, we can write that (for $\alpha_2 > 0$ small)

$$
\begin{aligned}
&|(F_4\,|\,\delta_\varepsilon)_{\dot H^s}|\\
&= |(\operatorname{div}(\delta_\varepsilon\otimes W_\varepsilon)\,|\,\delta_\varepsilon)_{\dot H^s}| = |(|D|^{s+\alpha_2}(\delta_\varepsilon\cdot W_\varepsilon)\,|\,|D|^{s-\alpha_2}\nabla\delta_\varepsilon)_{L_2}|\\
&\le C\,\|(|D|^{s+\alpha_2})\cdot W_\varepsilon+\delta_\varepsilon\cdot|D|^{s+\alpha_2}W_\varepsilon+M_{s+\alpha_2}(\delta_\varepsilon,W_\varepsilon)\|_{L^{6/(3+2\alpha_2)}}\cdot\||D|^{s-\alpha_2}\nabla\delta_\varepsilon\|_{L^{6/(3-2\alpha_2)}}\\
&\le C\Big(\||D|^{s+\alpha_2}\delta_\varepsilon\|_{L^2}\|W_\varepsilon\|_{L^{3/\alpha_2}}+\|\delta_\varepsilon\|_{L^3}\||D|^{s+\alpha_2}W_\varepsilon\|_{L^{6/(1+2\alpha_2)}}+\|\delta_\varepsilon\|_{\dot H^s}\|W_\varepsilon\|_{\dot B^{\alpha_2}_{3/\alpha_2,2}}\Big)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\times\||D|^{s-\alpha_2}\nabla\delta_\varepsilon\|_{\dot H^{\alpha_2}}\\
&\le C\,\|\delta_\varepsilon\|_{\dot H^s}^{1-\alpha_2}\|\delta_\varepsilon\|_{\dot H^{s+1}}^{1+\alpha_2}\|W_\varepsilon\|_{L^{3/\alpha_2}}+C\Big(\|\delta_\varepsilon\|_{L^3}\||D|^{s+\alpha_2}W_\varepsilon\|_{L^{6/(1+2\alpha_2)}}+\|\delta_\varepsilon\|_{\dot H^s}\|W_\varepsilon\|_{\dot B^{\alpha_2}_{3/\alpha_2,2}}\Big)\|\delta_\varepsilon\|_{\dot H^{s+1}}\\
&\le \frac{\nu}{16}\|\delta_\varepsilon\|_{\dot H^{s+1}}^2\\
&\qquad+C\,\|\delta_\varepsilon\|_{\dot H^s}^2\bigg(\frac{1}{\nu^{\frac{1+\alpha_2}{1-\alpha_2}}}\|W_\varepsilon\|_{L^{3/\alpha_2}}^{\frac{2}{1-\alpha_2}}+\frac{1}{\nu}\|W_\varepsilon\|_{\dot B^{\alpha_2}_{3/\alpha_2,2}}^2\bigg)+\frac{C}{\nu}\|\delta_\varepsilon\|_{\dot H^{1/2}}^2\||D|^{s+\alpha_2}W_\varepsilon\|_{L^{6/(1+2\alpha_2)}}^2. \qquad (2\text{-}35)
\end{aligned}
$$

Finally we estimate F_8 with the same method, but the term $M_{s+\alpha_3}(W_\varepsilon, W_\varepsilon)$ has to be estimated differently (otherwise we end up with the same problem as explained in the beginning of this section): instead of estimating it as for the other terms by

$$\|W_\varepsilon\|_{\dot H^s}\|W_\varepsilon\|_{\dot B^{\alpha_3}_{3/\alpha_3,2}}$$

(the first term being L^∞, and the second L^2 in time), we will estimate it by

$$\|W_\varepsilon\|_{\dot H^{s+\alpha_3-\beta\delta}}\|W_\varepsilon\|_{\dot B^{\beta\delta}_{3/\alpha_3,2}},$$

for small enough $\alpha_3, \beta > 0$ so that the first term keeps L^∞ in time and the second one is L^2 (we try to be as close as possible to the forbidden choice $\beta = 0$). As we will make precise below, dealing with

$$\|W_\varepsilon\|_{L^\infty\dot H^s}^{2(1-\alpha_3)}\|W_\varepsilon\|_{L^2\dot H^s}^{2\alpha_3}\|W_\varepsilon\|_{L^2L^{3/\alpha_3}}^2$$

(for the first term) will only lead to $\gamma < \frac{\delta}{4}$, whereas

$$\|W_\varepsilon\|_{L^\infty\dot H^{s+\alpha_3}}^2\|W_\varepsilon\|_{L^2L^{3/\alpha_3}}^2$$

will allow us to reach $\gamma < \frac{\delta}{2}$. For the same reason we will estimate the other term by

$$\|W_\varepsilon\|_{L^2\dot B^{\beta\delta}_{3/\alpha_3,2}}$$

instead of

$$\|W_\varepsilon\|_{L^{2/(1-\alpha_3)}\dot B^{\beta\delta}_{3/\alpha_3,2}}.$$

Although this choice seems very close to the other, it allows us to use a smaller p in the Strichartz estimates, which allows a slightly wider range for θ helping us to reach $\gamma < \frac{\delta}{2}$ instead of $\gamma < \frac{\delta}{4}$. Once more, we try to obtain as close as possible to what we would get if it Proposition 29 could be applied for $s_1 = s+\alpha_3$ and $s_2 = 0$:

$$\begin{aligned}|(F_8\,|\,\delta_\varepsilon)_{\dot H^s}| &= \||D|^{s+\alpha_3}(W_\varepsilon\otimes W_\varepsilon)\|_{L^{6/(3+2\alpha_3)}}\||D|^{s-\alpha_3}\nabla\delta_\varepsilon\|_{L^{6/(3-2\alpha_3)}}\\ &\le (2\||D|^{s+\alpha_3}W_\varepsilon\|_{L^2}\|W_\varepsilon\|_{L^{3/\alpha_3}}+\|W_\varepsilon\|_{\dot H^{s+\alpha_3-\beta\delta}}\|W_\varepsilon\|_{\dot B^{\beta\delta}_{3/\alpha_3,2}})\cdot\||D|^{s-\alpha_3}\nabla\delta_\varepsilon\|_{\dot H^{\alpha_3}}\\ &\le \frac{\nu}{16}\|\delta_\varepsilon\|^2_{\dot H^{s+1}}+\frac{C}{\nu}(\|W_\varepsilon\|^2_{\dot H^{s+\alpha_3}}\|W_\varepsilon\|^2_{L^{3/\alpha_3}}+\|W_\varepsilon\|^2_{\dot H^{s+\alpha_3-\beta\delta}}\|W_\varepsilon\|^2_{\dot B^{\beta\delta}_{3/\alpha_3,2}}).\end{aligned}\tag{2-36}$$

We can now gather all the external force term estimates (2-26), (2-27), (2-35), (2-34), (2-36) and taking the $\dot H^s$-inner product of system (1-11) with δ_ε, we obtain that for all $s\in\left[\frac12,\frac12+\eta\delta\right]$ and all $t<T^*_\varepsilon$

$$\begin{aligned}&\frac12\frac{d}{dt}\|\delta_\varepsilon\|^2_{\dot H^s}+\nu\|\nabla\delta_\varepsilon\|^2_{\dot H^s}\\ &\quad\le\left(C\|\delta_\varepsilon\|_{\dot H^{1/2}}+8\frac{\nu}{16}\right)\|\delta_\varepsilon\|^2_{\dot H^s}\\ &\qquad+\frac{C}{\nu}\|\delta_\varepsilon\|^2_{\dot H^s}\left\{\|\nabla\widetilde U_{\mathrm{QG}}\|^2_{\dot H^{1/2}}\left(1+\frac{1}{\nu^2}\|\widetilde U_{\mathrm{QG}}\|^2_{\dot H^{1/2}}\right)+\frac{1}{\nu^{\frac{2\alpha_2}{1-\alpha_2}}}\|W_\varepsilon\|^{\frac{2}{1-\alpha_2}}_{L^{3/\alpha_2}}+\|W_\varepsilon\|^2_{\dot B^{\alpha_2}_{3/\alpha_2,2}}\right\}\\ &\qquad\quad+\frac{C}{\nu}\Big[\|\widetilde U_{\mathrm{QG}}\|^{2(1-\alpha_1)}_{\dot H^s}\|\widetilde U_{\mathrm{QG}}\|^{2\alpha_1}_{\dot H^{s+1}}\|W_\varepsilon\|^2_{L^{3/\alpha_1}}+\|\widetilde U_{\mathrm{QG}}\|^2_{\dot H^{1/2}}\||D|^{s+\alpha_1}W_\varepsilon\|^2_{L^{6/(1+2\alpha_1)}}\\ &\qquad\qquad+\|\widetilde U_{\mathrm{QG}}\|^2_{\dot H^s}\|W_\varepsilon\|^2_{\dot B^{\alpha_1}_{3/\alpha_1,2}}+\|\delta_\varepsilon\|^2_{\dot H^{1/2}}\||D|^{s+\alpha_2}W_\varepsilon\|^2_{L^{6/(1+2\alpha_2)}}\\ &\qquad\qquad\qquad+\|W_\varepsilon\|^2_{\dot H^{s+\alpha_3}}\|W_\varepsilon\|^2_{L^{3/\alpha_3}}+\|W_\varepsilon\|^2_{\dot H^{s+\alpha_3-\beta\delta}}\|W_\varepsilon\|^2_{\dot B^{\beta\delta}_{3/\alpha_3,2}}\Big].\end{aligned}\tag{2-37}$$

In order to perform the bootstrap argument (we refer to in [Charve 2004; 2006]), let us now define

$$T_\varepsilon \overset{\text{def}}{=} \sup\left\{t\in[0,T^*_\varepsilon[\,:\text{ for all } t'\le t,\ \|\delta_\varepsilon(t')\|_{\dot H^{1/2}}\le\frac{\nu}{4C}\right\}.\tag{2-38}$$

Due to the assumptions, $\|\delta_\varepsilon(0)\|_{H^{1/2+\delta}}\le\mathbb{C}_0\varepsilon^{\alpha_0}$ so that we are sure that $T_\varepsilon>0$ if

$$\varepsilon\le\left(\frac{\nu}{8C\mathbb{C}_0}\right)^{\frac{1}{\alpha_0}}.$$

Thanks to the Gronwall and Young estimates, and estimating the first terms in the last block as

$$\begin{aligned}&\int_0^\infty\|\widetilde U_{\mathrm{QG}}\|^{2(1-\alpha_1)}_{\dot H^s}\|\widetilde U_{\mathrm{QG}}\|^{2\alpha_1}_{\dot H^{s+1}}\|W_\varepsilon\|^2_{L^{3/\alpha_1}}\,d\tau\\ &\qquad\le\left(\int_0^\infty\|\widetilde U_{\mathrm{QG}}\|^2_{\dot H^{s+1}}\,d\tau\right)^{\alpha_1}\left(\int_0^\infty\|W_\varepsilon\|^{\frac{2}{1-\alpha_1}}_{L^{3/\alpha_1}}\|\widetilde U_{\mathrm{QG}}\|^2_{\dot H^s}\,d\tau\right)^{1-\alpha_1},\end{aligned}\tag{2-39}$$

we can now state that for all $s\in\left[\frac12,\frac12+\eta\delta\right]$ and all $t\le T_\varepsilon$, we have (as W_ε and $\widetilde U_{\mathrm{QG}}$ are globally defined, each time integral in the right-hand side is over $\mathbb{R}_+$)

$$\begin{aligned}
&\|\delta_\varepsilon(t)\|_{\dot H^s}^2+\frac{\nu}{2}\int_0^t\|\nabla\delta_\varepsilon(\tau)\|_{\dot H^s}^2\,d\tau\\
&\le\Bigg[\|\delta_\varepsilon(0)\|_{\dot H^s}^2+\frac{C}{\nu}\Big(\|\widetilde U_{\mathrm{QG}}\|_{L^\infty\dot H^s}^{2(1-\alpha_1)}\|\widetilde U_{\mathrm{QG}}\|_{L^2\dot H^{s+1}}^{2\alpha_1}\|W_\varepsilon\|_{L^{2/(1-\alpha_1)}L^{3/\alpha_1}}^2\\
&\qquad+\|\widetilde U_{\mathrm{QG}}\|_{L^\infty\dot H^{1/2}}^2\||D|^{s+\alpha_1}W_\varepsilon\|_{L^2L^{6/(1+2\alpha_1)}}^2+\|\widetilde U_{\mathrm{QG}}\|_{L^\infty\dot H^s}^2\|W_\varepsilon\|_{L^2\dot B^{\alpha_1}_{3/\alpha_1,2}}^2\\
&\qquad+\||D|^{s+\alpha_2}W_\varepsilon\|_{L^2L^{6/(1+2\alpha_2)}}^2+\|W_\varepsilon\|_{L^\infty\dot H^{s+\alpha_3}}^2\|W_\varepsilon\|_{L^2L^{3/\alpha_3}}^2\\
&\qquad+\|W_\varepsilon\|_{L^\infty\dot H^{s+\alpha_3-\beta\delta}}^2\|W_\varepsilon\|_{L^2\dot B^{\beta\delta}_{3/\alpha_3,2}}^2\Big)\Bigg]\\
&\quad\times\exp\frac{C}{\nu}\Bigg\{\|\nabla\widetilde U_{\mathrm{QG}}\|_{L^2\dot H^{1/2}}^2\Big(1+\frac{1}{\nu^2}\|\widetilde U_{\mathrm{QG}}\|_{L^\infty\dot H^{1/2}}^2\Big)\\
&\qquad+\frac{1}{\nu^{\frac{2\alpha_2}{1-\alpha_2}}}\|W_\varepsilon\|_{L^{2/(1-\alpha_2)}L^{3/\alpha_2}}^{\frac{2}{1-\alpha_2}}+\|W_\varepsilon\|_{L^2\dot B^{\alpha_2}_{3/\alpha_2,2}}^2\Bigg\}.
\end{aligned}\tag{2-40}$$

It is now time to properly use the new Strichartz estimates we proved in the present article (see the Appendix for Proposition 47 and its proof).

Let us begin with the case $(d,p,r,q)=\big(s+\alpha,2,\frac{6}{1+2\alpha},2\big)$, for all $\theta\in\big]0,\frac{1-\alpha}{1-4\alpha}\big[\cap]0,1]=]0,1]$. Thanks to Proposition 40 (for simplicity we will not track the dependency in ν),

$$\begin{aligned}
&\||D|^{s+\alpha}W_\varepsilon\|_{L^2_tL^{6/(1+2\alpha)}}\\
&\qquad\le C\||D|^{s+\alpha}W_\varepsilon\|_{\widetilde L^2_t\dot B^0_{6/(1+2\alpha),2}}\\
&\qquad\le C_{F,\nu,p,\theta,\alpha}\varepsilon^{\frac{\theta}{12}(1-4\alpha)}\Big(\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot H^{s+(\theta/6)(1-4\alpha)}}+\int_0^t\|G^b(\tau)\|_{\dot H^{s+(\theta/6)(1-4\alpha)}}\,d\tau\Big),
\end{aligned}\tag{2-41}$$

and if we choose $\alpha\in\big]0,\frac14\big[$, and

$$\theta=\frac{6(\delta+\frac12-s)}{1-4\alpha}$$

$\big($which is in $]0,1]$ if $\delta\le s-\frac13-\frac{2\alpha}{3}$, recall that $s\sim\frac12\big)$, then we obtain (thanks to Proposition 25)

$$\begin{aligned}
\||D|^{s+\alpha}W_\varepsilon\|_{L^2_tL^{6/(1+2\alpha)}}&\le C\||D|^{s+\alpha}W_\varepsilon\|_{\widetilde L^2_t\dot B^0_{6/(1+2\alpha),2}}\\
&\le C_{F,\nu,s,\delta,\alpha}\varepsilon^{\frac12(\delta+\frac12-s)}\Big(\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot H^{1/2+\delta}}+\int_0^t\|G^b(\tau)\|_{\dot H^{1/2+\delta}}\,d\tau\Big)\\
&\le C_{F,\nu,s,\delta,\alpha}\varepsilon^{\frac12(\delta+\frac12-s)}\mathbb{D}_0(\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot H^{1/2+\delta}}+1).
\end{aligned}\tag{2-42}$$

Let us continue with the case $(d,p,r,q)=\big(\alpha,2,\frac{3}{\alpha},2\big)$, for all

$$\theta\in\Bigg]0,\frac{\frac12-\frac{\alpha}{3}}{1-\frac{4\alpha}{3}}\Bigg[,$$

if we assume $\alpha\in]0,\frac34[$, and choose $\theta=\frac{6\delta}{3-4\alpha}$, then

$$\|W_\varepsilon\|_{\widetilde L^2_t\dot B^\alpha_{3/\alpha,2}}\le C_{F,\nu,\delta,\alpha}\varepsilon^{\frac{\delta}{2}}\mathbb{D}_0(\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot H^{1/2+\delta}}+1).\tag{2-43}$$

For the case $(d, p, r, q) = \big(0, \frac{2}{1-\alpha}, \frac{3}{\alpha}, 2\big)$, for all

$$\theta \in \left]0, \frac{\frac{1}{2}-\frac{\alpha}{3}}{1-\frac{4\alpha}{3}}\right[,$$

if $\alpha \in \left]0, \frac{3}{4}\right[$, and if we choose $\theta = \frac{6\delta}{3-4\alpha}$, then

$$\|W_\varepsilon\|_{L_t^{2/(1-\alpha)}L^{3/\alpha}} \le C\|W_\varepsilon\|_{\widetilde{L}_t^{2/(1-\alpha)}\dot{B}^0_{3/\alpha,2}} \le C_{F,\nu,\delta,\alpha}\varepsilon^{\frac{\delta}{2}}\mathbb{D}_0(\|U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^{1/2+\delta}}+1). \tag{2-44}$$

All these estimates are verified for $\alpha_1 = \alpha_2 = \alpha = \frac{1}{16}$ if $\delta \le \frac{1}{8}$. Then we turn to the last two terms from (2-36), let us begin with the first one: as announced, due to the first factor (estimated thanks to (2-25)), doing as before will only allow us to get $\varepsilon^{\frac{\delta}{2}}\mathbb{D}_0(\|U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^{1/2+\delta}}+1)^2$, which leads to $\gamma < \frac{\delta}{4}$. In order to reach the announced bound $\frac{\delta}{2}$, we will try to take a slightly smaller p which will allow us to widen the range for θ. But taking $p = 2$ instead of $\frac{2}{1-\alpha}$ requires that $\|W_\varepsilon\|_{\dot{H}^{s+\alpha_3}}$ is L^∞; that is, we need that $s+\alpha_3 \le \frac{1}{2}+\delta$. More precisely with $(d, p, r, q) = \big(0, 2, \frac{3}{\alpha}, 2\big)$, we have

$$\|W_\varepsilon\|_{L_t^2 L^{3/\alpha_3}} \le C\|W_\varepsilon\|_{\widetilde{L}_t^2\dot{B}^0_{3/\alpha_3,2}} \le C_{F,\nu,\theta,s}\varepsilon^{\frac{\theta}{4}(1-\frac{4\alpha_3}{3})}\mathbb{D}_0(\|U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^{1/2-\alpha_3+(\theta/2)(1-4\alpha_3/3)}}+1) \tag{2-45}$$

and as we want

$$\alpha_3 + s = \tfrac{1}{2}+\delta = \tfrac{1}{2}-\alpha_3+\tfrac{\theta}{2}\big(1-\tfrac{4\alpha_3}{3}\big)$$

we choose

$$(\alpha_3, \theta) = \left(\delta+\frac{1}{2}-s, \frac{2(\delta+\alpha_3)}{1-\frac{4\alpha_3}{3}}\right).$$

This is possible (according to the condition from Proposition 47) when

$$\theta < \frac{\frac{1}{2}-\frac{\alpha_3}{3}}{1-\frac{4\alpha_3}{3}},$$

that is if

$$\delta < \tfrac{7s-2}{13}, \tag{2-46}$$

which is realized $\big($recall that $s \in \big[\frac{1}{2}, \frac{1}{2}+\eta\delta\big]\big)$ when $\delta \le \frac{1}{10} < \frac{3}{26}$. Then we have

$$\|W_\varepsilon\|_{L_t^2 L^{3/\alpha_3}} \le C\|W_\varepsilon\|_{\widetilde{L}_t^2\dot{B}^0_{3/\alpha_3,2}} \le C_{F,\nu,\delta,s}\varepsilon^{\frac{1}{2}(2\delta+\frac{1}{2}-s)}\mathbb{D}_0(\|U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^{1/2+\delta}}+1). \tag{2-47}$$

Now, for the last term, α_3 is fixed and we will adjust θ and β. For $(d, p, r, q) = \big(\beta\delta, 2, \frac{3}{\alpha_3}, 2\big)$, we choose θ so that the corresponding σ (see Proposition 47) is equal to $\frac{1}{2}+\delta$; that is

$$\tfrac{\theta}{2}\big(1-\tfrac{4\alpha_3}{3}\big) = (2-\beta)\delta+\tfrac{1}{2}-s,$$

which is possible when

$$\theta \in \left]0, \frac{\frac{1}{2}-\frac{\alpha_3}{3}}{1-\frac{4\alpha_3}{3}}\right[,$$

that is $\delta < \frac{7s-2}{13-6\beta}$, which is realized when (2-46) is true (when $\beta \in]0,1[$). In this case, we end up with

$$\|W_\varepsilon\|_{\widetilde{L}_t^2 \dot{B}^{\beta\delta}_{3/\alpha_3,2}} \le C_{F,\nu,\alpha,\delta,s}\varepsilon^{\frac{1}{2}((2-\beta)\delta+\frac{1}{2}-s)}\mathbb{D}_0(\|U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^{1/2+\delta}}+1). \tag{2-48}$$

Combining (2-40) with all these Strichartz estimates, namely (2-42), (2-43), (2-44), (2-47) and (2-48), we end up with for all $s \in \left[\frac{1}{2}, \frac{1}{2}+\eta\delta\right]$, all $\beta > 0$ small and all $t \le T_\varepsilon$

$$\begin{aligned}
&\|\delta_\varepsilon(t)\|^2_{\dot{H}^s} + \frac{\nu}{2}\int_0^t \|\nabla\delta_\varepsilon(\tau)\|^2_{\dot{H}^s}\,d\tau \\
&\le \Big[\|\delta_\varepsilon(0)\|^2_{\dot{H}^s} + \mathbb{D}_0\big((\varepsilon^{\delta+\frac{1}{2}-s}+\varepsilon^\delta)(\|U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^{1/2+\delta}}+1)^2 \\
&\qquad + (\varepsilon^{2\delta+\frac{1}{2}-s}+\varepsilon^{(2-\beta)\delta+\frac{1}{2}-s})(\|U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^{1/2+\delta}}+1)^4\big)\Big] \\
&\qquad\qquad \times \exp\{\mathbb{D}_0(1+\varepsilon^\delta(\|U_{0,\varepsilon,\text{osc}}\|_{\dot{H}^{1/2+\delta}}+1)^2)\} \\
&\le \mathbb{D}_0[\varepsilon^{2\alpha_0} + (\varepsilon^{\delta+\frac{1}{2}-s}+\varepsilon^\delta)\|U_{0,\varepsilon,\text{osc}}\|^2_{\dot{H}^{1/2+\delta}} + (\varepsilon^{2\delta+\frac{1}{2}-s}+\varepsilon^{(2-\beta)\delta+\frac{1}{2}-s})\|U_{0,\varepsilon,\text{osc}}\|^4_{\dot{H}^{1/2+\delta}}] \\
&\qquad\qquad \times \exp\{\mathbb{D}_0(1+\varepsilon^\delta\|U_{0,\varepsilon,\text{osc}}\|^2_{\dot{H}^{1/2+\delta}} + (\varepsilon^{\frac{\delta}{2}}\|U_{0,\varepsilon,\text{osc}}\|^2_{\dot{H}^{1/2+\delta}})^{\frac{2}{1-\alpha_2}})\}.
\end{aligned} \tag{2-49}$$

As $s \in \left[\frac{1}{2}, \frac{1}{2}+\eta\delta\right]$, we can write that

$$\begin{aligned}
&\|\delta_\varepsilon(t)\|^2_{\dot{H}^s} + \frac{\nu}{2}\int_0^t \|\nabla\delta_\varepsilon(\tau)\|^2_{\dot{H}^s}\,d\tau \\
&\qquad \le \mathbb{D}_0[\varepsilon^{2\alpha_0}+\varepsilon^{(1-\eta)\delta-2\gamma}+\varepsilon^{(2-\eta)\delta-4\gamma}+\varepsilon^{(2-\eta-\beta)\delta-4\gamma}]e^{\mathbb{D}_0(1+\varepsilon^{\delta-2\gamma})},
\end{aligned} \tag{2-50}$$

so that we need

$$\gamma < \min\big((1-\eta)\tfrac{\delta}{2}, \big(1-\tfrac{\eta}{2}\big)\tfrac{\delta}{2}, \big(1-\tfrac{\beta+\eta}{2}\big)\tfrac{\delta}{2}\big).$$

If we fix $\beta = \eta$, the condition is reduced to $\gamma < (1-\eta)\frac{\delta}{2}$, so that if $0 < \gamma < \frac{\delta}{2}$, we define $\eta = \frac{1}{2}\big(1-\frac{2\gamma}{\delta}\big)$ $\big($or equivalently $\gamma = (1-2\eta)\frac{\delta}{2}\big)$; then with $\beta = \eta$, for all $s \in \left[\frac{1}{2}, \frac{1}{2}+\eta\delta\right]$ and $t \le T_\varepsilon$, we end up with (as soon as $\varepsilon \le 1$)

$$\|\delta_\varepsilon(t)\|^2_{\dot{H}^s} + \frac{\nu}{2}\int_0^t \|\nabla\delta_\varepsilon(\tau)\|^2_{\dot{H}^s}\,d\tau \le \mathbb{D}_0 e^{2\mathbb{D}_0}\varepsilon^{2\min(\alpha_0,\frac{\eta\delta}{2})}. \tag{2-51}$$

We can now conclude the bootstrap argument: there exists $\varepsilon_0 > 0$ such that for any $0 < \varepsilon < \varepsilon_0$ the previous quantity is bounded by $\big(\frac{\nu}{8C}\big)^2$, so that $\big($in particular for $s = \frac{1}{2}\big)$ if we assume by contradiction that $T_\varepsilon < T_\varepsilon^*$, then

$$\|\delta_\varepsilon\|_{L^\infty_{T_\varepsilon}\dot{H}^{1/2}} \le \frac{\nu}{8C},$$

which contradicts the maximality of T_ε $\big($in this case, we would have $\|\delta_\varepsilon(T_\varepsilon)\|_{\dot{H}^{1/2}} = \frac{\nu}{4C}\big)$. Then $T_\varepsilon = T_\varepsilon^*$ and the previous estimates hold true for any $t < T_\varepsilon^*$, so that by the blowup criterion $T_\varepsilon = T_\varepsilon^* = \infty$ and the previous estimate is true for all $t \ge 0$ and all $s \in \left[\frac{1}{2}, \frac{1}{2}+\eta\delta\right]$:

$$\|\delta_\varepsilon\|_{\dot{E}^s} \le \mathbb{B}_0\varepsilon^{\min(\alpha_0,\frac{\eta\delta}{2})}.$$

Finally, to prove the last part of the theorem, we only have to remark that the previous argument is then true for any $s \in \left[\frac{1}{2}-\eta\delta, \frac{1}{2}+\eta\delta\right]$ when we require $\delta < \frac{3}{26+14\eta}$ (instead of $\delta < \frac{3}{26}$; see (2-46)), and use

Lemma 27:

$$\|\delta_\varepsilon\|_{L^2L^\infty} \le \|\delta_\varepsilon\|_{L^2\dot{B}^0_{\infty,1}} \le (\|\delta_\varepsilon\|_{L^2\dot{H}^{3/2-\eta\delta}}\|\delta_\varepsilon\|_{L^2\dot{H}^{3/2+\eta\delta}})^{\frac{1}{2}} \le \mathbb{B}_0\varepsilon^{\min(\alpha_0,\frac{\eta\delta}{2})}.$$

For $(d,p,r,q)=(0,2,\infty,1)$ and for all $\theta\in\left]0,\frac{1}{2}\right[$, from Proposition 47,

$$\|W_\varepsilon\|_{L^2L^\infty} \le \mathbb{C}_0\varepsilon^{\frac{\theta}{4}}\left(\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{B}^{1/2+\theta/2}_{2,1}} + \int_0^\infty \|G^b(\tau)\|_{\dot{B}^{1/2+\theta/2}_{2,1}}\,d\tau\right).$$

Using Lemma 27 with $(\alpha,\beta)=\left(\frac{\theta}{2},k\frac{\theta}{2}\right)$, and if $\theta=\frac{2\delta}{1+k}$ (for some small $k>0$),

$$\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{B}^{1/2+\theta/2}_{2,1}} \le \|U_{0,\varepsilon,\mathrm{osc}}\|^{\frac{k}{1+k}}_{\dot{H}^{1/2}}\|U_{0,\varepsilon,\mathrm{osc}}\|^{\frac{1}{1+k}}_{\dot{H}^{1/2+(1+k)(\theta/2)}} \le \mathbb{C}_0\varepsilon^{-\gamma}. \tag{2-52}$$

Choosing $k=\frac{\eta}{1-\eta}$, we get

$$\|W_\varepsilon\|_{L^2L^\infty} \le \mathbb{D}_0\varepsilon^{\frac{\delta}{2}(\frac{1}{1+k}-(1-2\eta))} = \mathbb{D}_0\varepsilon^{\frac{\eta\delta}{2}},$$

and the conclusion follows from the fact that $U_\varepsilon-\widetilde{U}_{\mathrm{QG}}=\delta_\varepsilon+W_\varepsilon$. □

Remark 32. Going back to (2-49), in the case $s=\frac{1}{2}$ if we only seek for global well-posedness, we retrieve here the same condition as in [Iwabuchi et al. 2017], except for the last term because Proposition 29 imposes $\beta>0$, so that the condition for global well-posedness is still $\gamma<\frac{\delta}{2}$. If β could reach zero, the conditions would be:

- $\|U_{0,\varepsilon,\mathrm{osc}}\|^2_{\dot{H}^{1/2}\cap\dot{H}^{1/2+\delta}}\varepsilon^\delta\le c$, with c some fixed small constant, if we want global well-posedness.
- $\|U_{0,\varepsilon,\mathrm{osc}}\|^2_{\dot{H}^{1/2}\cap\dot{H}^{1/2+\delta}}\varepsilon^\delta\to 0$ as $\varepsilon\to 0$ if we want δ_ε to go to zero.
- $\gamma<\frac{\delta}{2}$, if we want δ_ε to go to zero as a positive power of ε (which is what we originally searched for).

In our case, due to this $\beta>0$ these three conditions coincide.

2.3. *The general case.*

2.3.1. *Estimates on W_ε^T.* Let us begin by recalling the energy estimates for W_ε^T (we refer to Theorem 21 for m,M).

Proposition 33. *Assume $M<\frac{1}{2}$, there exist $\varepsilon_0=\varepsilon_0(\nu,\nu',M)>0$, and $\mathbb{B}_0=\mathbb{B}_0(\mathbb{C}_0,\nu,\nu',F,\delta)\ge 1$ such that for any $0<\varepsilon\le\varepsilon_0$ and $s\in\left[\frac{1}{2}-\frac{3}{2}\delta,\frac{1}{2}+\delta\right]$, we have*

$$\|W_\varepsilon^T\|^2_{L^\infty(\mathbb{R}_+,\dot{H}^s)} + \nu_0\|W_\varepsilon^T\|^2_{L^2(\mathbb{R}_+,\dot{H}^{s+1})} \le \mathbb{B}_0(\varepsilon^{-2\gamma}+1). \tag{2-53}$$

Proof. We know from [Charve 2004] that there exists a constant $C_F>0$ such that for any $s\in[0,1]$ and $t\in\mathbb{R}_+$, we have

$$\|W_\varepsilon^T\|_{\dot{E}^s_t} \le e^{\int_0^t\|G^b(\tau)\|_{\dot{H}^s}\,d\tau} \times\left(\|W_\varepsilon^T(0)\|^2_{\dot{H}^s}+C_F(1+\varepsilon R_\varepsilon^2|\nu-\nu'|)^2\int_0^t\left(\|G^b(\tau)\|_{\dot{H}^s}+\frac{1}{\nu_0}\|G^l(\tau)\|_{\dot{H}^{s-1}}\right)d\tau\right). \tag{2-54}$$

Combined with (A-86), Proposition 25 allows us to obtain that when $s \in \left[\frac{1}{2}-\frac{3}{2}\delta, \frac{1}{2}+\delta\right]$

$$\|W_\varepsilon^T\|_{\dot{E}_t^s} \leq C_F(1+|\nu-\nu'|\varepsilon R_\varepsilon^2)^2 e^{\frac{C_F}{\nu_0}C_{\delta,\nu_0}\mathbb{C}_0^{2+\frac{1}{\delta}}} \times\left(\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^s}^2+\left(\frac{1}{\nu_0}+\frac{|\nu-\nu'|^2}{\nu_0^2}\right)C_{\delta,\nu_0}\mathbb{C}_0^{2+\frac{1}{\delta}}\right). \tag{2-55}$$

We have $|\nu-\nu'|\varepsilon R_\varepsilon^2 \leq 1$ as soon as $M>\frac{1}{2}$ and $\varepsilon \leq \varepsilon_0 = |\nu-\nu'|^{\frac{-1}{1-2M}}$, which leads to (2-53). □

2.3.2. *Estimates on* δ_ε. As explained in the previous section (see also [Charve 2004; 2018a]), as $U_{0,\varepsilon} \in \dot{H}^s$ for all $s \in \left[\frac{1}{2}-\frac{3}{2}\delta, \frac{1}{2}+\delta\right]$, in particular it lies in $\dot{H}^{\frac{1}{2}}$ and thanks to the Fujita–Kato theorem there exists a unique local strong solution $U_\varepsilon \in L_T^\infty \dot{H}^{\frac{1}{2}} \cap L_T^2 \dot{H}^{\frac{3}{2}}$ for all $0<T<T_\varepsilon^*$, where $T_\varepsilon^*>0$ denotes the maximal lifespan. In addition, if T_ε^* is finite then we have

$$\int_0^{T_\varepsilon^*} \|\nabla U_\varepsilon(\tau)\|_{\dot{H}^{1/2}(\mathbb{R}^3)}^2\, d\tau = \infty.$$

Moreover, as our initial data enjoys additional regularity properties, they are transmitted to the solution: for all $s \in \left[\frac{1}{2}-\frac{3}{2}\delta, \frac{1}{2}+\delta\right]$ and $T<T_\varepsilon^*$, we have $U_\varepsilon \in L_T^\infty \dot{H}^s \cap L_T^2 \dot{H}^{s+1}$. As before, with a view to a bootstrap argument, let us now define

$$T_\varepsilon \overset{\text{def}}{=} \sup\left\{t \in [0, T_\varepsilon^*[: \text{for all } t' \leq t,\ \|\delta_\varepsilon(t')\|_{\dot{H}^{1/2}} \leq \frac{\nu}{4C}\right\}. \tag{2-56}$$

Thanks to (2-59), we are sure that $\|\delta_\varepsilon(0)\|_{\dot{H}^{1/2}} \leq \frac{\nu}{8C}$ (and then $T_\varepsilon>0$) if ε is small enough. Assuming that $T_\varepsilon<T_\varepsilon^*$, the computations from the previous case imply that, for all $s \in \left[\frac{1}{2}-\eta\delta, \frac{1}{2}+\eta\delta\right]$, and all $t \leq T_\varepsilon$,

$$\begin{aligned}
&\|\delta_\varepsilon(t)\|_{\dot{H}^s}^2+\frac{\nu_0}{2}\int_0^t \|\nabla\delta_\varepsilon(\tau)\|_{\dot{H}^s}^2\,d\tau \\
&\leq \Bigg[\|\delta_\varepsilon(0)\|_{\dot{H}^s}^2+\frac{C}{\nu_0}\Big(\nu_0\|f^b\|_{L^1\dot{H}^s}+\|f^l\|_{L^2\dot{H}^{s-1}}^2+\|\widetilde{U}_{\mathrm{QG}}\|_{L^\infty\dot{H}^s}^{2(1-\alpha_1)}\|\widetilde{U}_{\mathrm{QG}}\|_{L^2\dot{H}^{s+1}}^{2\alpha_1}\|W_\varepsilon^T\|_{L^{2/(1-\alpha_1)}L^{3/\alpha_1}}^2 \\
&\qquad+\|\widetilde{U}_{\mathrm{QG}}\|_{L^\infty\dot{H}^{1/2}}^2\||D|^{s+\alpha_1}W_\varepsilon^T\|_{L^2L^{6/(1+2\alpha_1)}}^2+\|\widetilde{U}_{\mathrm{QG}}\|_{L^\infty\dot{H}^s}^2\|W_\varepsilon^T\|_{L^2\dot{B}^{\alpha_1}_{3/\alpha_1,2}}^2 \\
&\qquad+\||D|^{s+\alpha_2}W_\varepsilon^T\|_{L^2L^{6/(1+2\alpha_2)}}^2+\|W_\varepsilon^T\|_{L^\infty\dot{H}^s}^{2(1-\alpha_3)}\|W_\varepsilon^T\|_{L^2\dot{H}^{s+1}}^{2\alpha_3}\|W_\varepsilon^T\|_{L^{2/(1-\alpha_3)}L^{3/\alpha_3}}^2 \\
&\qquad+\|W_\varepsilon^T\|_{L^\infty\dot{H}^s}^2\|W_\varepsilon^T\|_{L^2\dot{B}^{\alpha_3}_{3/\alpha_3,2}}^2\Big)\Bigg] \\
&\quad\times\exp\frac{C}{\nu_0}\Bigg\{\nu_0\|f^b\|_{L^1\dot{H}^s}+\|\nabla\widetilde{U}_{\mathrm{QG}}\|_{L^2\dot{H}^{1/2}}^2\left(1+\frac{1}{\nu_0^2}\|\widetilde{U}_{\mathrm{QG}}\|_{L^\infty\dot{H}^{1/2}}^2\right) \\
&\qquad+\frac{1}{\nu_0^{\frac{2\alpha_2}{1-\alpha_2}}}\|W_\varepsilon^T\|_{L^{2/(1-\alpha_2)}L^{3/\alpha_2}}^{\frac{2}{1-\alpha_2}}+\|W_\varepsilon^T\|_{L^2\dot{B}^{\alpha_2}_{3/\alpha_2,2}}^2\Bigg\}.
\end{aligned} \tag{2-57}$$

Compared to (2-40), the only differences are:

- The force terms $f^{b,l}$ (dealt with as in [Charve 2004; 2006]).

- The simpler estimates for F_8: as precision will be imposed by the truncated terms, we only write

$$|(F_8|\delta_\varepsilon)_{\dot H^s}| \le \frac{\nu_0}{16}\|\delta_\varepsilon\|^2_{\dot H^{s+1}} + \frac{C}{\nu_0}\big(\|W_\varepsilon^T\|^{2(1-\alpha_3)}_{L^\infty \dot H^s}\|W_\varepsilon^T\|^{2\alpha_3}_{L^2\dot H^{s+1}}\|W_\varepsilon\|^2_{L^{2/(1-\alpha_3)}L^{3/\alpha_3}} + \|W_\varepsilon\|^2_{L^\infty \dot H^s}\|W_\varepsilon\|^2_{L^2\dot B^{\alpha_3}_{3/\alpha_3,2}}\big). \tag{2-58}$$

2.3.3. *Estimates for the truncated quantities.* We will now bound much more precisely than in [Charve 2004; 2006] the external force terms and initial data (see (1-9)):

Proposition 34. *There exists a constant $\mathbb{B}_0 \ge 1$ such that for all $s \in \big[\frac12 - \eta\delta, \frac12 + \eta\delta\big]$,*

$$\begin{aligned}
\|f^b\|_{L^1\dot H^s} &\le \mathbb{B}_0\big(\varepsilon^{1-2M} + \varepsilon^{M(1-\eta)\delta} + \varepsilon^{\frac{m}{6}-M(\frac56+\eta\delta)}\big),\\
\|f^l\|_{L^2\dot H^{s-1}} &\le \mathbb{B}_0\big(\varepsilon^{1-2M} + \varepsilon^{M(1-\eta)\delta} + \varepsilon^{m(\frac12-\eta\delta)}\big),\\
\|\delta_\varepsilon(0)\|_{\dot H^s} &\le \mathbb{B}_0\big(\varepsilon^{\alpha_0} + \varepsilon^{1-2M-\gamma} + \varepsilon^{\delta(M-\eta m)-\gamma} + \varepsilon^{\delta((\frac12-\eta)m-M)-\gamma}\big).
\end{aligned} \tag{2-59}$$

Remark 35. Note that as we want positive powers of ε, the previous estimates imply the conditions

$$\begin{aligned}
&M, \eta, \eta\delta \in \big]0, \tfrac12\big[,\\
&\eta < \frac{M}{m} < \min\big(\tfrac{1}{5+6\eta\delta}, \tfrac12 - \eta\big),\\
&\gamma < \min\big(1-2M, \delta(M-\eta m), \delta\big((\tfrac12-\eta)m - M\big)\big).
\end{aligned} \tag{2-60}$$

Proof of Proposition 34. Let us begin with the terms involving G: thanks to (A-84), and Remark 10 and Proposition 25, we immediately obtain that there exists a constant $\mathbb{B}_0$ (only depending on $\mathbb{C}_0, \nu, \nu'$ and F) such that for all $s \in \big[\frac12 - \eta\delta, \frac12 + \eta\delta\big]$

$$\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}\mathbb{P}_2 G^b\|_{L^1\dot H^s} + \|\mathcal{P}_{r_\varepsilon,R_\varepsilon}\mathbb{P}_2 G^l\|_{L^2\dot H^{s-1}} \le \mathbb{B}_0 \varepsilon R_\varepsilon^2.$$

Thanks to Lemma 42 (see the Appendix), Proposition 25 and (2-17), the second term in f_1 can be bounded (for all $s \in \big[\frac12 - \eta\delta, \frac12 + \eta\delta\big]$) according to

$$\begin{aligned}
\|(\mathrm{Id}-\mathcal{P}_{r_\varepsilon,R_\varepsilon})G^b\|_{L^1\dot H^s} &\le \left\|\left(\mathrm{Id}-\chi\left(\frac{|D|}{R_\varepsilon}\right)\right)G^b\right\|_{L^1\dot H^s} + \left\|\chi\left(\frac{|D|}{R_\varepsilon}\right)\chi\left(\frac{|D_3|}{r_\varepsilon}\right)G^b\right\|_{L^1\dot H^s}\\
&\le \frac{1}{R_\varepsilon^{\frac12+\delta-s}}\left\|\left(\mathrm{Id}-\chi\left(\frac{|D|}{R_\varepsilon}\right)\right)G^b\right\|_{L^1\dot H^{1/2+\delta}} + R_\varepsilon^s\left\|\chi\left(\frac{|D|}{R_\varepsilon}\right)\chi\left(\frac{|D_3|}{r_\varepsilon}\right)G^b\right\|_{L^1L^2}\\
&\le \frac{1}{R_\varepsilon^{\frac12+\delta-s}}\|G^b\|_{L^1\dot H^{1/2+\delta}} + R_\varepsilon^s(R_\varepsilon^2 r_\varepsilon)^{\frac23-\frac12}\|\widetilde U_{\mathrm{QG}}\cdot\nabla\widetilde U_{\mathrm{QG}}\|_{L^1L^{3/2}}\\
&\le \frac{1}{R_\varepsilon^{\frac12+\delta-s}}\|G^b\|_{L^1\dot H^{1/2+\delta}} + R_\varepsilon^{s+\frac13}r_\varepsilon^{\frac16}\int_0^\infty \|\widetilde U_{\mathrm{QG}}(\tau)\|_{L^6}\|\nabla\widetilde U_{\mathrm{QG}}(\tau)\|_{L^2}\,d\tau\\
&\le \mathbb{B}_0\left(\frac{1}{R_\varepsilon^{\frac12+\delta-s}} + R_\varepsilon^{s+\frac13}r_\varepsilon^{\frac16}\right),
\end{aligned} \tag{2-61}$$

which implies the first estimates in (2-59) for all $s \in \left[\frac{1}{2}-\eta\delta, \frac{1}{2}+\eta\delta\right]$. Similarly, we have

$$\left\|\left(\mathrm{Id}-\chi\left(\frac{|D|}{R_\varepsilon}\right)\right)G^l\right\|_{L^2\dot{H}^{s-1}}^2 \le \frac{\mathbb{B}_0}{R_\varepsilon^{2(\frac{1}{2}+\delta-s)}}, \tag{2-62}$$

and using that the expression of G^l (see (1-9)) features some derivative ∂_3, for all $s \in \left[\frac{1}{2}-\eta\delta, \frac{1}{2}+\eta\delta\right]$ we have

$$\begin{aligned}\left\|\chi\left(\frac{|D|}{R_\varepsilon}\right)\chi\left(\frac{|D_3|}{r_\varepsilon}\right)G^l\right\|_{L^2\dot{H}^{s-1}} &\le C_F|\nu-\nu'|\left\|\chi\left(\frac{|D|}{R_\varepsilon}\right)\chi\left(\frac{|D_3|}{r_\varepsilon}\right)\partial_3\nabla\widetilde{U}_{\mathrm{QG}}\right\|_{L^2\dot{H}^{s-1}}\\ &\le C_F|\nu-\nu'|r_\varepsilon^s\|\partial_3^{1-s}\nabla\widetilde{U}_{\mathrm{QG}}\|_{L^2\dot{H}^{s-1}}\\ &\le C_F|\nu-\nu'|r_\varepsilon^s\|\widetilde{U}_{\mathrm{QG}}\|_{L^2\dot{H}^1}.\end{aligned} \tag{2-63}$$

Let us now turn to bound the initial data $\delta_\varepsilon(0)$:

$$\begin{aligned}\|\delta_\varepsilon(0)\|_{\dot{H}^s} &\le \|U_{0,\varepsilon,\mathrm{QG}}-\widetilde{U}_{0,\mathrm{QG}}\|_{\dot{H}^s}+\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}\mathbb{P}_2U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^s}+\|(\mathrm{Id}-\mathcal{P}_{r_\varepsilon,R_\varepsilon})U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^s}\\ &\le \mathbb{C}_0\varepsilon^{\alpha_0}+\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}\mathbb{P}_2U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^s}\\ &\qquad+\left\|\left(\mathrm{Id}-\chi\left(\frac{|D|}{R_\varepsilon}\right)\right)U_{0,\varepsilon,\mathrm{osc}}\right\|_{\dot{H}^s}+\left\|\chi\left(\frac{|D|}{R_\varepsilon}\right)\chi\left(\frac{|D_3|}{r_\varepsilon}\right)U_{0,\varepsilon,\mathrm{osc}}\right\|_{\dot{H}^s}.\end{aligned} \tag{2-64}$$

As before, we easily estimate the second and third terms for all $s \in \left[\frac{1}{2}-\eta\delta, \frac{1}{2}+\eta\delta\right]$ by

$$C_F|\nu-\nu'|\varepsilon R_\varepsilon^2\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^s}+\frac{C_F}{R_\varepsilon^{\frac{1}{2}+\delta-s}}\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^{1/2+\delta}} \le \mathbb{B}_0\varepsilon^{-\gamma}\left[\varepsilon R_\varepsilon^2+\frac{1}{R_\varepsilon^{\frac{1}{2}+\delta-s}}\right]. \tag{2-65}$$

It is here that the $\dot{B}_{q,q}^{\frac{1}{2}}$-assumption on the initial data will be specifically used (everywhere else we only use the fact that this space is embedded in $\dot{H}^{\frac{1}{2}-\frac{3}{2}\delta}$). To bound the last term, thanks to Proposition 40 let us write that (we recall that $q=\frac{2}{1+\delta}<2$)

$$\begin{aligned}\left\|\chi\left(\frac{|D|}{R_\varepsilon}\right)\chi\left(\frac{|D_3|}{r_\varepsilon}\right)U_{0,\varepsilon,\mathrm{osc}}\right\|_{\dot{H}^s} &= \left\|\chi\left(\frac{|D|}{R_\varepsilon}\right)\chi\left(\frac{|D_3|}{r_\varepsilon}\right)|D|^sU_{0,\varepsilon,\mathrm{osc}}\right\|_{L^2}\\ &\le C(R_\varepsilon^2r_\varepsilon)^{\frac{1}{q}-\frac{1}{2}}R_\varepsilon^{s-\frac{1}{2}}\left\|\chi\left(\frac{|D|}{R_\varepsilon}\right)\chi\left(\frac{|D_3|}{r_\varepsilon}\right)|D|^{\frac{1}{2}}U_{0,\varepsilon,\mathrm{osc}}\right\|_{L^q}\\ &\le CR_\varepsilon^{\delta+s-\frac{1}{2}}r_\varepsilon^{\frac{\delta}{2}}\||D|^{\frac{1}{2}}U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{B}_{q,q}^0}\\ &\le CR_\varepsilon^{\delta+s-\frac{1}{2}}r_\varepsilon^{\frac{\delta}{2}}\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{B}_{q,q}^{1/2}}.\end{aligned} \tag{2-66}$$

Note that this can be done only if $s \ge \frac{1}{2}$. In the case $s \in \left[\frac{1}{2}-\eta\delta, \frac{1}{2}\right[$, we simply go back to (2-64) and write that (taking advantage of the frequency localization)

$$\begin{aligned}\|(\mathrm{Id}-\mathcal{P}_{r_\varepsilon,R_\varepsilon})U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^s} &\le \frac{1}{r_\varepsilon^{\frac{1}{2}-s}}\|(\mathrm{Id}-\mathcal{P}_{r_\varepsilon,R_\varepsilon})U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^{1/2}}\\ &\le \frac{C_F}{r_\varepsilon^{\frac{1}{2}-s}}\left(\frac{1}{R_\varepsilon^\delta}\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^{1/2+\delta}}+R_\varepsilon^\delta r_\varepsilon^{\frac{\delta}{2}}\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{B}_{q,q}^{1/2}}\right).\end{aligned} \tag{2-67}$$

We can sum up as follows: for all $s \in \left[\frac{1}{2}-\eta\delta, \frac{1}{2}+\eta\delta\right]$

$$\|(\mathrm{Id}-\mathcal{P}_{r_\varepsilon,R_\varepsilon})U_{0,\varepsilon,\mathrm{osc}}\|_{\dot H^s} \le \mathbb{C}_0\varepsilon^{-\gamma}\times\begin{cases}\dfrac{1}{R_\varepsilon^{(1-\eta)\delta}}+R_\varepsilon^{(1+\eta)\delta}r_\varepsilon^{\frac{\delta}{2}} & \text{if } s\in\left[\frac12,\frac12+\eta\delta\right],\\ \dfrac{1}{r_\varepsilon^{\eta\delta}}\Big(\dfrac{1}{R_\varepsilon^\delta}+R_\varepsilon^\delta r_\varepsilon^{\frac{\delta}{2}}\Big) & \text{if } s\in\left[\frac12-\eta\delta,\frac12\right]\end{cases}$$

$$\le \mathbb{C}_0\varepsilon^{-\gamma}\times\begin{cases}\varepsilon^{M\delta(1-\eta)}+\varepsilon^{\delta(\frac{m}{2}-(1+\eta)M)} & \text{if } s\in\left[\frac12,\frac12+\eta\delta\right],\\ \varepsilon^{\delta(M-m\eta)}+\varepsilon^{\delta((\frac12-\eta)m-M)} & \text{if } s\in\left[\frac12-\eta\delta,\frac12\right].\end{cases} \tag{2-68}$$

As

$$M(1-\eta)-(M-m\eta)=\eta(m-M), \quad \text{and} \quad \Big(\frac{m}{2}-(1+\eta)M\Big)-\Big(\big(\tfrac12-\eta\big)m-M\Big)=\eta(m-M),$$

and as $m>M$ (see (2-60)), we obtain the announced result. ☐

2.3.4. *Strichartz estimates for W_ε^T.* We will need the following Strichartz estimates to complete our bootstrap argument:

Proposition 36. *There exist $\varepsilon_0, \mathbb{B}_0>0$ such that for any $\alpha>0$ and $\varepsilon<\varepsilon_0$, W_ε^T satisfies*

$$\begin{aligned}\|W_\varepsilon^T\|_{\widetilde L^2\dot B^\alpha_{3/\alpha,2}} &\le \mathbb{B}_0\varepsilon^{\frac14-\frac{\alpha}{3}-M(\frac92-4\alpha-\delta)-m(\frac92-2\alpha)}\le\mathbb{B}_0\varepsilon^{\frac14-\frac{\alpha}{3}-\frac92(M+m)},\\ \|W_\varepsilon^T\|_{L^{2/(1-\alpha)}L^{3/\alpha}} &\le \mathbb{B}_0\varepsilon^{\frac14-\frac{\alpha}{3}-M(\frac92-3\alpha-\delta)-m(\frac92-3\alpha)}\le\mathbb{B}_0\varepsilon^{\frac14-\frac{\alpha}{3}-\frac92(M+m)},\\ \||D|^{s+\alpha}W_\varepsilon^T\|_{L^2L^{6/(1+2\alpha)}} &\le \mathbb{B}_0\varepsilon^{\frac{1}{12}-\frac{\alpha}{3}-M(\frac72-3\alpha)-m(\frac72-2\alpha)}\le\mathbb{B}_0\varepsilon^{\frac{1}{12}-\frac{\alpha}{3}-\frac72(M+m)}.\end{aligned}\tag{2-69}$$

Proof. Using Proposition 51 in the case $(d,p,r,q)=\big(\alpha,2,\frac{3}{\alpha},2\big)$, we obtain that

$$\begin{aligned}&\|W_\varepsilon^T\|_{\widetilde L^2\dot B^\alpha_{3/\alpha,2}}\\ &\le \mathbb{B}_0\varepsilon^{\frac14(1-\frac{4\alpha}{3})}\frac{R_\varepsilon^{4-3\alpha}}{r_\varepsilon^{\frac72-2\alpha}}\Big(\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}U_{0,\varepsilon,\mathrm{osc}}\|_{\dot H^\alpha}+\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}G^b\|_{L^1\dot H^\alpha}+\frac{1}{\nu_0^{\frac12}r_\varepsilon}\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}G^l\|_{L^2\dot H^\alpha}\Big)\\ &\le \mathbb{B}_0\varepsilon^{\frac14(1-\frac{4\alpha}{3})}\frac{R_\varepsilon^{4-3\alpha}}{r_\varepsilon^{\frac72-2\alpha}}\times\Big(\frac{1}{r_\varepsilon^{\frac12-\frac{3\delta}{2}-\alpha}}\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot H^{1/2-3\delta/2}}+\|G^b\|_{L^1\dot H^\alpha}+\frac{1}{\nu_0^{\frac12}r_\varepsilon}R_\varepsilon^{\frac12-\delta-\alpha}\|G^l\|_{L^2\dot H^{1/2+\delta}}\Big)\\ &\le \mathbb{B}_0\varepsilon^{\frac14-\frac{\alpha}{3}-M(4-3\alpha)-m(\frac72-2\alpha)}\big(\varepsilon^{-\gamma-m(\frac12-\frac{3\delta}{2}-\alpha)}+\varepsilon^{-m-M(\frac12-\delta-\alpha)}\big).\end{aligned}\tag{2-70}$$

From (2-60), we know that $\gamma<\delta M$ so that

$$m+M\big(\tfrac12-\delta-\alpha\big)-\big(\gamma+m\big(\tfrac12-\tfrac{3\delta}{2}-\alpha\big)\big)=M\big(\tfrac12-\delta-\alpha\big)+m\big(\tfrac12+\tfrac{3\delta}{2}+\alpha\big)-\gamma>0,$$

which leads to the first estimate. Similarly, considering Proposition 51 in the case $(d,p,r,q)=\big(0,\frac{2}{1-\alpha},\frac{3}{\alpha},2\big)$, we get (thanks to Proposition 40)

$$\begin{aligned}\|W_\varepsilon^T\|_{L^{2/(1-\alpha)}L^{3/\alpha}} &\le \|W_\varepsilon^T\|_{\widetilde L^2\dot B^0_{3/\alpha,2}}\\ &\le \mathbb{B}_0\varepsilon^{\frac14(1-\frac{4\alpha}{3})}\frac{R_\varepsilon^{4-3\alpha}}{r_\varepsilon^{\frac72-3\alpha}}\Big(\frac{1}{r_\varepsilon^{\frac12-\frac{3\delta}{2}}}\|U_{0,\varepsilon,\mathrm{osc}}\|_{\dot H^{1/2-3\delta/2}}+\|G^b\|_{L^1L^2}\\ &\qquad+\frac{1}{\nu_0^{\frac12}r_\varepsilon}R_\varepsilon^{\frac12-\delta}\|G^l\|_{L^2\dot H^{1/2+\delta}}\Big)\\ &\le \mathbb{B}_0\varepsilon^{\frac14-\frac{\alpha}{3}-M(4-3\alpha)-m(\frac72-3\alpha)}\big(\varepsilon^{-\gamma-m(\frac12-\frac{3\delta}{2})}+\varepsilon^{-m-M(\frac12-\delta)}\big),\end{aligned}\tag{2-71}$$

which leads to the second estimate. In the case $(d,p,r,q)=\big(s+\alpha,2,\frac{6}{1+2\alpha},2\big)$, we obtain that (provided $0<\alpha<\delta+\frac12-s$)

$$\begin{aligned}
\||D|^{s+\alpha}W_\varepsilon^T\|_{L^2L^{6/(1+2\alpha)}} &\le \|W_\varepsilon^T\|_{\widetilde{L}^2\dot{B}^{s+\alpha}_{6/(1+2\alpha),2}}\\
&\le \mathbb{B}_0\varepsilon^{\frac{1}{12}(1-4\alpha)}\frac{R_\varepsilon^{\frac52-3\alpha}}{r_\varepsilon^{\frac52-2\alpha}}\Big(\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}U_{0,\varepsilon,\mathrm{osc}}\|_{\dot{H}^{s+\alpha}}+\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}G^b\|_{L^1\dot{H}^{s+\alpha}}\\
&\qquad\qquad+\frac{1}{\nu_0^{\frac12}r_\varepsilon}\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}G^l\|_{L^2\dot{H}^{s+\alpha}}\Big)\\
&\le \mathbb{B}_0\varepsilon^{\frac{1}{12}(1-4\alpha)}\frac{R_\varepsilon^{\frac52-3\alpha}}{r_\varepsilon^{\frac52-2\alpha}}\Big(\varepsilon^{-\gamma}+1+\frac{1}{r_\varepsilon}R_\varepsilon^{s+\alpha+\frac12-\delta}\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}\widetilde{U}_{\mathrm{QG}}\|_{L^2\dot{H}^{3/2+\delta}}\Big)\\
&\le \mathbb{B}_0\varepsilon^{\frac{1}{12}-\frac{\alpha}{3}-M(\frac52-3\alpha)-m(\frac52-2\alpha)}(\varepsilon^{-\gamma}+\varepsilon^{-m-M}),
\end{aligned}\tag{2-72}$$

which concludes the proof. □

2.3.5. *Bootstrap.* We are now able to conclude the bootstrap argument (see the previous section and [Charve 2004; 2006]). Gathering (2-57), (2-59) and (2-69), we obtain that for all $t\le T_\varepsilon$

$$\begin{aligned}
&\|\delta_\varepsilon(t)\|^2_{\dot{H}^s}+\frac{\nu_0}{2}\int_0^t\|\nabla\delta_\varepsilon(\tau)\|^2_{\dot{H}^s}\,d\tau\\
&\quad\le \mathbb{D}_0\big[\varepsilon^{2\alpha_0}+\varepsilon^{2(1-2M-\gamma)}+\varepsilon^{2(\delta(M-\eta m)-\gamma)}+\varepsilon^{2(\delta((\frac12-\eta)m-M)-\gamma)}+\varepsilon^{1-2M}\\
&\qquad+\varepsilon^{M(1-\eta)\delta}+\varepsilon^{\frac{m}{6}-M(\frac56+\eta\delta)}+\varepsilon^{2(1-2M)}+\varepsilon^{2M(1-\eta)\delta}\\
&\qquad\qquad+\varepsilon^{2m(\frac12-\eta\delta)}+\varepsilon^{\frac14-\frac{\alpha}{3}-\frac92(M+m)-\gamma}+\varepsilon^{\frac{1}{12}-\frac{\alpha}{3}-\frac72(M+m)}\big]\\
&\quad\times\exp\frac{C}{\nu_0}\{1+\varepsilon^{1-2M}+\varepsilon^{M(1-\eta)\delta}+\varepsilon^{\frac{m}{6}-M(\frac56+\eta\delta)}+\varepsilon^{\min(2,\frac{2}{1-\alpha})(\frac14-\frac{\alpha}{3}-\frac92(M+m))}\}.
\end{aligned}\tag{2-73}$$

For simplicity we will require, instead of the second condition from (2-60), that

$$2\eta\le\frac{M}{m}\le\frac12\min\big(\tfrac{1}{5+6\eta\delta},\tfrac12-\eta\big).$$

This obviously implies that $\eta\le\frac{1}{10}$, so we will finally ask that

$$\begin{gathered}
M\in\left]0,\tfrac14\right[,\quad \eta\in\left]0,\tfrac{1}{10}\right[,\\
2\eta\le\tfrac{M}{m}\le\tfrac12\tfrac{1}{5+\delta},\\
\gamma<\min\big(\tfrac12(1-2M),\tfrac12\delta(M-\eta m),\tfrac12\delta\big((\tfrac12-\eta)m-M\big)\big).
\end{gathered}\tag{2-74}$$

Moreover, if we take $\alpha=\gamma$ and ask that

$$\begin{aligned}
&\tfrac92(M+m)\le\tfrac18 &&\text{and}\quad \tfrac43\delta\le\tfrac12\big(\tfrac14-\tfrac92(M+m)\big),\\
&\tfrac72(M+m)\le\tfrac{1}{24} &&\text{and}\quad \tfrac{\delta}{3}\le\tfrac12\big(\tfrac{1}{12}-\tfrac72(M+m)\big).
\end{aligned}\tag{2-75}$$

As $M\le\frac{m}{10}$, this is realized when

$$m\in\left]0,\tfrac{1}{100}\right[,\quad 2\eta\le\tfrac{M}{m}\le\tfrac12\tfrac{1}{5+\delta}.$$

When

$$\gamma \le \min\left(\tfrac{M\delta}{4}, \tfrac{m\delta}{16}, \tfrac{m}{12}, \tfrac{1}{32}\right) = \tfrac{M\delta}{4},$$

we obtain that all powers of ε in the exponential are positive so that for small enough ε, we get that for all $s \in \left[\frac{1}{2} - \eta\delta, \frac{1}{2} + \eta\delta\right]$ and $t \le T_\varepsilon$

$$\|\delta_\varepsilon(t)\|_{\dot{H}^s}^2 + \frac{\nu_0}{2}\int_0^t \|\nabla\delta_\varepsilon(\tau)\|_{\dot{H}^s}^2 \, d\tau \le \mathbb{D}_0 e^{2\mathbb{D}_0} \varepsilon^{\min(2\alpha_0, \frac{M\delta}{2})}, \tag{2-76}$$

so that we finally end up with (for small enough ε), $\delta_\varepsilon(T_\varepsilon) \le \frac{\nu_0}{8C}$, which clearly contradicts the maximality of T_ε. We can conclude that $T_\varepsilon = T_\varepsilon^*$ and then the previous estimate is valid for all $t < T_\varepsilon^*$, which implies for $s = \frac{1}{2}$ that the integral in (1-1) is finite. Therefore $T_\varepsilon^* = \infty$ and (2-76) is then valid for all $t \ge 0$. The rest of the theorem is done as for the case $\nu = \nu'$. □

Appendix

A.1. *Notation and Sobolev spaces.* For $s \in \mathbb{R}$, $\dot{H}^s$ and H^s are the classical homogeneous/inhomogeneous Sobolev spaces in $\mathbb{R}^3$ endowed with the norms

$$\|u\|_{\dot{H}^s}^2 = \int_{\mathbb{R}^3} |\xi|^{2s} |\hat{u}(\xi)| \, d\xi \quad \text{and} \quad \|u\|_{H^s}^2 = \int_{\mathbb{R}^3} (1 + |\xi|^2)^s |\hat{u}(\xi)| \, d\xi.$$

We also use the following notation: if E is a Banach space and $T > 0$,

$$\mathcal{C}_T E = \mathcal{C}([0, T], E) \quad \text{and} \quad L_T^p E = L^p([0, T], E).$$

Let us recall the Sobolev injections, and product laws:

Proposition 37. *There exists a constant $C > 0$ such that if $s < \frac{3}{2}$, then for any $u \in \dot{H}^s(\mathbb{R}^3)$, we have $u \in L^p(\mathbb{R}^3)$ with $p = \frac{6}{3-2s}$ and*

$$\|u\|_{L^p} \le C\|u\|_{\dot{H}^s}.$$

Proposition 38 [Bahouri et al. 2011, Chapter 2]. *There exists a constant C such that for any $(u, v) \in \dot{H}^{s_1}(\mathbb{R}^3) \times \dot{H}^{s_2}(\mathbb{R}^3)$, if $s_1, s_2 \in \left]-\frac{3}{2}, \frac{3}{2}\right[$ and $s_1 + s_2 > 0$ then $uv \in \dot{H}^{s_1+s_2-\frac{3}{2}}(\mathbb{R}^3)$ and we have*

$$\|uv\|_{\dot{H}^{s_1+s_2-3/2}} \le C\|u\|_{\dot{H}^{s_1}}\|v\|_{\dot{H}^{s_2}}.$$

A.2. *Besov spaces.* We refer to Chapter 2 from [Bahouri et al. 2011] for an in-depth presentation of the classical homogeneous and inhomogeneous Besov and Sobolev spaces. We also refer to the appendix of [Charve 2018a] for a quick presentation.

Let us just recall that ψ is a smooth radial function supported in the ball $B\left(0, \frac{4}{3}\right)$, equal to 1 in a neighborhood of $B\left(0, \frac{3}{4}\right)$ and such that $r \mapsto \psi(r \cdot e_r)$ is nonincreasing over $\mathbb{R}_+$. If we set $\varphi(\xi) = \psi\left(\frac{\xi}{2}\right) - \psi(\xi)$, then φ is compactly supported in the annulus $\mathcal{C} = \left\{\xi \in \mathbb{R}^d : c_0 = \frac{3}{4} \le |\xi| \le C_0 = \frac{8}{3}\right\}$ and we define the homogeneous dyadic blocks: for all $j \in \mathbb{Z}$,

$$\dot{\Delta}_j u := \varphi(2^{-j} D)u = 2^{jd} h(2^j \cdot) * u, \quad \text{with } h = \mathcal{F}^{-1}\varphi.$$

We recall that $\widehat{\phi(D)u}(\xi) = \phi(\xi)\hat{u}(\xi)$ and we can define the homogeneous Besov norms and spaces:

Definition 39. For $s \in \mathbb{R}$ and $1 \le p, r \le \infty$, we set

$$\|u\|_{\dot{B}^s_{p,r}} := \Big(\sum_{l\in\mathbb{Z}} 2^{rls}\|\dot{\Delta}_l u\|^r_{L^p}\Big)^{\frac{1}{r}} \quad \text{if } r<\infty \qquad \text{and} \qquad \|u\|_{\dot{B}^s_{p,\infty}} := \sup_l 2^{ls}\|\dot{\Delta}_l u\|_{L^p}.$$

The homogeneous Besov space $\dot{B}^s_{p,r}$ is the subset of tempered distributions such that

$$\lim_{j\to-\infty}\|\dot{S}_j u\|_{L^\infty} = 0$$

and $\|u\|_{\dot{B}^s_{p,r}}$ is finite (where $\dot{S}_j u = \sum_{l\le j-1}\dot{\Delta}_l u = \psi(2^{-j}D)u$).

- The space $\dot{B}^s_{p,r}$ is complete whenever $s < d/p$, or $s \le d/p$ and $r = 1$.
- For any $p \in [1,\infty]$, we have the continuous embedding $\dot{B}^0_{p,1} \hookrightarrow L^p \hookrightarrow \dot{B}^0_{p,\infty}$.
- If $\sigma\in\mathbb{R}$, $1\le p_1\le p_2\le\infty$, and $1\le r_1\le r_2\le\infty$, then we have $\dot{B}^\sigma_{p_1,r_1} \hookrightarrow \dot{B}^{\sigma-d(\frac{1}{p_1}-\frac{1}{p_2})}_{p_2,r_2}$.
- The space $\dot{B}^{\frac{d}{p}}_{p,1}$ is continuously embedded in the set of bounded continuous functions (going to 0 at infinity if $p<\infty$).
- $\dot{H}^s = \dot{B}^s_{2,2}$.
- Interpolation: if $1\le p, r_1, r_2, r\le\infty$, $\sigma_1\ne\sigma_2$, and $\theta\in(0,1)$,

$$\|f\|_{\dot{B}^{\theta\sigma_2+(1-\theta)\sigma_1}_{p,r}} \lesssim \|f\|^{1-\theta}_{\dot{B}^{\sigma_1}_{p,r_1}}\,\|f\|^{\theta}_{\dot{B}^{\sigma_2}_{p,r_2}}. \tag{A-77}$$

Proposition 40 [Bahouri et al. 2011]. *We have the following continuous injections*:

- $\dot{B}^0_{p,1} \hookrightarrow L^p$ *for any* $p \ge 1$.
- $\dot{B}^0_{p,2} \hookrightarrow L^p$ *for any* $p\in[2,\infty[$.
- $\dot{B}^0_{p,p} \hookrightarrow L^p$ *for any* $p\in[1,2]$.

Let us then define the spaces $\widetilde{L}^\rho_T\dot{B}^s_{p,r}$ from the following norm:

Definition 41. For $T>0$, $s\in\mathbb{R}$, and $1\le r,\rho\le\infty$, we set

$$\|u\|_{\widetilde{L}^\rho_T\dot{B}^s_{p,r}} := \big\|2^{js}\|\dot{\Delta}_q u\|_{L^\rho_T L^p}\big\|_{\ell^r(\mathbb{Z})}.$$

Any product of two distributions u and v may be formally written through the Bony decomposition:

$$uv = T_u v + T_v u + R(u,v), \tag{A-78}$$

where

$$T_u v := \sum_l \dot{S}_{l-1}u\dot{\Delta}_l v, \quad T_v u := \sum_l \dot{S}_{l-1}v\dot{\Delta}_l u \quad \text{and} \quad R(u,v) := \sum_l\sum_{|l'-l|\le 1}\dot{\Delta}_l u\,\dot{\Delta}_{l'}v.$$

The above operator T is called a "paraproduct", whereas R is called a "remainder". We refer to [Bahouri et al. 2011] for general properties and for paraproduct and remainder estimates but we can recall that (if $\frac{1}{r}=\frac{1}{r_1}+\frac{1}{r_2}$ and $\frac{1}{p}=\frac{1}{p_1}+\frac{1}{p_2}$):

- For any $s\in\mathbb{R}$, we have $\|T_u v\|_{\dot{B}^s_{p,r}} \lesssim \|u\|_{L^\infty}\|v\|_{\dot{B}^s_{p,r}}$.

- For any $(s,t) \in \mathbb{R}_-^* \times \mathbb{R}$, we have $\|T_u v\|_{\dot{B}^{s+t}_{p,r}} \lesssim \|u\|_{\dot{B}^s_{p_1,r_1}} \|v\|_{\dot{B}^t_{p_2,r_2}}$.
- For any $s,t \in \mathbb{R}$ with $s+t>0$, we have $\|R(u,v)\|_{\dot{B}^{s+t}_{p,r}} \lesssim \|u\|_{\dot{B}^s_{p_1,r_1}} \|v\|_{\dot{B}^t_{p_2,r_2}}$.

A.3. ***Dispersion and Strichartz estimates.*** Consider the system

$$\begin{cases} \partial_t f - \left(L - \frac{1}{\varepsilon}\mathbb{P}\mathcal{A}\right) f = F_{\text{ext}}, \\ f_{|t=0} = f_0. \end{cases} \tag{A-79}$$

If we apply the Fourier transform, the equation becomes (see [Charve 2005] for details)

$$\partial_t \hat{f} - \mathbb{B}(\xi,\varepsilon)\hat{f} = \widehat{F}_{\text{ext}},$$

where

$$\mathbb{B}(\xi,\varepsilon) = \widehat{L - \frac{1}{\varepsilon}\mathbb{P}\mathcal{A}} = \begin{pmatrix} -\nu|\xi|^2 + \xi_1\xi_2/(\varepsilon|\xi|^2) & (\xi_2^2+\xi_3^2)/(\varepsilon|\xi|^2) & 0 & \xi_1\xi_3/(\varepsilon F|\xi|^2) \\ -(\xi_1^2+\xi_3^2)/(\varepsilon|\xi|^2) & -\nu|\xi|^2 - \xi_1\xi_2/(\varepsilon|\xi|^2) & 0 & \xi_2\xi_3/(\varepsilon F|\xi|^2) \\ \xi_2\xi_3/(\varepsilon|\xi|^2) & -\xi_1\xi_3/(\varepsilon|\xi|^2) & -\nu|\xi|^2 & -(\xi_1^2+\xi_2^2)/(\varepsilon F|\xi|^2) \\ 0 & 0 & 1/(\varepsilon F) & -\nu'|\xi|^2 \end{pmatrix}.$$

For $0<r<R$ we will denote by $\mathcal{C}_{r,R}$ the set

$$\mathcal{C}_{r,R} = \{\xi \in \mathbb{R}^3 : |\xi| \le R \text{ and } |\xi_3| \ge r\}.$$

We also introduce the following frequency truncation operator on $\mathcal{C}_{r,R}$:

$$\mathcal{P}_{r,R} = \chi\left(\frac{|D|}{R}\right)\left(1 - \chi\left(\frac{|D_3|}{r}\right)\right), \tag{A-80}$$

where χ is the smooth cut-off function introduced before and ($\mathcal{F}^{-1}$ is the inverse Fourier transform)

$$\chi\left(\frac{|D|}{R}\right) f = \mathcal{F}^{-1}\left(\chi\left(\frac{|\xi|}{R}\right)\hat{f}(\xi)\right) \quad \text{and} \quad \chi\left(\frac{|D_3|}{r}\right) f = \mathcal{F}^{-1}\left(\chi\left(\frac{|\xi_3|}{r}\right)\hat{f}(\xi)\right),$$

and $|D|^s$ the classical derivation operator $|D|^s f = \mathcal{F}^{-1}(|\xi|^s \hat{f}(\xi))$.

In what follows we will use it for particular radii $r_\varepsilon = \varepsilon^m$ and $R_\varepsilon = \varepsilon^{-M}$, where m and M will be made precise later. Let us end with the following anisotropic Bernstein-type result (we refer to [Charve 2005], and to [Iftimie 1999b] for more general anisotropic estimates):

Lemma 42. *There exists a constant $C>0$ such that for any function f, $\alpha>0$, $1 \le q \le p \le \infty$, and all $0<r<R$, we have*

$$\begin{aligned} \left\|\chi\left(\frac{|D|}{R}\right)\chi\left(\frac{|D_3|}{r}\right)f\right\|_{L^p} &\le C\|f\|_{L^p}, \\ \left\|\chi\left(\frac{|D|}{R}\right)\chi\left(\frac{|D_3|}{r}\right)f\right\|_{L^p} &\le C(R^2 r)^{\frac{1}{q}-\frac{1}{p}}\left\|\chi\left(\frac{|D|}{R}\right)\chi\left(\frac{|D_3|}{r}\right)f\right\|_{L^q}. \end{aligned} \tag{A-81}$$

Moreover if f has its frequencies located in $\mathcal{C}_{r,R}$, then

$$\||D|^\alpha f\|_{L^p} \le CR^\alpha \|f\|_{L^p}.$$

□

A.3.1. *Eigenvalues, projectors.* We begin with the eigenvalues and eigenvectors of matrix the $\mathbb{B}(\xi,\varepsilon)$. We refer to [Charve 2004; 2005; 2006; 2018a; Charve and Ngo 2011] for details about the following proposition. We will only state the results and skip details as the proof is an adaptation of Proposition 3.1 from [Charve and Ngo 2011] (there in the anisotropic case).

Proposition 43. *If $\nu \neq \nu'$ there exists $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$, for all $r_\varepsilon = \varepsilon^m$, and $R_\varepsilon = \varepsilon^{-M}$, with $M < \frac{1}{4}$ and $3M + m < 1$, and for all $\xi \in \mathcal{C}_{r_\varepsilon, R_\varepsilon}$, the matrix $\mathbb{B}(\xi,\varepsilon) = \widehat{L - \frac{1}{\varepsilon}\mathbb{P}\mathcal{A}}$ is diagonalizable and its eigenvalues have the following asymptotic expansions with respect to ε*:

$$\begin{aligned} \mu_0 &= -\nu|\xi|^2, & \lambda &= -\tau(\xi)|\xi|^2 + i\frac{|\xi|_F}{\varepsilon F|\xi|} + \varepsilon E(\xi,\varepsilon), \\ \mu &= -(\nu\xi_1^2 + \nu\xi_2^2 + \nu' F^2\xi_3^2)\frac{|\xi|^2}{|\xi|_F^2} + \varepsilon^2 D(\xi,\varepsilon), & \bar{\lambda} &= -\tau(\xi)|\xi|^2 - i\frac{|\xi|_F}{\varepsilon F|\xi|} + \varepsilon \bar{E}(\xi,\varepsilon), \end{aligned} \tag{A-82}$$

where $|\xi|_F^2 = \xi_1^2 + \xi_2^2 + F^2\xi_3^3$, and D, E denote remainder terms satisfying for all $\xi \in \mathcal{C}_{r_\varepsilon, R_\varepsilon}$

$$\begin{aligned} \varepsilon^2|D(\xi,\varepsilon)| &\le C_F|\nu-\nu'|^3\varepsilon^2|\xi|^6 \le C_F|\nu-\nu'|^3\varepsilon^{2-6M} \ll 1, \\ \varepsilon|E(\xi,\varepsilon)| &\le C_F|\nu-\nu'|^2\varepsilon|\xi|^4 \le C_F|\nu-\nu'|^2\varepsilon^{1-4M} \ll 1, \\ \varepsilon|\partial_{\xi_2}E(\xi,\varepsilon)| &\le C_F|\nu-\nu'|^2\varepsilon|\xi|^3 \le C_F|\nu-\nu'|^2\varepsilon^{1-3M} \ll 1, \end{aligned}$$

and

$$\tau(\xi) = \frac{\nu}{2}\left(1 + \frac{F^2\xi_3^2}{|\xi|_F^2}\right) + \frac{\nu'}{2}\left(1 - \frac{F^2\xi_3^2}{|\xi|_F^2}\right) \ge \min(\nu,\nu') > 0.$$

Moreover, if we denote by $\mathcal{P}_i(\xi,\varepsilon)$ the projectors onto the eigenspaces corresponding to μ, λ and $\bar{\lambda}$ ($i \in \{2,3,4\}$), and set

$$\mathbb{P}_i(u) = \mathcal{F}^{-1}(\mathcal{P}_i(\xi,\varepsilon)(\hat{u}(\xi))), \tag{A-83}$$

then for any divergence-free vector field f whose Fourier transform is supported in $\mathcal{C}_{r_\varepsilon,R_\varepsilon}$ and $s \in \mathbb{R}$, we have the estimates:

$$\|\mathbb{P}_2 f\|_{\dot{H}^s} \le C_F\|f\|_{\dot{H}^s} \times \begin{cases} 1 & \text{if } \Omega(f) \neq 0, \\ |\nu-\nu'|\varepsilon R_\varepsilon^2 = |\nu-\nu'|\varepsilon^{1-2M} & \text{if } \Omega(f) = 0, \end{cases} \tag{A-84}$$

and, for $i = 3, 4$,

$$\|\mathbb{P}_i f\|_{\dot{H}^s} \le C_F\frac{R_\varepsilon}{r_\varepsilon}\|f\|_{\dot{H}^s} = C_F\varepsilon^{-(m+M)}\|f\|_{\dot{H}^s}. \tag{A-85}$$

Finally, if we define $\mathbb{P}_{3+4}f \overset{\text{def}}{=} (\mathbb{P}_3+\mathbb{P}_4)f = (I_d - \mathbb{P}_2)f$ (as $\operatorname{div} f = 0$), then

$$\|\mathbb{P}_{3+4}f\|_{\dot{H}^s} \le C_F(1 + |\nu-\nu'|\varepsilon R_\varepsilon^2)\|f\|_{\dot{H}^s}. \tag{A-86}$$

Remark 44. In the case $\nu = \nu'$ everything is simpler: the eigenvalues have simple explicit expressions: $-\nu|\xi|^2$ (double, μ and μ_0 coincide), $-\nu|\xi|^2 \pm (i|\xi|_F)/(\varepsilon F|\xi|)$, the eigenvectors do not depend on ε and are mutually orthogonal (so that $\mathbb{P}_i$ are of norm 1) and this basis exactly corresponds to the QG/osc decomposition (for divergence-free vector fields): $\mathcal{P} = \mathbb{P}_{3+4}$ and $\mathcal{Q} = \mathbb{P}_2$ so that the quasigeostrophic part only depends on W_2 and the oscillating part only depends on $W_{3,4}$. Finally the operator Γ reduces to a simple anisotropic Laplace operator. We refer to [Charve 2006, Appendix B; 2016; 2018a] for more details.

Remark 45. We emphasize that the leading term in μ is the Fourier symbol of the quasigeostrophic operator Γ. Moreover, the dispersion is related to the term $(i|\xi|_F)/(\varepsilon F|\xi|)$, and when $F=1$ this term reduces to the constant $\frac{i}{\varepsilon}$. This is why dispersion does not occur in the case $F=1$ (we refer to [Chemin 1997; Charve 2018b] for a study of the asymptotics in the special case $F=1$).

A.3.2. *Dispersion, Strichartz estimates.* Combining Proposition 3 from [Charve 2018a] (covering the range $p\geq 4$) with the convolution arguments from the appendix of [Charve 2004] allows us to cover the full range $p\geq 1$ and obtain the following Strichartz estimates satisfied by the last two projections of the solution of system(A-79):

Proposition 46. *Assume that f satisfies (A-79) on $[0,T[$, where $\operatorname{div} f_0=0$ and the frequencies of f_0 and F are localized in $\mathcal{C}_{r_\varepsilon,R_\varepsilon}$. Then there exists a constant $C=C_{F,p,\nu,\nu'}>0$ such that, for $i\in\{3,4\}$ and $p\geq 1$, we have*

$$\|\mathbb{P}_i f\|_{L^p_T L^\infty}\leq CK(\varepsilon)\biggl(\|f_0\|_{L^2}+\int_0^T\|F_{\text{ext}}(\tau)\|_{L^2}\,d\tau\biggr).$$

where

$$K(\varepsilon)=\begin{cases}\varepsilon^{\frac14}\dfrac{R_\varepsilon^4}{r_\varepsilon^{\frac52+\frac2p}}\left[\dfrac{4}{\nu_0}\left(\dfrac1p-\dfrac14\right)\right]^{\frac1p-\frac14}=\varepsilon^{\frac14-(4M+(\frac52+\frac2p)m)}\left[\dfrac{4}{\nu_0}\left(\dfrac1p-\dfrac14\right)\right]^{\frac1p-\frac14} & \text{if } p\in[1,4],\\[2ex] \varepsilon^{\frac1p}\dfrac{R_\varepsilon^{\frac52+\frac6p}}{r_\varepsilon^{2+\frac4p}}=\varepsilon^{\frac1p-((\frac52+\frac6p)M+(2+\frac4p)m)} & \text{if } p\geq 4.\end{cases}$$

Unfortunately these estimates would be completely useless in our case: we need more flexibility than only L^p-L^∞-estimates, and in the case $\nu\neq\nu'$ we need to take into account the second term G^l as done in [Charve 2004]. We begin with the case $\nu=\nu'$, where we have to deal with the fact that we obtain Strichartz estimates on W_ε, which is not frequency-localized (we improve the method from [Charve 2006, Appendix B]). Then we deal with the case $\nu\neq\nu'$.

A.3.3. *Strichartz estimates in the case $\nu=\nu'$.* The main result of this section is stated as follows:

Proposition 47. *There exists a constant $C_F>0$ such that for any $d\in\mathbb{R}$, $r>4$, $q\geq 1$, and*

$$\theta\in\biggl]0,\frac{\frac12-\frac1r}{1-\frac4r}\biggl[\cap\,]0,1],\qquad p\in\biggl[1,\frac{4}{\theta\bigl(1-\frac4r\bigr)}\biggr],$$

if f solves (A-79) for initial data f_0 and external force F_{ext} both with zero divergence and potential vorticity, then (c_0 refers to the smaller constant appearing in the Littlewood–Paley decomposition, usually $c_0=\frac34$).

$$\||D|^d f\|_{\tilde{L}^p_t\dot{B}^0_{r,q}}\leq C_F\frac{C_{p,\theta,r}}{\nu^{\frac1p-\frac\theta4(1-\frac4r)}}\varepsilon^{\frac\theta4(1-\frac4r)}\times\biggl(\|f_0\|_{\dot{B}^\sigma_{2,q}}+\int_0^t\|F_{\text{ext}}(\tau)\|_{\dot{B}^\sigma_{2,q}}\,d\tau\biggr),\tag{A-87}$$

where

$$\sigma = d + \tfrac{3}{2} - \tfrac{3}{r} - \tfrac{2}{p} + \tfrac{\theta}{2}\big(1 - \tfrac{4}{r}\big), \quad C_{p,\theta,r} = \big[\tfrac{2}{c_0^2}\big(\tfrac{1}{p} - \tfrac{\theta}{4}(1-\tfrac{4}{r})\big)\big]^{\frac{1}{p}-\frac{\theta}{4}(1-\frac{4}{r})} \frac{2^{\frac{1}{2}(1-\frac{2}{r}-2\theta(1-\frac{4}{r}))}}{1-2^{-\frac{1}{2}(1-\frac{2}{r}-2\theta(1-\frac{4}{r}))}}.$$

Remark 48. It is interesting to compare our Strichartz estimates with the ones from [Iwabuchi et al. 2017; Scrobogna 2017] (see Proposition 5). In our estimates we use the range $r > 4$, whereas in Proposition 5 they consider the case $r \in]2,4[$ and they use it for r close to 3. Our index p is mostly equal to 2 but we can reach $p = 1$ (which is useful when there are derivatives), whereas in [Iwabuchi et al. 2017], $p > 1/\big(1 - \frac{2}{r}\big) > 2$.

Proof of Proposition 47. Let us first assume that $F_{\text{ext}} = 0$. As $\nu = \nu'$, the fact that f_0 is divergence-free and with zero potential vorticity implies that

$$f_0 = \mathbb{P} f_0 = \mathcal{P}\mathbb{P} f_0 = \mathbb{P}_{3+4}\mathbb{P} f_0 = \mathbb{P}_{3+4} f_0,$$

so that we only consider the last two eigenvalues (we recall the eigenvectors are orthogonal). The idea is to push further the Strichartz estimates without the frequency truncation we obtained in [Charve 2006]: we will once more use a simple nonstationary phase argument (see for example the works of Chemin, Desjardins, Gallagher and Grenier [Chemin et al. 2000; 2002; 2006]). As outlined previously, in this special case there is no need to truncate in frequency through the operator $\mathcal{P}_{r_\varepsilon,R_\varepsilon}$ but within the computations we will truncate considering the vertical Littlewood–Paley decomposition ($\dot\Delta_k^v u = \varphi(2^{-j} D_3)u$):

$$\|\dot\Delta_j f\|_{L_t^p L_x^r} = \|\dot\Delta_j f\|_{L^p L^r} = \sum_{k=-\infty}^{j+1} \|\dot\Delta_k^v \dot\Delta_j f\|_{L^p L^r}.$$

Now we will use the methods leading to the general Strichartz estimates (previously used when frequencies are truncated on some $\mathcal{C}_{r,R}$) as in our case $r = c_0 2^k$ and $R = C_0 2^j$. We recall that φ is the truncation function involved in the Littlewood–Paley decomposition, we denote by φ_1 another smooth truncation function, with support in a slightly larger annulus than φ and equal to 1 on $\operatorname{supp}\varphi$, and by $\mathcal{B}$ the set

$$\mathcal{B} \overset{\text{def}}{=} \{\psi \in \mathcal{C}_0^\infty(\mathbb{R}_+, \mathbb{R}^3) : \|\psi\|_{L^{\bar p}(\mathbb{R}_+, L^{\bar r}(\mathbb{R}^3))} \le 1\}.$$

Then following the same classical steps as in [Charve 2006] we get that (we choose for simplicity to write it only for the third eigenvalue) for any $\beta \ge 1$

$$\begin{aligned}
&\|\dot\Delta_k^v \dot\Delta_j f\|_{L^p L^r}\\
&\quad = \sup_{\psi\in\mathcal{B}} \int_0^\infty \dot\Delta_k^v \dot\Delta_j f(t,x)\psi(t,x)\,dx\,dt\\
&\quad = C \sup_{\psi\in\mathcal{B}} \int_0^\infty \int_{\mathbb{R}^3} e^{-\nu t|\xi|^2 + i\frac{t}{\varepsilon}\frac{|\xi|_F}{F|\xi|}} \widehat{\dot\Delta_j f_0}(\xi)\varphi_1(2^{-j}\xi)\varphi(2^{-k}|\xi_3|)\hat\psi(t,\xi)\,d\xi\,dt\\
&\quad \le C \sup_{\psi\in\mathcal{B}} \|\dot\Delta_j f_0\|_{L^2} \bigg[\int_0^\infty\int_0^\infty\int_{\mathbb{R}^3} K\Big(\nu(t+s), \frac{|t-s|}{\varepsilon}, x\Big)\cdot(\psi(t) * \bar\psi(s))(x)\,dx\,ds\,dt\bigg]^{\frac{1}{2}},\\
&\quad \le C \sup_{\psi\in\mathcal{B}} \|\dot\Delta_j f_0\|_{L^2} \bigg[\int_0^\infty\int_0^\infty \Big\|K\Big(\nu(t+s), \frac{|t-s|}{\varepsilon}, \cdot\Big)\Big\|_{L^{\bar\beta}} \|\psi(t) * \bar\psi(s)\|_{L^\beta}\,ds\,dt\bigg]^{\frac{1}{2}},
\end{aligned} \tag{A-88}$$

with K defined as follows (we refer to [Charve 2006] for details):

$$K(\sigma,\tau,x)=\int_{A_{j,k}} e^{ix\cdot\xi-\sigma|\xi|^2+i\tau\frac{|\xi|_F}{F|\xi|}}\varphi_1(2^{-j}|\xi|)^2\varphi(2^{-k}|\xi_3|)^2\,d\xi,$$

where

$$A_{j,k}\overset{\text{def}}{=}\{\xi\in\mathbb{R}^3: c_0 2^j\le|\xi|\le C_0 2^j \text{ and } c_0 2^k\le|\xi_3|\le C_0 2^k\}. \tag{A-89}$$

Interpolating the following estimates (we refer to [Charve 2006, Section B.2] for more details), and using as in [Charve 2018a, Section 3.2] that for all $a,b>0$ and $\theta\in[0,1]$ we have $\min(a,b)\le a^{1-\theta}b^\theta$,

$$\|K(\sigma,\tau,\cdot)\|_{L^\infty}\le C_F e^{-c_0^2\sigma 2^{2j}}2^{3j}\min\Big(2^{k-j},\frac{1}{\tau^{\frac12}2^{k-j}}\Big),$$
$$\|K(\sigma,\tau,\cdot)\|_{L^2}\le C_F e^{-\frac{c_0^2}{2}\sigma 2^{2j}}2^{\frac{3j}{2}}2^{\frac{k-j}{2}},$$

we get for any $r\in[2,\infty]$, $\frac1r=\frac{1-\alpha}{\infty}+\frac{\alpha}{2}=\frac{\alpha}{2}$, and $\theta\in[0,1]$

$$\begin{aligned}\|K(\sigma,\tau,\cdot)\|_{L^r}&\le C_F e^{-\frac{c_0^2}{2}\sigma 2^{2j}}\Big(2^{3j}\frac{2^{(k-j)(1-2\theta)}}{\tau^{\frac{\theta}{2}}}\Big)^{1-\frac2r}(2^{\frac{3j}{2}}2^{\frac{k-j}{2}})^{\frac2r}\\&\le C_F e^{-\frac{c_0^2}{2}\sigma 2^{2j}}2^{3j(1-\frac1r)}\frac{2^{(k-j)[1-\frac1r-2\theta(1-\frac2r)]}}{\tau^{\frac{\theta}{2}(1-\frac2r)}}.\end{aligned} \tag{A-90}$$

Now we can go back to (A-88), by the Cauchy–Schwarz inequality, fixing $\beta\ge1$ so that

$$\|\psi(t)*\bar\psi(s)\|_{L^\beta}\le\|\psi(t)\|_{L^{\bar r}}\|\psi(s)\|_{L^{\bar r}},$$

that is, choosing $\bar\beta=\frac{\beta}{\beta-1}=\frac r2$ (which implies that $r\ge4$), and using (A-90), we obtain that

$$\begin{aligned}\|\dot\Delta_k^v\dot\Delta_j f\|_{L^pL^r}\le C_F\sup_{\psi\in\mathcal{B}}\|\dot\Delta_j f_0\|_{L^2}\varepsilon^{\frac{\theta}{4}(1-\frac4r)}2^{\frac{3j}{2}(1-\frac2r)}2^{\frac{k-j}{2}(1-\frac2r-2\theta(1-\frac4r))}\\\times\Big[\int_0^\infty\int_0^\infty\frac{h(t)h(s)}{|t-s|^{\frac{\theta}{2}(1-\frac4r)}}\,ds\,dt\Big]^{\frac12},\end{aligned} \tag{A-91}$$

with

$$h(t)=e^{-\frac{c_0^2}{2}\nu t2^{2j}}\|\psi(t)\|_{L^{\bar r}}.$$

Next we will use the Hardy–Littlewood–Sobolev estimates, which we recall in $\mathbb{R}$ for the convenience of the reader (we refer to [Hardy and Littlewood 1930; Sobolev 1938; Lieb 1983]):

Proposition 49. *There exists a constant $C>0$ such that for any function $h_i\in L^{q_i}(\mathbb{R})$ ($q_i>1$ for $i=1,2$) and any $\alpha>0$, with $\frac1{q_1}+\frac1{q_2}+\alpha=2$, we have*

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{h_1(t)h_2(s)}{|t-s|^\alpha}\,dt\,ds\le C\|h_1\|_{L^{q_1}}\|h_2\|_{L^{q_2}}.$$

Choosing $h_1 = h_2 = h\mathbf{1}_{\mathbb{R}_+}$, $\alpha = \frac{\theta}{2}\big(1-\frac{4}{r}\big) > 0$, and $\frac{1}{q} = 1 - \frac{\theta}{4}\big(1-\frac{4}{r}\big)$, we get that

$$\int_0^\infty\int_0^\infty \frac{h(t)h(s)}{|t-s|^{\frac{\theta}{2}(1-\frac{4}{r})}}\,ds\,dt \le C\|h\|_{L^q}^2 \le C(\|e^{-\frac{c_0^2}{2}\nu 2^{2j}t}\|_{L^m}\|\psi\|_{L^{\bar p}L^{\bar r}})^2 \le C\bigg(\frac{1}{\nu^{\frac{1}{m}}}\bigg[\frac{2}{mc^2}\bigg]^{\frac{1}{m}} 2^{-\frac{2j}{m}}\|\psi\|_{L^{\bar p}L^{\bar r}}\bigg)^2 \tag{A-92}$$

for $m \in [1,\infty]$ chosen so that $\frac{1}{m} + \frac{1}{\bar p} = \frac{1}{q}$, that is,

$$\frac{1}{m} = \frac{1}{p} - \frac{\theta}{4}\big(1-\frac{4}{r}\big).$$

Remark 50. Note that this implies the condition $p \le \frac{4}{\theta}\big(1/\big(1-\frac{4}{r}\big)\big)$.

Combining with (A-91), we can write that

$$\|\dot\Delta_k^v\dot\Delta_j f\|_{L^pL^r} \le C_F\|\dot\Delta_j f_0\|_{L^2}\varepsilon^{\frac{\theta}{4}(1-\frac{4}{r})}2^{j(\frac{3}{2}-\frac{3}{r}-\frac{2}{p}+\frac{\theta}{2}(1-\frac{4}{r}))} \times\frac{2^{\frac{k-j}{2}(1-\frac{2}{r}-2\theta(1-\frac{4}{r}))}}{\nu^{\frac{1}{p}-\frac{\theta}{4}(1-\frac{4}{r})}}\Big[\frac{2}{c^2}\Big(\frac{1}{p}-\frac{\theta}{4}\Big(1-\frac{4}{r}\Big)\Big)\Big]^{\frac{1}{p}-\frac{\theta}{4}(1-\frac{4}{r})}. \tag{A-93}$$

It is possible to sum this for $k \le j+1$ if and only if

$$1-\frac{2}{r}-2\theta\big(1-\frac{4}{r}\big) > 0,$$

that is, as $r > 4$, when

$$\theta < \frac{1-\frac{2}{r}}{2\big(1-\frac{4}{r}\big)}.$$

Summing over k we obtain that for all j,

$$\|\dot\Delta_j f\|_{L^pL^r} \le C_F\frac{C_{p,\theta,r}}{\nu^{\frac{1}{p}-\frac{\theta}{4}(1-\frac{4}{r})}}\varepsilon^{\frac{\theta}{4}(1-\frac{4}{r})}2^{j(\frac{3}{2}-\frac{3}{r}-\frac{2}{p}+\frac{\theta}{2}(1-\frac{4}{r}))}\|\dot\Delta_j f_0\|_{L^2}, \tag{A-94}$$

which leads to the desired result in the homogeneous case. The inhomogeneous case (i.e., when $F_{\text{ext}} \ne 0$) easily follows thanks to the Duhamel formula. □

A.3.4. *Strichartz estimates in the case $\nu \ne \nu'$.*

Proposition 51. *There exists a constant $C_{F,\omega} > 0$ $\big($where $\omega = \frac{\max(\nu,\nu')}{\nu_0}\big)$ such that for any $d \in \mathbb{R}$, $r > 4$, and $p < 4/\big(1-\frac{4}{r}\big)$, if f solves (A-79) for initial data f_0 and external force $F_{\text{ext}} = F^b + F^l$, all three of them with zero divergence and potential vorticity, then for $i = 3, 4$*

$$\||D|^d\mathbb{P}_i\mathcal{P}_{r_\varepsilon,R_\varepsilon}f\|_{\widetilde L_t^p\dot B_{r,q}^0} \le C_{F,\omega}\frac{D_{p,r}}{\nu_0^{\frac{1}{p}-\frac{1}{4}(1-\frac{4}{r})}}\varepsilon^{\frac{1}{4}(1-\frac{4}{r})}\frac{R_\varepsilon^{4-\frac{9}{r}}}{r_\varepsilon^{\frac{5}{2}+\frac{2}{p}-\frac{6}{r}}} \times\bigg(\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}f_0\|_{\dot B_{2,q}^d}+\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}F^b\|_{L^1\dot B_{2,q}^d}+\frac{1}{\nu_0^{\frac{1}{2}}r_\varepsilon}\|\mathcal{P}_{r_\varepsilon,R_\varepsilon}F^l\|_{L^2\dot B_{2,q}^d}\bigg), \tag{A-95}$$

where $D_{p,r} = \max(b_{p,r}, d_{p,r})$ *with*

$$b_{p,r} = \left(\tfrac{2}{\nu c^2}\right)^{\frac{1}{p}-\frac{1}{4}(1-\frac{4}{r})}\left(\tfrac{1}{p}-\tfrac{1}{4}\left(1-\tfrac{4}{r}\right)\right)^{\frac{1}{p}-\frac{1}{4}(1-\frac{4}{r})},$$

$$d_{p,r} = 2^{\frac{1}{p}}\left(\tfrac{8}{c^2 p}\right)^{\frac{1}{p}-\frac{1}{4}(1-\frac{4}{r})}\left(\int_0^\infty \frac{e^{-x}}{x^{\frac{p}{4}(1-\frac{4}{r})}}\,dx\right)^{\frac{1}{p}}.$$

Remark 52. We could prove like in the previous section some refined estimate with $\theta \in]0,1]$ (allowing $p \le 4/\left(\theta\left(1-\frac{4}{r}\right)\right)$) but we will only need the case $\theta = 1$ and p close to 2 in this article.

Proof of Proposition 51. Let us first assume that $F_{\text{ext}} = 0$. With the same notation as in the previous section, we get that (see previous section, as well as [Charve and Ngo 2011; Charve 2018a] for details)

$$\begin{aligned}
&\|\mathbb{P}_i \mathcal{P}_{r_\varepsilon,R_\varepsilon} f\|_{\widetilde{L}^p(\mathbb{R}_+,L^r(\mathbb{R}^3))}\\
&\quad= \sup_{\psi\in\mathcal{B}} \int_0^\infty\int_{\mathbb{R}^3} \mathbb{P}_i \mathcal{P}_{r_\varepsilon,R_\varepsilon} f(t,x)\psi(t,x)\,dx\,dt\\
&\quad= \sup_{\psi\in\mathcal{B}} \int_0^\infty\int_{\mathbb{R}^3} e^{-t\tau(\xi)|\xi|^2+it\frac{|\xi|_F}{\varepsilon F|\xi|}+\varepsilon t E(\xi,\varepsilon)}\widehat{\mathbb{P}_i \mathcal{P}_{r_\varepsilon,R_\varepsilon} f_0}(t,\xi)\chi\left(\frac{|\xi|}{2R_\varepsilon}\right)\left(1-\chi\left(\frac{2|\xi_3|}{r_\varepsilon}\right)\right)\hat{\psi}(t,\xi)\,d\xi\,dt\\
&\quad\le C\sup_{\psi\in\mathcal{B}} \|\mathbb{P}_i \mathcal{P}_{r_\varepsilon,R_\varepsilon} f_0\|_{L^2}\left[\int_0^\infty\int_0^\infty \|L(t,s,\varepsilon,\cdot)\|_{L^{r/2}}\|\psi(t)*\bar{\psi}(s)\|_{L^{r/(r-2)}}\,ds\,dt\right]^{\frac{1}{2}},
\end{aligned}\tag{A-96}$$

where

$$L(t,s,\varepsilon,x) = \int_{\mathbb{R}^3} e^{ix\cdot\xi-(t+s)\tau(\xi)|\xi|^2+i(t-s)\frac{|\xi|_F}{\varepsilon F|\xi|}+\varepsilon t E(\xi,\varepsilon)+\varepsilon s\overline{E}(\xi,\varepsilon)}\chi\left(\frac{|\xi|}{2R_\varepsilon}\right)^2\left(1-\chi\left(\frac{2|\xi_3|}{r_\varepsilon}\right)\right)^2 d\xi.$$

Like before, to obtain the $L^{\frac{r}{2}}$-norm, we will interpolate between L^2 and L^∞. It is easy to obtain

$$\|L(s,t,\varepsilon,\cdot)\|_{L^2} \le C_F R_\varepsilon^{\frac{3}{2}} e^{-c^2\frac{\nu_0}{4}(t+s)r_\varepsilon^2},$$

and we refer to [Charve and Ngo 2011; Charve 2018a] where we proved that (there we were working with local-in-time solutions, and we dropped the exponential)

$$\|L(s,t,\varepsilon,\cdot)\|_{L^\infty} \le C_{F,\omega}\frac{R_\varepsilon^3}{r_\varepsilon^2}\min\left(1, \frac{R_\varepsilon^3}{r_\varepsilon^2}\left(\frac{\varepsilon}{|t-s|}\right)^{\frac{1}{2}}\right)e^{-c^2\frac{\nu_0}{4}(t+s)r_\varepsilon^2},$$

so that we obtain for any $\beta \ge 2$

$$\|L(s,t,\varepsilon,\cdot)\|_{L^\beta} \le C_{F,\omega} e^{-c^2\frac{\nu_0}{4}(t+s)r_\varepsilon^2}\frac{R_\varepsilon^{6-\frac{9}{\beta}}}{r_\varepsilon^{4-\frac{8}{\beta}}}\left(\frac{\varepsilon}{|t-s|}\right)^{\frac{1}{2}(1-\frac{2}{\beta})}.$$

Thanks to (A-85), and doing the same as previously, we end up with $\left(\beta = \frac{r}{2}\right)$

$$\|\mathbb{P}_i \mathcal{P}_{r_\varepsilon,R_\varepsilon} f\|_{L^pL^r} = C_{F,\omega}\|\mathcal{P}_{r_\varepsilon,R_\varepsilon} f_0\|_{L^2}\frac{R_\varepsilon^{4-\frac{9}{r}}}{r_\varepsilon^{3-\frac{8}{r}}}\varepsilon^{\frac{1}{4}(1-\frac{4}{r})}\sup_{\psi\in\mathcal{B}}\left[\int_0^\infty\int_0^\infty \frac{g(t)g(s)}{|t-s|^{\frac{1}{2}(1-\frac{4}{r})}}\,ds\,dt\right]^{\frac{1}{2}},\tag{A-97}$$

with

$$g(t) = e^{-\frac{c^2}{2}\nu t r_\varepsilon^2} \|\psi(t)\|_{L^{\bar{r}}}.$$

Using once more Proposition 49, we end up with

$$\|\mathbb{P}_i \mathcal{P}_{r_\varepsilon,R_\varepsilon} f\|_{L^p L^r} \le C_F \frac{b_{p,r}}{\nu_0^{\frac{1}{p}-\frac{1}{4}(1-\frac{4}{r})}} \varepsilon^{\frac{1}{4}(1-\frac{4}{r})} \frac{R_\varepsilon^{4-\frac{9}{r}}}{r_\varepsilon^{\frac{5}{2}+\frac{2}{p}-\frac{6}{r}}} \|\mathcal{P}_{r_\varepsilon,R_\varepsilon} f_0\|_{L^2}. \tag{A-98}$$

Then it is easy to deduce the nonhomogeneous case with F^b only. Let us now focus on the other external force term; we extend the method from [Charve 2004]. If we denote by $S(t) f_0$ the solution of system (A-79) with $F_{\text{ext}} = 0$, we have by the Duhamel formula

$$\begin{aligned}
&\left\| \int_0^t S(t-t')\mathcal{P}_{r_\varepsilon,R_\varepsilon}\mathbb{P}_i F^l(t')\,dt' \right\|_{L_t^p L^r} \\
&= \sup_{\psi\in\mathcal{B}} \int_0^\infty \int_{\mathbb{R}^3} \widehat{\mathbb{P}_i \mathcal{P}_{r_\varepsilon,R_\varepsilon} F^l}(t',\xi) \\
&\qquad \times \int_{t'}^\infty e^{-(t-t')\tau(\xi)|\xi|^2 + i(t-t')\frac{|\xi|_F}{\varepsilon F|\xi|} + \varepsilon t E(\xi,\varepsilon)} \chi\left(\frac{|\xi|}{2R_\varepsilon}\right)\left(1-\chi\left(\frac{2|\xi_3|}{r_\varepsilon}\right)\right)\hat{\psi}(t,\xi)\,dt\,d\xi\,dt' \\
&\le C \sup_{\psi\in\mathcal{B}} \|\mathbb{P}_i \mathcal{P}_{r_\varepsilon,R_\varepsilon} F^l\|_{L^2L^2} \left[\int_0^\infty \int_{t'}^\infty \int_{t'}^\infty \|L(t-t',s-t',\varepsilon,\cdot)\|_{L^{r/2}} \|\psi(t) * \bar{\psi}(s)\|_{L^{r/(r-2)}}\,ds\,dt \right]^{\frac{1}{2}} \\
&\le C_{F,\omega} \|\mathcal{P}_{r_\varepsilon,R_\varepsilon} F^l\|_{L^2L^2} \frac{R_\varepsilon^{4-\frac{9}{r}}}{r_\varepsilon^{3-\frac{8}{r}}} \varepsilon^{\frac{1}{4}(1-\frac{4}{r})} \\
&\qquad \times \sup_{\psi\in\mathcal{B}} \left[\int_0^\infty \int_0^\infty \int_0^\infty \mathbf{1}_{\{t'\le \min(t,s)\}} \frac{e^{-c^2\frac{\nu_0}{4}(t+s-2t')r_\varepsilon^2}}{|t-s|^{\frac{1}{2}(1-\frac{4}{r})}} \|\psi(t)\|_{L^{\bar{r}}} \|\psi(s)\|_{L^{\bar{r}}}\,ds\,dt\,dt' \right]^{\frac{1}{2}}.
\end{aligned} \tag{A-99}$$

Computing the integral in t',

$$\int_0^{\min(s,t)} e^{c^2\frac{\nu_0}{2}t' r_\varepsilon^2}\,dt' \le \frac{2}{\nu_0 r_\varepsilon^2} e^{c^2\frac{\nu_0}{2}\min(t,s) r_\varepsilon^2},$$

and using the fact that $|t-s| = s+t-2\min(s,t)$, we get

$$\begin{aligned}
&\left\| \int_0^t S(t-t')\mathcal{P}_{r_\varepsilon,R_\varepsilon}\mathbb{P}_i F^l(t')\,dt' \right\|_{L_t^p L^r} \\
&\le C_{F,\omega} \|\mathcal{P}_{r_\varepsilon,R_\varepsilon} F^l\|_{L^2L^2} \frac{R_\varepsilon^{4-\frac{9}{r}}}{r_\varepsilon^{4-\frac{8}{r}}} \varepsilon^{\frac{1}{4}(1-\frac{4}{r})} \\
&\qquad\qquad \times \sup_{\psi\in\mathcal{B}} \left[\int_0^\infty \int_0^\infty \frac{e^{-c^2\frac{\nu_0}{4}|t-s|r_\varepsilon^2}}{|t-s|^{\frac{1}{2}(1-\frac{4}{r})}} \|\psi(t)\|_{L^{\bar{r}}} \|\psi(s)\|_{L^{\bar{r}}}\,ds\,dt \right]^{\frac{1}{2}}.
\end{aligned} \tag{A-100}$$

Then setting

$$k(\tau) = e^{-c^2\frac{\nu_0}{4}|\tau| r_\varepsilon^2} |\tau|^{-\frac{1}{2}(1-\frac{4}{r})},$$

we just have to estimate a convolution

$$\int_0^\infty \int_0^\infty k(t-s)\|\psi(t)\|_{L^{\bar{r}}}\|\psi(s)\|_{L^{\bar{r}}}\,ds\,dt \le \|k\|_{L^{p/2}}\|\psi\|^2_{L^{\bar{p}}L^{\bar{r}}}, \tag{A-101}$$

provided that $p \ge 2$ and $\frac{p}{4}\left(1-\frac{4}{r}\right) < 1$ so that $k \in L^{\frac{p}{2}}$, whose norm is featured in the constant $d_{p,r}$ and we have

$$\|\mathbb{P}_i \mathcal{P}_{r_\varepsilon,R_\varepsilon} f\|_{L^pL^r} \le C_{F,\omega} \frac{D_{p,r}}{\nu_0^{\frac{1}{p}-\frac{1}{4}(1-\frac{4}{r})}} \varepsilon^{\frac{1}{4}(1-\frac{4}{r})} \frac{R_\varepsilon^{4-\frac{9}{r}}}{r_\varepsilon^{\frac{5}{2}+\frac{2}{p}-\frac{6}{r}}} \times \left(\|\mathcal{P}_{r_\varepsilon,R_\varepsilon} f_0\|_{L^2} + \|\mathcal{P}_{r_\varepsilon,R_\varepsilon} F^b\|_{L^1L^2} + \frac{1}{\nu_0^{\frac{1}{2}} r_\varepsilon}\|\mathcal{P}_{r_\varepsilon,R_\varepsilon} F^l\|_{L^2L^2}\right). \tag{A-102}$$

Finally, to obtain the announced estimates, we just have to apply this estimate to $\dot{\Delta}_j |D|^d f$. □

Acknowledgements

This work was supported by the ANR project INFAMIE, ANR-15-CE40-0011.

References

[Abidi and Hmidi 2008] H. Abidi and T. Hmidi, "On the global well-posedness of the critical quasi-geostrophic equation", *SIAM J. Math. Anal.* **40**:1 (2008), 167–185. MR Zbl

[Babin et al. 2001] A. Babin, A. Mahalov, and B. Nicolaenko, "Strongly stratified limit of 3D primitive equations in an infinite layer", pp. 1–11 in *Advances in wave interaction and turbulence* (South Hadley, MA, 2000), edited by P. A. Milewski et al., Contemp. Math. **283**, Amer. Math. Soc., Providence, RI, 2001. MR Zbl

[Bahouri et al. 2011] H. Bahouri, J.-Y. Chemin, and R. Danchin, *Fourier analysis and nonlinear partial differential equations*, Grundlehren der Mathematischen Wissenschaften **343**, Springer, 2011. MR Zbl

[Bougeault and Sadourny 1998] P. Bougeault and R. Sadourny, *Dynamique de l'atmosphère et de l'océan*, École Polytechnique, Palaiseau, 1998.

[Bourgeois and Beale 1994] A. J. Bourgeois and J. T. Beale, "Validity of the quasigeostrophic model for large-scale flow in the atmosphere and ocean", *SIAM J. Math. Anal.* **25**:4 (1994), 1023–1068. MR Zbl

[Calderón 1990] C. P. Calderón, "Existence of weak solutions for the Navier–Stokes equations with initial data in L^p", *Trans. Amer. Math. Soc.* **318**:1 (1990), 179–200. MR Zbl

[Charve 2004] F. Charve, "Global well-posedness and asymptotics for a geophysical fluid system", *Comm. Partial Differential Equations* **29**:11-12 (2004), 1919–1940. MR Zbl

[Charve 2005] F. Charve, "Convergence of weak solutions for the primitive system of the quasigeostrophic equations", *Asymptot. Anal.* **42**:3-4 (2005), 173–209. MR Zbl

[Charve 2006] F. Charve, "Asymptotics and vortex patches for the quasigeostrophic approximation", *J. Math. Pures Appl.* (9) **85**:4 (2006), 493–539. MR Zbl

[Charve 2008] F. Charve, "Global well-posedness for the primitive equations with less regular initial data", *Ann. Fac. Sci. Toulouse Math.* (6) **17**:2 (2008), 221–238. MR Zbl

[Charve 2016] F. Charve, "A priori estimates for the 3D quasi-geostrophic system", *J. Math. Anal. Appl.* **444**:2 (2016), 911–946. MR Zbl

[Charve 2018a] F. Charve, "Asymptotics and lower bound for the lifespan of solutions to the primitive equations", *Acta Appl. Math.* **158** (2018), 11–47. MR Zbl

[Charve 2018b] F. Charve, "Global well-posedness and asymptotics for a penalized Boussinesq-type system without dispersion", *Commun. Math. Sci.* **16**:3 (2018), 791–807. MR Zbl

[Charve and Ngo 2011] F. Charve and V.-S. Ngo, "Global existence for the primitive equations with small anisotropic viscosity", *Rev. Mat. Iberoam.* **27**:1 (2011), 1–38. MR Zbl

[Chemin 1997] J.-Y. Chemin, "À propos d'un problème de pénalisation de type antisymétrique", *J. Math. Pures Appl.* (9) **76**:9 (1997), 739–755. MR Zbl

[Chemin et al. 2000] J.-Y. Chemin, B. Desjardins, I. Gallagher, and E. Grenier, "Fluids with anisotropic viscosity", *M2AN Math. Model. Numer. Anal.* **34**:2 (2000), 315–335. MR Zbl

[Chemin et al. 2002] J.-Y. Chemin, B. Desjardins, I. Gallagher, and E. Grenier, "Anisotropy and dispersion in rotating fluids", pp. 171–192 in *Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. XIV* (Paris, 1997/1998), edited by D. Cioranescu, Stud. Math. Appl. **31**, North-Holland, Amsterdam, 2002. MR Zbl

[Chemin et al. 2006] J.-Y. Chemin, B. Desjardins, I. Gallagher, and E. Grenier, *Mathematical geophysics: an introduction to rotating fluids and the Navier–Stokes equations*, Oxford Lecture Series in Mathematics and its Applications **32**, Oxford University Press, 2006. MR Zbl

[Córdoba and Córdoba 2004] A. Córdoba and D. Córdoba, "A maximum principle applied to quasi-geostrophic equations", *Comm. Math. Phys.* **249**:3 (2004), 511–528. MR Zbl

[Cushman-Roisin 1994] B. Cushman-Roisin, *Introduction to geophysical fluid dynamics*, Prentice Hall, Englewood Cliffs, NJ, 1994.

[Dutrifoy 2004] A. Dutrifoy, "Slow convergence to vortex patches in quasigeostrophic balance", *Arch. Ration. Mech. Anal.* **171**:3 (2004), 417–449. MR Zbl

[Dutrifoy 2005] A. Dutrifoy, "Examples of dispersive effects in non-viscous rotating fluids", *J. Math. Pures Appl.* (9) **84**:3 (2005), 331–356. MR Zbl

[Gallagher 1998] I. Gallagher, "Applications of Schochet's methods to parabolic equations", *J. Math. Pures Appl.* (9) **77**:10 (1998), 989–1054. MR Zbl

[Gallagher and Planchon 2002] I. Gallagher and F. Planchon, "On global infinite energy solutions to the Navier–Stokes equations in two dimensions", *Arch. Ration. Mech. Anal.* **161**:4 (2002), 307–337. MR Zbl

[Giga et al. 2008] Y. Giga, K. Inui, A. Mahalov, and J. Saal, "Uniform global solvability of the rotating Navier–Stokes equations for nondecaying initial data", *Indiana Univ. Math. J.* **57**:6 (2008), 2775–2791. MR Zbl

[Hardy and Littlewood 1930] G. H. Hardy and J. E. Littlewood, "Notes on the theory of series, XII: On certain inequalities connected with the calculus of variations", *J. London Math. Soc.* **5**:1 (1930), 34–39. MR Zbl

[Hieber and Shibata 2010] M. Hieber and Y. Shibata, "The Fujita–Kato approach to the Navier–Stokes equations in the rotational framework", *Math. Z.* **265**:2 (2010), 481–491. MR Zbl

[Hmidi 2005] T. Hmidi, "Régularité höldérienne des poches de tourbillon visqueuses", *J. Math. Pures Appl.* (9) **84**:11 (2005), 1455–1495. MR Zbl

[Hmidi and Keraani 2007] T. Hmidi and S. Keraani, "Global solutions of the super-critical 2D quasi-geostrophic equation in Besov spaces", *Adv. Math.* **214**:2 (2007), 618–638. MR Zbl

[Iftimie 1999a] D. Iftimie, "Approximation of the quasigeostrophic system with the primitive systems", *Asymptot. Anal.* **21**:2 (1999), 89–97. MR Zbl

[Iftimie 1999b] D. Iftimie, "The resolution of the Navier–Stokes equations in anisotropic spaces", *Rev. Mat. Iberoamericana* **15**:1 (1999), 1–36. MR Zbl

[Iwabuchi et al. 2017] T. Iwabuchi, A. Mahalov, and R. Takada, "Global solutions for the incompressible rotating stably stratified fluids", *Math. Nachr.* **290**:4 (2017), 613–631. MR Zbl

[Koba et al. 2012] H. Koba, A. Mahalov, and T. Yoneda, "Global well-posedness for the rotating Navier–Stokes–Boussinesq equations with stratification effects", *Adv. Math. Sci. Appl.* **22**:1 (2012), 61–90. MR Zbl

[Koh et al. 2014a] Y. Koh, S. Lee, and R. Takada, "Dispersive estimates for the Navier–Stokes equations in the rotational framework", *Adv. Differential Equations* **19**:9-10 (2014), 857–878. MR Zbl

[Koh et al. 2014b] Y. Koh, S. Lee, and R. Takada, "Strichartz estimates for the Euler equations in the rotational framework", *J. Differential Equations* **256**:2 (2014), 707–744. MR Zbl

[Lee and Takada 2017] S. Lee and R. Takada, "Dispersive estimates for the stably stratified Boussinesq equations", *Indiana Univ. Math. J.* **66**:6 (2017), 2037–2070. MR Zbl

[Lieb 1983] E. H. Lieb, "Sharp constants in the Hardy–Littlewood–Sobolev and related inequalities", *Ann. of Math.* (2) **118**:2 (1983), 349–374. MR Zbl

[Ngo 2009] V.-S. Ngo, "Rotating fluids with small viscosity", *Int. Math. Res. Not.* **2009**:10 (2009), 1860–1890. MR Zbl

[Ngo and Scrobogna 2018] V.-S. Ngo and S. Scrobogna, "Dispersive effects of weakly compressible and fast rotating inviscid fluids", *Discrete Contin. Dyn. Syst.* **38**:2 (2018), 749–789. MR Zbl

[Pedlosky 1979] J. Pedlosky, *Geophysical fluid dynamics*, Springer, 1979.

[Scrobogna 2017] S. Scrobogna, "Derivation of limit equations for a singular perturbation of a 3D periodic Boussinesq system", *Discrete Contin. Dyn. Syst.* **37**:12 (2017), 5979–6034. MR Zbl

[Scrobogna 2018] S. Scrobogna, "Highly rotating fluids with vertical stratification for periodic data and vanishing vertical viscosity", *Rev. Mat. Iberoam.* **34**:1 (2018), 1–58. MR Zbl

[Sobolev 1938] S. L. Sobolev, "On a theorem of functional analysis", *Mat. Sbornik* **5** (1938), 471–479. In Russian; translated in *Amer. Math. Soc. Transl.* **1963**:2 (1963), 39–68.

[Stein 1970] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series **30**, Princeton University Press, 1970. MR Zbl

[Takada 2016] R. Takada, "Long time existence of classical solutions for the 3D incompressible rotating Euler equations", *J. Math. Soc. Japan* **68**:2 (2016), 579–608. MR Zbl

[Widmayer 2018] K. Widmayer, "Convergence to stratified flow for an inviscid 3D Boussinesq system", *Commun. Math. Sci.* **16**:6 (2018), 1713–1728. MR Zbl

Received 28 Oct 2019. Accepted 7 Jan 2020.

FRÉDÉRIC CHARVE: frederic.charve@u-pec.fr
Laboratoire d'Analyse et de Mathématiques Appliquées (UMR 8050), Université Paris-Est Créteil, Créteil, France

PURE and APPLIED
ANALYSIS

Vol. 2, No. 2, 2020
https://doi.org/10.2140/paa.2020.2.519

DISPERSIVE ESTIMATES, BLOW-UP AND FAILURE OF STRICHARTZ ESTIMATES FOR THE SCHRÖDINGER EQUATION WITH SLOWLY DECAYING INITIAL DATA

RAINER MANDEL

The initial value problem for the homogeneous Schrödinger equation is investigated for radially symmetric initial data with slow decay rates and not too wild oscillations. Our global well-posedness results apply to initial data for which Strichartz estimates fail.

1. Introduction

In this paper we investigate the initial value problem for the Schrödinger equation

$$i\partial_t \psi + \Delta\psi = 0 \quad \text{in } \mathbb{R}^n, \qquad \psi(0) = \phi, \tag{1}$$

for radial initial data ϕ with slow decay at infinity. In particular, we are interested in a solution theory for (1) without assuming ϕ to belong to one of the Lebesgue spaces $L^r(\mathbb{R}^n)$ with $r \in [1,2]$. In this case Strichartz estimates are not available and local or global well-posedness results for (1) are unknown. Surprisingly, we could not find a complete statement about Strichartz estimates for such initial data in the literature, so we clarify this point here.

Theorem 1. *Let $n \in \mathbb{N}$ and $p, q \in [1,\infty]$, $r > 2$. Then there is no Strichartz estimate*

$$\|e^{it\Delta}\phi\|_{L^p_t(\mathbb{R};L^q(\mathbb{R}^n))} \lesssim \|\phi\|_{L^r(\mathbb{R}^n)}. \tag{2}$$

This theorem partly generalizes the known fact that for any given $t > 0$ the Schrödinger propagator $e^{it\Delta}$ is unbounded as a map from $L^r(\mathbb{R}^n)$ to $L^q(\mathbb{R}^n)$ for all $r > 2$, $q \in [1,\infty]$; see [Linares and Ponce 2015, p. 63] for the case $q = r$. Theorem 1 may seem surprising in view of the fact that the optimal conditions for Strichartz estimates in the most important special case $r = 2$ do not provide any obvious reason why the estimates should break down completely for $r > 2$. Recall that these conditions are given by

$$p, q \geq 2, \quad (p,q,n) \neq (2,\infty,2), \quad \frac{2}{p} + \frac{n}{q} = \frac{n}{2};$$

see for instance [Cazenave 2003, Theorem 2.3.3]. We refer to [Strichartz 1977; Keel and Tao 1998; Ginibre and Velo 1985] for three milestone contributions related to the discovery of these conditions. At least for $n \geq 3$, each of the above conditions has a counterpart in the range $r > 2$. The scaling invariance of the Schrödinger equation implies $2/p + n/q = n/r$ so that $q \geq r$ would be an immediate consequence that

MSC2010: primary 35Q41; secondary 35B40, 35B44.
Keywords: Schrödinger equation, failure of Strichartz estimates, dispersive blow-up.

replaces the condition $q \geq 2$. As we discuss in the Appendix $p \geq 2$ generalizes to $p \geq 2r/(2n - r(n-1))_+$. In particular, there is no evident reason for the necessity of $r \leq 2$ so that Theorem 1 seems to fill a gap in the literature. Its short proof relies on a thorough analysis of a counterexample due to Bona, Ponce, Saut, and Sparber [Bona et al. 2014]. The main feature of their solution is that the corresponding initial datum oscillates quadratically with respect to the distance to the origin, which produces L^∞-blow-up (or dispersive blow-up) of the solution at some prescribable finite time; see [Bona et al. 2014, Lemma 2.1]. We reconsider this self-similar blow-up analysis for partly more general initial data and estimate the blow-up rate in $L^q(\mathbb{R}^n)$, which eventually leads to Theorem 1. Accordingly, our proof even reveals that local Strichartz estimates cannot hold either and that no improvement in the radial situation is possible.

Given that Strichartz estimates fail, the question arises of how well-posedness results for the Schrödinger equation can be achieved if the initial datum lies in $L^r(\mathbb{R}^n)$ only for $r > 2$. In view of the above-mentioned counterexample it seems reasonable to impose a condition on the oscillations of the initial datum. In the following we present one possible approach in the radially symmetric case which relies on suitably weighted Sobolev norms of the initial data. For instance we identify a class of initial data lying in $L^r(\mathbb{R}^n)$ only for $r > 2n/(n-1)$ with solutions that are bounded in time and uniformly localized in space; see Corollary 3. In that case dispersion need not occur because there are solutions of the form

$$\psi(x,t) = e^{-i\omega^2 t}\phi(x), \quad \text{where } \phi(x) = |x|^{(2-n)/2} J_{(n-2)/2}(\omega|x|) \text{ for some } \omega \in \mathbb{R}. \tag{3}$$

Here, $J_{(n-2)/2}$ denotes the Bessel function of the first kind and ϕ solves the linear Helmholtz equation $\Delta\phi + \omega^2\phi = 0$ in $\mathbb{R}^n$. Aiming for a more general result in this direction, we first consider initial profiles of the form $\phi(x) = e^{i\omega|x|}\phi_\omega(|x|)$, where ϕ_ω belongs to the function spaces X and Y_m for some $m \in \{0, \dots, n\}$, which we introduce now. The space X is defined to be the completion of $C_c^\infty(\mathbb{R}_{\geq 0}; \mathbb{C})$ with respect to the norm $\|\cdot\|_X := \|\cdot\|_{X_1} + \|\cdot\|_{X_2}$ given by

$$\|f\|_{X_1} := \sup_{z>0} z^{(1-n)/2} \int_0^z (|f(r)| r^{n-2} + |f'(r)| r^{n-1})\, dr,$$
$$\|f\|_{X_2} := \int_0^\infty \left| \frac{d}{dr}(f(r) r^{(n-1)/2}) \right| dr + \sup_{z>0} \left(z \int_z^\infty |f(r)| r^{(n-5)/2}\, dr \right).$$

Similarly, we define Y_m to be the completion of $C_c^\infty(\mathbb{R}_{\geq 0}; \mathbb{C})$ with respect to the norm

$$\|f\|_{Y_m} := \sum_{k=0}^{m} \int_0^\infty |f^{(k)}(r)| r^{n-m+k-1}\, dr \quad \text{if } m \in \{0, \dots, n-1\},$$
$$\|f\|_{Y_n} := \sum_{k=1}^{n} \int_0^\infty |f^{(k)}(r)| r^{k-1}\, dr + \sup_{z>0} z^{-2} \int_0^z f(r) r\, dr + |f(0)|.$$

One main feature of these spaces is that slow decay rates of their elements are admissible only provided that their derivatives decay fast enough. For instance, we have that $r \mapsto (1+r)^{-\alpha}$ lies in X if and only if $\alpha \geq (n-1)/2$ and in Y_m if and only if $\alpha > n - m$, whereas $r \mapsto e^{ir}(1+r)^{-\alpha}$ belongs to X if and only if $\alpha > (n+1)/2$ and to Y_m if and only if $\alpha > n$. Using these spaces we find a local well-posedness theory for at most linearly oscillating radial initial data.

Theorem 2. *Let $n \in \mathbb{N}$, $m \in \{0, \dots, n\}$, and $\phi(x) = \phi_\omega(|x|)e^{i\omega|x|}$ for some $\omega \in \mathbb{R}$:*

(i) *If $\phi_\omega \in Y_m$ then (1) has a unique global solution ψ satisfying*

$$|\psi(x,t)| \le C(\sqrt{t})^{m-n} \|\phi_\omega\|_{Y_m}.$$

(ii) *If $\phi_\omega \in X$ then (1) has a unique global solution ψ satisfying*

$$|\psi(x,t)| \le C|x|^{(1-n)/2} \|\phi_\omega\|_X.$$

In (i) *and* (ii) *the constant C does not depend on ω.*

Here we used the term "global solution" to indicate a distributional solution of the Schrödinger equation (1) away from $t = 0$, i.e., on $\mathbb{R} \setminus \{0\} \times \mathbb{R}^n$. Given that the test functions are dense in both X and Y_m and that the above estimates hold, one even gets that such solutions are limits of smooth classical solutions with respect to uniform convergence on all compact sets avoiding $t = 0$ and $x = 0$. By the estimate (i) for $m = n$ this convergence is even uniform on $\mathbb{R} \times \mathbb{R}^n$ if we assume $\phi_\omega \in Y_n$. Combining the estimates (i) and (ii) we deduce the following.

Corollary 3. *Let $n \in \mathbb{N}$ and assume $\phi(x) = \int_{\mathbb{R}} \phi_\omega(|x|)e^{i\omega|x|}\, d\mu(\omega)$ for some Borel measure μ on $\mathbb{R}$. Then (1) has a unique global solution satisfying*

$$|\psi(x,t)| \le C(1+|x|)^{(1-n)/2} \int_{\mathbb{R}} (\|\phi_\omega\|_X + \|\phi_\omega\|_{Y_n})\, d\mu(\omega),$$

provided the right-hand side is finite.

Corollary 4. *Let $n \in \mathbb{N}$, $m \in \{0, \dots, n\}$, and assume $\phi(x) = \int_{\mathbb{R}} \phi_\omega(|x|)e^{i\omega|x|}\, d\mu(\omega)$ for some Borel measure μ on $\mathbb{R}$. Then (1) has a unique global solution satisfying*

$$|\psi(x,t)| \le C(1+t)^{-m/2} \int_{\mathbb{R}} (\|\phi_\omega\|_{Y_{n-m}} + \|\phi_\omega\|_{Y_n})\, d\mu(\omega),$$

provided the right-hand side is finite.

Remark 1. (a) In the case $n \ge 2$ we can apply Corollary 3 to initial conditions that are sufficiently regular superpositions of radially symmetric Herglotz waves. The corresponding densities a are assumed to be Lebesgue measurable and to satisfy $\int_{\mathbb{R}} |a(\omega)|(|\omega|^{-1/2} + |\omega|^{(n-2)/2}) < \infty$. Using the asymptotic expansions of the Bessel functions at infinity one finds

$$\phi(x) := \int_0^\infty a(\omega)|x|^{(2-n)/2} J_{(n-2)/2}(\omega|x|)\, d\omega = \int_{\mathbb{R}} \phi_\omega(|x|)e^{i\omega|x|}\, d\mu(\omega),$$

where $d\mu(\omega) = d\delta_0(\omega) + d\omega$ (δ_0 is the Dirac delta distribution) and

$$\phi_\omega(r) = r^{1-n} \cdot \begin{cases} a(\omega)\omega^{-n/2} B_n(\omega r) & \text{if } \omega > 0, \\ \int_0^\infty a(\xi)\xi^{-n/2} A_n(\xi r)\, d\xi & \text{if } \omega = 0, \\ a(-\omega)|\omega|^{-n/2} \overline{B_n(-\omega r)} & \text{if } \omega < 0. \end{cases}$$

Here we anticipated the notation from (6). So Proposition 6 and a few computations yield that $\omega \mapsto \|\phi_\omega\|_X + \|\phi_\omega\|_{Y_n}$ is μ-integrable because of

$$\|\phi_0\|_X + \int_{\mathbb{R}} \|\phi_\omega\|_X \, d\omega \lesssim \int_{\mathbb{R}} |a(\omega)| |\omega|^{-1/2} \, d\omega,$$
$$\|\phi_0\|_{Y_n} + \int_{\mathbb{R}} \|\phi_\omega\|_{Y_n} \, d\omega \lesssim \int_{\mathbb{R}} |a(\omega)| |\omega|^{(n-2)/2} \, d\omega.$$

In particular, Corollary 3 provides an abstract framework for the observation that the solutions of the Schrödinger equation with initial data given by sufficiently regular superpositions of radially symmetric Herglotz waves as in (3) remain bounded in time and uniformly localized in space.

(b) Theorem 2(i) is a generalized version of the fact that integrable initial data yield bounded solutions. Indeed, the latter statement corresponds to $m = 0$ in the theorem. So we get that less integrability of the initial datum is still sufficient for the absence of finite time L^∞-blow-up provided that the derivatives decay sufficiently fast. Notice that some kind of control on the derivatives seems necessary given that there are initial data in $L^s(\mathbb{R}^n)$ for any given $s > 1$, the corresponding solutions of which become unbounded in finite time; see [Bona et al. 2014, Lemma 2.1 and Remark 2.2].

(c) The decay rates from Corollary 3 improve once we add regularity assumptions on μ. In the simplest situation $d\mu(\omega) = b(\omega)\, d\omega$ for $b \in W_0^{k,1}(\mathbb{R})$ and $\omega \mapsto \phi_\omega \in W^{k,\infty}(\mathbb{R}; X \cap Y_n)$, the decay rate improves to $(1+|x|)^{(1-n)/2-k}$. This is proved using integration by parts as in the method of stationary phase. In the Appendix we discuss the densities $b(\omega) = (\omega - 1)^{-\delta} \mathbb{1}_{[1,2]}(\omega)$ with $\delta \in (0, 1)$ and find the intermediate decay rates $(1+|x|)^{(1-n)/2-(1-\delta)}$.

(d) Theorem 2(i) tells us that nondispersive solutions of the Schrödinger equation (1) can only occur for radial initial data $\phi(x) = \phi_{\mathrm{rad}}(|x|)$ that satisfy $\|\phi_{\mathrm{rad}}\|_{Y_m} = \infty$ for all $m \in \{0, \dots, n-1\}$. For smooth initial data this essentially means that for some $k \in \{0, \dots, n-1\}$ the function $|\phi_{\mathrm{rad}}^{(k)}(r)|$ does not decay faster than $(1+r)^{-1-k}$ as $r \to \infty$. We conclude that the lack of dispersion is a phenomenon related to slowly decaying or heavily oscillating initial data.

2. Proof of Theorem 2

In the following let ψ denote the unique solution of the Schrödinger equation (1) with initial datum ϕ given by $\phi(x) = \phi_{\mathrm{rad}}(|x|) = \phi_\omega(|x|)e^{i\omega|x|}$. By density, it suffices to prove the estimates for $\phi_\omega \in C_c^\infty(\mathbb{R}_{\geq 0}; \mathbb{C})$. In the following we use the abbreviations

$$f(r) := \phi_{\mathrm{rad}}(r) r^{n/2} J_{(n-2)/2}\left(\frac{r|x|}{2t}\right) \quad \text{and} \quad g(\rho) := f(2\sqrt{t}\rho). \tag{4}$$

We first recall the representation formula of the solution for radial initial data.

Proposition 5. *We have for all $x \in \mathbb{R}^n$, $t > 0$,*

$$\psi(x,t) = |x|^{(2-n)/2} (\sqrt{t})^{-1} e^{i(|x|^2/(4t) - n\pi/4)} \int_0^\infty g(\rho) e^{i\rho^2} \, d\rho. \tag{5}$$

Proof. From (4.2) in [Linares and Ponce 2015] or (2.2.5) in [Cazenave 2003] we get

$$\begin{aligned}
\psi(x,t) &= \frac{1}{(4i\pi t)^{n/2}} \int_{\mathbb{R}^n} e^{i|x-y|^2/(4t)} \phi(y)\,dy \\
&= \frac{1}{(4i\pi t)^{n/2}} \int_0^\infty \phi_{\mathrm{rad}}(r) r^{n-1} e^{i(|x|^2+r^2)/(4t)} \left(\int_{\partial B_1(0)} e^{ir\langle x,\omega\rangle/(2t)}\, d\sigma(\omega) \right) dr \\
&= \frac{1}{(4i\pi t)^{n/2}} \int_0^\infty \phi_{\mathrm{rad}}(r) r^{n-1} e^{i(|x|^2+r^2)/(4t)} \left(\int_{\partial B_1(0)} e^{ir|x|\omega_1/(2t)}\, d\sigma(\omega) \right) dr \\
&= \frac{1}{(4i\pi t)^{n/2}} \int_0^\infty \phi_{\mathrm{rad}}(r) r^{n-1} e^{i(|x|^2+r^2)/(4t)} |\partial B_1(0)| \Gamma\left(\frac{n}{2}\right) 2^{(n-2)/2} \left(\frac{r|x|}{2t}\right)^{(2-n)/2} J_{(n-2)/2}\left(\frac{r|x|}{2t}\right) dr \\
&= |\partial B_1(0)| \Gamma\left(\frac{n}{2}\right) (4\pi i)^{-n/2} 2^{n-2} |x|^{(2-n)/2} t^{-1} e^{i|x|^2/(4t)} \int_0^\infty f(r) e^{ir^2/(4t)}\, dr \\
&= |\partial B_1(0)| \Gamma\left(\frac{n}{2}\right) (4\pi i)^{-n/2} 2^{n-1} |x|^{(2-n)/2} (\sqrt{t})^{-1} e^{i|x|^2/(4t)} \int_0^\infty f(2\sqrt{t}\rho) e^{i\rho^2}\, d\rho.
\end{aligned}$$

So the claim follows from $|\partial B_1(0)| = 2\pi^{n/2}/\Gamma(n/2)$. □

It will be convenient to split the integrand in (5) into three parts, $g = g_1 + g_2 + g_3$. The function g_1 will be identical to g for small arguments and the corresponding estimates rely on the behavior of the Bessel function $J_{(n-2)/2}$ on the interval $[0, 1]$. The sum $g_2 + g_3$ represents g for large arguments and their definitions are based on the asymptotic expansion of the Bessel function at infinity. To be more precise, we fix some cut-off function $\chi \in C_0^\infty(\mathbb{R})$ such that $\chi \equiv 1$ on $\left[0, \frac{1}{2}\right]$ and $\chi \equiv 0$ on $[1, \infty]$. Then, much as in [Watson 1944, p. 202], we write

$$z^{n/2} J_{(n-2)/2}(z) = A_n(z) + e^{iz} B_n(z) + e^{-iz} \overline{B_n(z)}, \tag{6}$$

where the functions A_n, B_n are given by

$$A_n(z) := \chi(z) z^{n/2} J_{(n-2)/2}(z), \qquad B_n(z) := (1-\chi(z)) e^{-i(n-1)\pi/4} \sum_{k=0}^{\infty} \alpha_k z^{(n-1)/2-k} \tag{7}$$

and the coefficients $\alpha_k \in \mathbb{C}$ are $\alpha_0 := 1/\sqrt{2\pi}$ and for $k \in \mathbb{N}$

$$\alpha_k := \frac{1}{\sqrt{2\pi}} \left(\frac{n-2}{2}, k\right) \left(\frac{i}{2}\right)^k, \quad \text{where } (\nu, k) := \frac{(4\nu^2 - 1^2)(4\nu^2 - 3^2)\cdots(4\nu^2 - (2k-1)^2)}{4^k k!}; \tag{8}$$

see [Watson 1944, p. 199]. The motivation for this decomposition is that A_n, B_n and its derivatives satisfy useful uniform estimates that we provide next.

Proposition 6. *We have* $\operatorname{supp}(A_n) \subset [0, 1]$, $\operatorname{supp}(B_n) \subset \left[\frac{1}{2}, \infty\right)$, *and for all* $j \in \mathbb{N}_0$, $z \in \mathbb{R}$,

$$|A_n^{(j)}(z)| \lesssim \begin{cases} |z|^{n-1-j} & \text{if } j \in \{0, \ldots, n-1\}, \\ |z| & \text{if } j \in \{n, n+2, n+4, \ldots\}, \\ 1 & \text{if } j \in \{n+1, n+3, \ldots\}, \end{cases}$$

$$\left| \frac{d^j}{dz^j} \left(B_n(z) z^{(1-n)/2}\right) \right| \lesssim \begin{cases} 1 & \text{if } j = 0, \\ |z|^{-1-j} & \text{if } j \geq 1. \end{cases}$$

Proof. The estimate for A_n follows from

$$
\begin{aligned}
A_n(z) &= \chi(z) z^{n/2} J_{(n-2)/2}(z) \\
&= \chi(z) z^{n/2} \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\,\Gamma(m+n/2)} \left(\frac{z}{2}\right)^{(n-2)/2+2m} \\
&= \chi(z) \sum_{m=0}^{\infty} \frac{(-1)^m}{2^{(n-2)/2+2m} m!\,\Gamma(m+n/2)} z^{n-1+2m}.
\end{aligned}
\tag{9}
$$

The estimate for B_n follows from its series representation (7); see also [Watson 1944, p. 206]. □

Given the definition of g in (4) and the splitting (6) of the Bessel function, we decompose the integrand g according to $g = g_1 + g_2 + g_3$, where, for $r := 2\sqrt{t}\rho$,

$$
\begin{aligned}
g_1(\rho) &:= \phi_{\mathrm{rad}}(r) A_n\left(\frac{r|x|}{2t}\right)\left(\frac{|x|}{2t}\right)^{-n/2} = e^{2i\sqrt{t}\omega\rho}\phi_\omega(r) A_n\left(\frac{r|x|}{2t}\right)\left(\frac{|x|}{2t}\right)^{-n/2}, \\
g_2(\rho) &:= e^{ir|x|/(2t)}\phi_{\mathrm{rad}}(r) B_n\left(\frac{r|x|}{2t}\right)\left(\frac{|x|}{2t}\right)^{-n/2} = e^{2i\sqrt{t}(\omega+|x|/(2t))\rho}\phi_\omega(r) B_n\left(\frac{r|x|}{2t}\right)\left(\frac{|x|}{2t}\right)^{-n/2}, \\
g_3(\rho) &:= e^{-ir|x|/(2t)}\phi_{\mathrm{rad}}(r) \overline{B_n\left(\frac{r|x|}{2t}\right)}\left(\frac{|x|}{2t}\right)^{-n/2} = e^{2i\sqrt{t}(\omega-|x|/(2t))\rho}\phi_\omega(r) \overline{B_n\left(\frac{r|x|}{2t}\right)}\left(\frac{|x|}{2t}\right)^{-n/2}.
\end{aligned}
$$

We now remove the linear phase factors by putting $g_{j,a_j}(\rho) := g_j(\rho)e^{-ia_j\rho}$ for

$$
a_1 := 2\sqrt{t}\omega, \quad a_2 := 2\sqrt{t}\left(\omega + \frac{|x|}{2t}\right), \quad a_3 := 2\sqrt{t}\left(\omega - \frac{|x|}{2t}\right). \tag{10}
$$

This implies, again for $r := 2\sqrt{t}\rho$,

$$
\begin{aligned}
g_{1,a_1}(\rho) &= \phi_\omega(r) A_n\left(\frac{r|x|}{2t}\right)\left(\frac{|x|}{2t}\right)^{-n/2}, \\
g_{2,a_2}(\rho) &= \phi_\omega(r) B_n\left(\frac{r|x|}{2t}\right)\left(\frac{|x|}{2t}\right)^{-n/2}, \\
g_{3,a_3}(\rho) &= \phi_\omega(r) \overline{B_n\left(\frac{r|x|}{2t}\right)}\left(\frac{|x|}{2t}\right)^{-n/2}.
\end{aligned}
\tag{11}
$$

So we infer from Proposition 5

$$
\psi(x,t)|x|^{(n-2)/2}\sqrt{t}e^{-i(|x|^2/(4t)-n\pi/4)} = \int_0^\infty g(\rho)e^{i\rho^2}\,d\rho = \sum_{j=1}^{3}\int_0^\infty g_{j,a_j}(\rho)e^{i(\rho^2+a_j\rho)}\,d\rho. \tag{12}
$$

In order to estimate these terms, we make use of the following auxiliary result.

Proposition 7. *Let $a \in \mathbb{R}$ and $\Xi_a^m \in C^\infty(\mathbb{R};\mathbb{C})$ for $m \in \mathbb{N}_0$ be inductively defined by*

$$
\Xi_a^0(s) := \int_s^\infty e^{i(\rho^2+a\rho)}\,d\rho, \qquad \Xi_a^m(s) := \int_s^\infty \Xi_a^{m-1}(\rho)\,d\rho.
$$

Then $|\Xi_a^m(s)| \le C_m$ for all $a \in \mathbb{R},\ s \ge 0$.

Proof. This follows from $\Xi_a^m(s) = e^{-ia^2/4}\Xi_0^m(s+a/2)$ once we have proved the estimate

$$|\Xi^m(s)| \le C_m(1+s_+)^{-m-1} \quad \text{for all } s \in \mathbb{R},$$

where $\Xi^m := \Xi_0^m$. The existence of the improper Fresnel integral $\Xi^0(s)$ is a well-known consequence of the residue theorem. Moreover, l'Hôpital's rule gives

$$\Xi^0(s)(-2is)e^{-is^2} \to 1 \quad \text{and} \quad s^2(1+2is\,\Xi^0(s)e^{-is^2}) \to \frac{i}{2} \quad \text{as } s \to \infty.$$

(For the second limit one may proceed as we do below in the computation of z_k.) Since the improper integral $\int_s^\infty (1/\rho)e^{i\rho^2}\,d\rho$ exists for $s > 1$ (again by the residue theorem), we obtain from the previous statement that the integral $\Xi^1(s)$ exists and

$$\Xi^1(s)(-2is)^2e^{-is^2} \to 1 \quad \text{and} \quad s^2(1-(-2is)^2\Xi^1(s)e^{-is^2}) \to \frac{3i}{2} \quad \text{as } s \to \infty.$$

By induction, we find that for all $k \in \mathbb{N},\ k \ge 2$ the (proper) integral $\Xi^k(s)$ exists and

$$\begin{aligned}
z_k &:= \lim_{s\to\infty} s^2(1-(-2is)^{k+1}\Xi^k(s)e^{-is^2}) \\
&= \lim_{s\to\infty} \frac{e^{is^2}(-2is)^{-k-1} - \Xi^k(s)}{e^{is^2}s^{-2}(-2is)^{-k-1}} \\
&= \lim_{s\to\infty} \frac{-e^{is^2}(-2is)^{-k} + 2(k+1)ie^{is^2}(-2is)^{-k-2} - (\Xi^k)'(s)}{-e^{is^2}s^{-2}(-2is)^{-k} - 2e^{is^2}s^{-3}(-2is)^{-k-1} + 2(k+1)ie^{is^2}s^{-2}(-2is)^{-k-2}} \\
&= \lim_{s\to\infty} \frac{-(-2is)^{-k} + 2(k+1)i(-2is)^{-k-2} + \Xi^{k-1}(s)e^{-is^2}}{-s^{-2}(-2is)^{-k} - 2s^{-3}(-2is)^{-k-1} + 2(k+1)is^{-2}(-2is)^{-k-2}} \\
&= \lim_{s\to\infty} \frac{-(k+1)is^{-2} - 2 + 2(-2is)^k\,\Xi^{k-1}(s)e^{-is^2}}{-2s^{-2} - 2is^{-4} - (k+1)is^{-4}} \\
&= \lim_{s\to\infty} \frac{(k+1)i + 2s^2(1-(-2is)^k\,\Xi^{k-1}(s)e^{-is^2})}{2 + 2is^{-2} + (k+1)is^{-2}} = \frac{(k+1)i}{2} + z_{k-1},
\end{aligned}$$

implying

$$z_k = \lim_{s\to\infty} s^2(1-(-2is)^{k+1}\Xi^k(s)e^{-is^2}) = \frac{(k+2)(k+1)i}{4}.$$

This yields the bounds for $\Xi^m(s)$ and we are done. □

Let us remark that the proof actually yields the stronger estimate $|\Xi_a^m(s)| \le C_m$ for $0 \le s \le a_-$ and $|\Xi_a^m(s)| \le 2^{m+1}C_m(1+s)^{-m-1}$ for $s \ge a_-$. However, given that these estimates depend on a, it seems difficult to make use of them. Moreover, the independence of a guarantees that our estimates below do not depend on ω since the latter is completely absorbed in the definitions of a_1, a_2, a_3 from (10). From (12) and Proposition 7 we deduce the following estimate for the solution ψ.

Proposition 8. *For all $x \in \mathbb{R}^n,\ t \in \mathbb{R},\ m \in \{0,\dots,n\}$, we have*

$$|\psi(x,t)| \le C_{m-1}|x|^{(2-n)/2}t^{-1/2}\left(\delta_{m,n}|g_{1,a_1}^{(n-1)}(0)| + \int_0^\infty |g_{1,a_1}^{(m)}(\rho)| + |g_{2,a_2}^{(m)}(\rho)| + |g_{3,a_3}^{(m)}(\rho)|\,d\rho\right), \quad (13)$$

where $g_{1,a_1}, g_{2,a_2}, g_{3,a_3}$ are given by (11).

Proof. In the following integration-by-parts scheme we use $(\Xi_{a_j}^m)' = -\Xi_{a_j}^{m-1}$ as well as

$$g_{j,a_j}(0) = g'_{j,a_j}(0) = \cdots = g_{j,a_j}^{(n-2)}(0) = 0, \quad g_{2,a_2}^{(n-1)}(0) = g_{3,a_3}^{(n-1)}(0) = 0, \tag{14}$$

which follows from (11) and Proposition 6. Recall that the support of B_n is contained in $\left[\frac{1}{2}, \infty\right)$ by choice of the cut-off function χ so that the above estimate is actually trivial for $j \in \{2, 3\}$. So we have for $m \in \{0, \dots, n-1\}$

$$\begin{aligned}
\int_0^\infty g_{j,a_j}(\rho) e^{i(\rho^2 + a_j\rho)}\, d\rho &\overset{(14)}{=} \lim_{M\to\infty} \int_0^M \left(\int_0^\rho g'_{j,a_j}(t)\, dt \right) e^{i(\rho^2+a_j\rho)}\, d\rho \\
&= \lim_{M\to\infty} \int_0^M g'_{j,a_j}(t) \left(\int_t^M e^{i(\rho^2+a_j\rho)}\, d\rho \right) dt \\
&= \int_0^\infty g'_{j,a_j}(t)\, \Xi_{a_j}^0(t)\, dt \\
&\overset{(14)}{=} \int_0^\infty \left(\int_0^t g''_{j,a_j}(s)\, ds \right) \Xi_{a_j}^0(t)\, dt \\
&= \int_0^\infty g''_{j,a_j}(s)\, \Xi_{a_j}^1(s)\, ds \\
&\ \vdots \\
&= \int_0^\infty g_{j,a_j}^{(m)}(\rho)\, \Xi_{a_j}^{m-1}(\rho)\, d\rho.
\end{aligned} \tag{15}$$

Notice that the limit $M \to \infty$ passes under the integral because g_{j,a_j} has compact support and the $\Xi_{a_j}^k$ are bounded by Proposition 7. So we obtain for $m \in \{0, \dots, n-1\}$

$$\begin{aligned}
|\psi(x,t)| &\overset{(12)}{\le} |x|^{(2-n)/2} (\sqrt{t})^{-1} \sum_{j=1}^3 \left| \int_0^\infty g_{j,a_j}^{(m)}(\rho)\, \Xi_{a_j}^{m-1}(\rho)\, d\rho \right| \\
&\le C_{m-1} |x|^{(2-n)/2} (\sqrt{t})^{-1} \sum_{j=1}^3 \int_0^\infty |g_{j,a_j}^{(m)}(\rho)|\, d\rho.
\end{aligned}$$

Moreover, using (15) for $m = n-1$ we get

$$\begin{aligned}
|\psi(x,t)| &= |x|^{(2-n)/2} (\sqrt{t})^{-1} \left| \sum_{j=1}^3 \int_0^\infty g_{j,a_j}^{(n-1)}(\rho)\, \Xi_{a_j}^{n-2}(\rho)\, d\rho \right| \\
&= |x|^{(2-n)/2} (\sqrt{t})^{-1} \left| \sum_{j=1}^3 \left(g_{j,a_j}^{(n-1)}(0) \int_0^\infty \Xi_{a_j}^{n-2}(\rho)\, d\rho + \int_0^\infty \left(\int_0^t g_{j,a_j}^{(n)}(\rho)\, d\rho \right) \Xi_{a_j}^{n-2}(t)\, dt \right) \right| \\
&= |x|^{(2-n)/2} (\sqrt{t})^{-1} \left| \sum_{j=1}^3 \left(g_{j,a_j}^{(n-1)}(0)\, \Xi_{a_j}^{n-1}(0) + \int_0^\infty g_{j,a_j}^{(n)}(\rho)\, \Xi_{a_j}^{n-1}(\rho)\, d\rho \right) \right| \\
&\overset{(14)}{\le} C_{n-1} |x|^{(2-n)/2} (\sqrt{t})^{-1} \left(|g_{1,a_1}^{(n-1)}(0)| + \sum_{j=1}^3 \int_0^\infty |g_{j,a_j}^{(n)}(\rho)|\, d\rho \right).
\end{aligned}$$

□

For notational convenience we write $x \lesssim y$, respectively $x \gtrsim y$, instead of $x \le cy$, respectively $x \ge cy$, for positive numbers c that are independent of ω, $|x|$, t, r but may depend on $m \in \{0, \dots, n\}$ or the space dimension $n \in \mathbb{N}$.

Proposition 9. *Let $m \in \{0, \dots, n\}$. Then the functions $g_{1,a_1}, g_{2,a_2}, g_{3,a_3}$ from (11) satisfy the following estimates for all $\rho \ge 0$ and $r = 2\sqrt{t}\rho$:*

$$|g_{1,a_1}^{(m)}(\rho)| \lesssim |x|^{(n-2)/2}(\sqrt{t})^{m-n+2}\sum_{k=0}^{m}|\phi_\omega^{(k)}(r)|r^{n-m+k-1}\cdot \mathbb{1}_{[0,1]}\left(\frac{r|x|}{2t}\right) \quad \text{if } m < n,$$

$$|g_{1,a_1}^{(n)}(\rho)| \lesssim \left(|x|^{(n-2)/2}(\sqrt{t})^{2}\sum_{k=1}^{n}|\phi_\omega^{(k)}(r)|r^{k-1} + |x|^{(n+2)/2}(\sqrt{t})^{-2}|\phi_\omega(r)|r\right)\cdot \mathbb{1}_{[0,1]}\left(\frac{r|x|}{2t}\right),$$

$$|g_{2,a_2}^{(m)}(\rho)| + |g_{3,a_3}^{(m)}(\rho)| \lesssim |x|^{-1/2}(\sqrt{t})^{1+m}\sum_{k=0}^{m}|\phi_\omega^{(k)}(r)|r^{(n-1)/2-m+k}\cdot \mathbb{1}_{[1/2,\infty)}\left(\frac{r|x|}{2t}\right).$$

Proof. We get for $r = 2\sqrt{t}\rho$ and $m \in \{0, \dots, n-1\}$

$$\begin{aligned}
|g_{1,a_1}^{(m)}(\rho)| &\overset{(11)}{=} (2\sqrt{t})^m\left|\frac{d^m}{dr^m}\left(\phi_\omega(r)A_n\left(\frac{r|x|}{2t}\right)\right)\right|\left(\frac{|x|}{2t}\right)^{-n/2}\\
&\lesssim (\sqrt{t})^m\sum_{k=0}^{m}\left|\phi_\omega^{(k)}(r)A_n^{(m-k)}\left(\frac{r|x|}{2t}\right)\right|\left(\frac{|x|}{2t}\right)^{m-k-n/2}\\
&\overset{\text{Prop. 6}}{\lesssim} (\sqrt{t})^m\sum_{k=0}^{m}|\phi_\omega^{(k)}(r)|\left(\frac{r|x|}{2t}\right)^{n-1-m+k}\left(\frac{|x|}{2t}\right)^{m-k-n/2}\cdot \mathbb{1}_{[0,1]}\left(\frac{r|x|}{2t}\right)\\
&\lesssim |x|^{(n-2)/2}(\sqrt{t})^{m-n+2}\sum_{k=0}^{m}|\phi_\omega^{(k)}(r)|r^{n-1-m+k}\cdot \mathbb{1}_{[0,1]}\left(\frac{r|x|}{2t}\right).
\end{aligned}$$

This implies the first estimate. For $m = n$ we use the estimate for $A_n^{(j)}$ from Proposition 6 for $j \in \{0, \dots, n\}$ and obtain

$$\begin{aligned}
|g_{1,a_1}^{(n)}(\rho)| &\overset{(11)}{=} (2\sqrt{t})^n\left|\frac{d^n}{dr^n}\left(\phi_\omega(r)A_n\left(\frac{r|x|}{2t}\right)\right)\right|\left(\frac{|x|}{2t}\right)^{-n/2}\\
&\lesssim (\sqrt{t})^n\sum_{k=0}^{n}\left|\phi_\omega^{(k)}(r)A_n^{(n-k)}\left(\frac{r|x|}{2t}\right)\right|\left(\frac{|x|}{2t}\right)^{n/2-k}\\
&\lesssim (\sqrt{t})^n\left(\sum_{k=1}^{n}|\phi_\omega^{(k)}(r)|\left(\frac{r|x|}{2t}\right)^{-1+k}\left(\frac{|x|}{2t}\right)^{n/2-k} + |\phi_\omega(r)|\frac{r|x|}{2t}\left(\frac{|x|}{2t}\right)^{n/2}\right)\cdot \mathbb{1}_{[0,1]}\left(\frac{r|x|}{2t}\right)\\
&\lesssim \left(|x|^{(n-2)/2}(\sqrt{t})^{2}\sum_{k=1}^{n}|\phi_\omega^{(k)}(r)|r^{k-1} + |x|^{(n+2)/2}(\sqrt{t})^{-2}|\phi_\omega(r)|r\right)\cdot \mathbb{1}_{[0,1]}\left(\frac{r|x|}{2t}\right).
\end{aligned}$$

This yields the second estimate. The third estimate results from

$$|B_n^{(j)}(z)| \lesssim |z|^{(n-1)/2-j}, \tag{16}$$

which is a consequence of Proposition 6. Thus

$$
\begin{aligned}
|g_{2,a_2}^{(m)}(\rho)|+|g_{2,a_3}^{(m)}(\rho)| &\overset{(11)}{=} 2(2\sqrt{t})^m\left|\frac{d^m}{dr^m}\left(\phi_\omega(r)B_n\left(\frac{r|x|}{2t}\right)\right)\right|\left(\frac{|x|}{2t}\right)^{-n/2}\\
&\lesssim (\sqrt{t})^m\sum_{k=0}^m|\phi_\omega^{(k)}(r)|\left|B_n^{(m-k)}\left(\frac{r|x|}{2t}\right)\right|\left(\frac{|x|}{2t}\right)^{m-k-n/2}\\
&\overset{(16)}{\lesssim} (\sqrt{t})^m\sum_{k=0}^m|\phi_\omega^{(k)}(r)|\left(\frac{r|x|}{2t}\right)^{(n-1)/2-m+k}\left(\frac{|x|}{2t}\right)^{m-k-n/2}\cdot\mathbb{1}_{[1/2,\infty)}\left(\frac{r|x|}{2t}\right)\\
&\lesssim |x|^{-1/2}(\sqrt{t})^{m+1}\sum_{k=0}^m|\phi_\omega^{(k)}(r)|r^{(n-1)/2-m+k}\cdot\mathbb{1}_{[1/2,\infty)}\left(\frac{r|x|}{2t}\right). \qquad\square
\end{aligned}
$$

Proof of Theorem 2(i). We combine the estimates from Propositions 8 and 9. Under the assumption $\phi_\omega\in C_c^\infty(\mathbb{R}_{\geq 0};\mathbb{C})$ we get for all $m\in\{0,\dots,n-1\}$

$$
\begin{aligned}
|\psi(x,t)| &\overset{\text{Prop. }8}{\lesssim} |x|^{(2-n)/2}(\sqrt{t})^{-1}\int_0^\infty |g_{1,a_1}^{(m)}(\rho)|+|g_{2,a_2}^{(m)}(\rho)|+|g_{3,a_3}^{(m)}(\rho)|\,d\rho\\
&\overset{\text{Prop. }9}{\lesssim} (\sqrt{t})^{m-n+1}\sum_{k=0}^m\int_0^{\sqrt{t}/|x|}|\phi_\omega^{(k)}(2\sqrt{t}\rho)|(2\sqrt{t}\rho)^{n-m+k-1}\,d\rho\\
&\quad+|x|^{(1-n)/2}(\sqrt{t})^m\sum_{k=0}^m\int_{\sqrt{t}/(2|x|)}^\infty|\phi_\omega^{(k)}(2\sqrt{t}\rho)|(2\sqrt{t}\rho)^{(n-1)/2-m+k}\,d\rho\\
&\lesssim (\sqrt{t})^{m-n}\sum_{k=0}^m\int_0^{2t/|x|}|\phi_\omega^{(k)}(r)|r^{n-m+k-1}\,dr\\
&\quad+(\sqrt{t})^{m-n}\left(\frac{t}{|x|}\right)^{(n-1)/2}\sum_{k=0}^m\int_{t/|x|}^\infty|\phi_\omega^{(k)}(r)|r^{(n-1)/2-m+k}\,dr\\
&\lesssim (\sqrt{t})^{m-n}\sum_{k=0}^m\left(\int_0^{2t/|x|}|\phi_\omega^{(k)}(r)|r^{n-m+k-1}\,dr+\int_{t/|x|}^\infty|\phi_\omega^{(k)}(r)|r^{n-m+k-1}\,dr\right)\\
&\lesssim (\sqrt{t})^{m-n}\|\phi_\omega\|_{Y_m}.
\end{aligned}
$$

In the case $m=n$ we use the second estimate in Proposition 9 instead of the first one. By density of $C_c^\infty(\mathbb{R}_{\geq 0};\mathbb{C})$ in Y_m the result follows. $\square$

Proof of Theorem 2(ii). For $r=2\sqrt{t}\rho$ we use

$$
|g'_{1,a_1}(\rho)|\lesssim |x|^{(n-2)/2}(\sqrt{t})^{3-n}(|\phi_\omega(r)|r^{n-2}+|\phi'_\omega(r)|r^{n-1})\cdot\mathbb{1}_{[0,1]}\left(\frac{r|x|}{2t}\right),
$$

as well as

$$
\begin{aligned}
&|g'_{2,a_2}(\rho)|+|g'_{2,a_3}(\rho)|\\
&\quad\overset{(11)}{=} 4\sqrt{t}\left|\frac{d}{dr}\left(\phi_\omega(r)B_n\left(\frac{r|x|}{2t}\right)\right)\right|\left(\frac{|x|}{2t}\right)^{-n/2}\\
&\quad\lesssim \sqrt{t}\left|\frac{d}{dr}(\phi_\omega(r)r^{(n-1)/2})\right|\left|r^{(1-n)/2}B_n\left(\frac{r|x|}{2t}\right)\right|\left(\frac{|x|}{2t}\right)^{-n/2}\\
&\qquad+\sqrt{t}|\phi_\omega(r)|r^{(n-1)/2}\left|\frac{d}{dr}\left(r^{(1-n)/2}B_n\left(\frac{r|x|}{2t}\right)\right)\right|\left(\frac{|x|}{2t}\right)^{-n/2}
\end{aligned}
$$

$$\overset{\text{Prop. 6}}{\lesssim} \sqrt{t}\left|\frac{d}{dr}(\phi_\omega(r)r^{(n-1)/2})\right| r^{(1-n)/2}\left(\frac{r|x|}{2t}\right)^{(n-1)/2}\left(\frac{|x|}{2t}\right)^{-n/2}\cdot \mathbb{1}_{[1/2,\infty)}\left(\frac{r|x|}{2t}\right)$$
$$+\sqrt{t}|\phi_\omega(r)|r^{(n-1)/2}\left(\frac{r|x|}{2t}\right)^{-2}\left(\frac{|x|}{2t}\right)^{1/2}\cdot \mathbb{1}_{[1/2,\infty)}\left(\frac{r|x|}{2t}\right)$$
$$\lesssim \left(|x|^{-1/2}(\sqrt{t})^2\left|\frac{d}{dr}(\phi_\omega(r)r^{(n-1)/2})\right| + |x|^{-3/2}(\sqrt{t})^4|\phi_\omega(r)|r^{(n-5)/2}\right)\cdot \mathbb{1}_{[1/2,\infty)}\left(\frac{r|x|}{2t}\right).$$

This implies

$$|\psi(x,t)| \overset{\text{Prop. 8}}{\lesssim} |x|^{(2-n)/2}(\sqrt{t})^{-1}\int_0^\infty |g^{(m)}_{1,a_1}(\rho)| + |g^{(m)}_{2,a_2}(\rho)| + |g^{(m)}_{3,a_3}(\rho)|\,d\rho$$
$$\overset{\text{Prop. 9}}{\lesssim} (\sqrt{t})^{2-n}\int_0^{\sqrt{t}/|x|}\left(|\phi_\omega(2\sqrt{t}\rho)|(2\sqrt{t}\rho)^{n-2} + |\phi'_\omega(2\sqrt{t}\rho)|(2\sqrt{t}\rho)^{n-1}\right)d\rho$$
$$+|x|^{(1-n)/2}\sqrt{t}\int_{\sqrt{t}/(2|x|)}^\infty \left|\frac{d}{dr}(\phi_\omega(r)r^{(n-1)/2})\right|_{r=2\sqrt{t}\rho} d\rho$$
$$+|x|^{(-1-n)/2}(\sqrt{t})^3\int_{\sqrt{t}/(2|x|)}^\infty |\phi_\omega(2\sqrt{t}\rho)|(2\sqrt{t}\rho)^{(n-5)/2}\,d\rho$$
$$\lesssim |x|^{(1-n)/2}\left(\frac{t}{|x|}\right)^{(1-n)/2}\int_0^{2t/|x|}(|\phi_\omega(r)|r^{n-2} + |\phi'_\omega(r)|r^{n-1})\,dr$$
$$+|x|^{(1-n)/2}\left(\int_{t/|x|}^\infty \left|\frac{d}{dr}(\phi_\omega(r)r^{(n-1)/2})\right| dr + \frac{t}{|x|}\int_{t/|x|}^\infty |\phi_\omega(r)|r^{(n-5)/2}\,dr\right)$$
$$\lesssim |x|^{(1-n)/2}\|\phi_\omega\|_X.$$

So we get the result by density of $C_c^\infty(\mathbb{R}_{\geq 0};\mathbb{C})$ in X. □

3. Proof of Theorem 1

We estimate the solution of the Schrödinger equation for the initial datum

$$\phi(x) = \phi_{\text{rad}}(|x|), \qquad \text{where } \phi_{\text{rad}}(\rho) := e^{-i\rho^2/4}\mathbb{1}_{\rho\geq 1}\rho^{-\sigma}$$

and σ is chosen according to $(n-3)/2 < \sigma < n$. In this case, the formula (5) is well-defined and provides a solution of the initial value problem (1). In [Bona et al. 2014, Section 2.1] it was shown that the solution ψ blows up in $L^\infty(\mathbb{R}^n)$ as $t \to 1$ provided that $n/2 < \sigma < n$ holds. In fact, in this case the function

$$a(x) := \mathbb{1}_{|x|\geq 1}|x|^{-\sigma}$$

lies in $L^2(\mathbb{R}^n)$ but not in $L^1(\mathbb{R}^n)$ so that [Bona et al. 2014, Remark 2.2] applies. We now generalize this analysis to the range $(n-3)/2 < \sigma < n$ and detect a self-similar blow-up in $L^q(\mathbb{R}^n)$ for all $q > n/(n-\sigma)$, and a lower estimate for the corresponding blow-up rate then implies $\|\psi\|_{L^p(\mathbb{R},L^q(\mathbb{R}^n))} = \infty$ for $p \geq 2q/((n-\sigma)q-n)_+$. From this we will finally deduce the nonvalidity of Strichartz estimates for initial data $\phi \in L^r(\mathbb{R}^n)$, where $r > 2$.

We set $k_t := \sqrt{1/(4t) - 1/4}$ for $0 \le t < 1$ and write $\psi(x) = \psi_{\mathrm{rad}}(|x|)$. We get for $|x| = 2tk_t z$

$$\begin{aligned}
2|\psi(x,t)|k_t^{n-\sigma} &\overset{(5)}{=} |x|^{(2-n)/2} t^{-1} k_t^{n-\sigma} \left| \int_0^\infty J_{(n-2)/2}\left(\frac{\rho|x|}{2t}\right) \phi_{\mathrm{rad}}(\rho)\rho^{n/2} e^{i\rho^2/(4t)}\, d\rho \right| \\
&= |x|^{(2-n)/2} t^{-1} k_t^{n-\sigma} \left| \int_1^\infty J_{(n-2)/2}\left(\frac{\rho|x|}{2t}\right) \rho^{n/2-\sigma} e^{i\rho^2 k_t^2}\, d\rho \right| \\
&= (2tk_t z)^{(2-n)/2} t^{-1} k_t^{n-\sigma} \left| \int_1^\infty J_{(n-2)/2}(\rho k_t z) \rho^{n/2-\sigma} e^{i\rho^2 k_t^2}\, d\rho \right| \\
&= (2tk_t z)^{(2-n)/2} t^{-1} k_t^{(n-2)/2} \left| \int_{k_t}^\infty J_{(n-2)/2}(sz) s^{n/2-\sigma} e^{is^2}\, ds \right| \\
&\to (2z)^{(2-n)/2} \left| \int_0^\infty J_{(n-2)/2}(sz) s^{n/2-\sigma} e^{is^2}\, ds \right| \quad \text{as } t \to 1
\end{aligned}$$

and the convergence is locally uniform in $z \in (0,\infty)$ due to $(n-3)/2 < \sigma < n$. Since the right-hand side is not identically zero we may find $\delta > 0$ and radii $0 < R_1 < R_2$ such that

$$|\psi(x,t)| \gtrsim k_t^{\sigma-n} \quad \text{for } R_1 k_t \le |x| \le R_2 k_t \text{ and } 1-\delta < t < 1.$$

Hence we get

$$\begin{aligned}
\int_{1-\delta}^1 \left(\int_{B_{R_2 k_t}(0)\setminus B_{R_1 k_t}(0)} |\psi(x,t)|^q\, dx \right)^{p/q} dt &\gtrsim \int_{1-\delta}^1 \left(\int_{R_1 k_t}^{R_2 k_t} r^{n-1} k_t^{(\sigma-n)q}\, dr \right)^{p/q} dt \\
&\gtrsim \int_{1-\delta}^1 (k_t^{n+(\sigma-n)q})^{p/q}\, dt \\
&\gtrsim \int_{1-\delta}^1 (1-t)^{p(n+(\sigma-n)q)/(2q)}\, dt.
\end{aligned}$$

This integral is finite if and only if $(p/(2q))(n + (\sigma - n)q) > -1$. So $\psi \in L^p(\mathbb{R}; L^q(\mathbb{R}^n))$ can only hold for $p < 2q/((n-\sigma)q - n)_+$. Moreover, the initial datum lies in $L^r(\mathbb{R}^n)$ if and only if $\sigma > n/r$. So, for any $r > 1$ we can consider the limit $\sigma \searrow \max\{n/r, (n-3)/2\}$, and we find that the validity of the Strichartz estimate (2) with initial datum in $L^r(\mathbb{R}^n)$, $r > 1$, implies

$$p \le \frac{2q}{((n - \max\{n/r, (n-3)/2\})q - n)_+} = \max\left\{ \frac{2qr}{n((r-1)q - r)_+}, \frac{4q}{((n+3)q - 2n)_+} \right\}. \tag{17}$$

On the other hand, the scaling invariance of the Schrödinger equation implies $2/p + n/q = n/r$ and thus $p = 2qr/(n(q-r))$. Plugging this into (17) we obtain $r \le 2$. Hence, the Strichartz estimate (2) cannot hold for any $r > 2$, which finishes the proof. □

Appendix

In this appendix we briefly discuss the restriction $p \ge 2$ in the context of Strichartz estimates of the form

$$\|e^{it\Delta}\phi\|_{L_t^p(\mathbb{R}; L^q(\mathbb{R}^n))} \lesssim \|\phi\|_{L^2(\mathbb{R}^n)},$$

which results from an abstract reasoning involving translation-invariant operators due to [Hörmander 1960]; see [Keel and Tao 1998, p. 970–971]. Here we provide a family of explicit counterexamples that not only implies $p \geq 2$ for square integrable initial data, but even shows $p \geq 2r/(2n - r(n-1))_+$ for initial data in $L^r(\mathbb{R}^n)$ with $r > 2n/(n+1)$. In order to avoid lengthy computations involving oscillatory integrals, we only sketch the proofs. The starting point is a reasonable choice of an initial condition. We choose

$$\phi(x) = |x|^{(2-n)/2} \int_1^2 (\omega - 1)^{-\delta} J_{(n-2)/2}(\omega|x|)\, d\omega$$

for $\delta \in (0, 1)$. This function corresponds to a singular superposition of Herglotz waves; see Remark 1(c). It is smooth and lengthy computations involving the van der Corput lemma [Stein 1993, p. 334] reveal

$$|\phi(x)| \sim 2|x|^{(1-n)/2-(1-\delta)} \int_0^\infty \mathrm{Re}(e^{i\rho}\alpha_0 e^{i(|x|-((n-1)\pi)/4)})\rho^{-\delta}\, d\rho \quad \text{as } |x| \to \infty,$$

where $\alpha_0 > 0$ is the dominant term in the series expansion of the Bessel function near infinity; see (8). In particular we get $\phi \in L^r(\mathbb{R}^n)$ if and only if $\delta < (n+1)/2 - n/r$.

The above choice for the initial datum allows to write down the corresponding solution of the Schrödinger equation semiexplicitly via

$$\widehat{\psi(\cdot\,, t)}(\xi) = e^{-it|\xi|^2}\hat{\phi}(\xi) = e^{-it|\xi|^2}|\xi|^{-n/2}(|\xi| - 1)^{-\delta}\mathbb{1}_{[1,2]}(|\xi|).$$

Hence, one gets

$$\psi(x, t) = |x|^{(2-n)/2} \int_1^2 e^{-it\omega^2}(\omega - 1)^{-\delta} J_{(n-2)/2}(\omega|x|)\, d\omega,$$

and the van der Corput Lemma implies $|\psi(x, t)| \gtrsim ct^{\delta-1}$ for small $|x|$ and large t. In particular, $\psi \in L^p(\mathbb{R}; L^q(\mathbb{R}^n))$ implies $p(1-\delta) > 1$. So we conclude that for any $r \in (2n/(n+1), \infty]$ we can consider the limit $\delta \nearrow \min\{1, (n+1)/2 - n/r\}$ in the above computations and obtain that $\psi \in L^p(\mathbb{R}; L^q(\mathbb{R}^n))$ implies

$$p \geq \frac{2r}{(2n - r(n-1))_+} \quad \text{provided } r > \frac{2n}{n+1}. \tag{18}$$

For $r = 2$, i.e., for square integrable initial data, this implies $p \geq 2$, which is all we wanted to demonstrate.

Let us mention that a detailed analysis of ψ reveals $\psi \in L^p(\mathbb{R}; L^q(\mathbb{R}^n))$ if and only if

$$q > \max\left\{\frac{2n-1}{n-\delta}, \frac{2n}{n+1-2\delta}\right\},$$
$$p > \max\left\{\frac{1}{1-\delta}, \frac{2q}{q(n-\delta)+1-2n}, \frac{2q}{q(n+1-2\delta)-2n}\right\}.$$

Keeping the scaling condition $2/p + n/q = n/r$ in mind, these a priori more restrictive conditions do however not result in stronger necessary conditions than (18), so that further necessary conditions cannot be deduced from this example.

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID 258734477-SFB 1173. The author thanks Dominic Scheider (Karlsruhe Institute of Technology) and Martin Spitz (University of Bielefeld) for valuable remarks and discussions leading to an improvement of the manuscript.

References

[Bona et al. 2014] J. L. Bona, G. Ponce, J.-C. Saut, and C. Sparber, "Dispersive blow-up for nonlinear Schrödinger equations revisited", *J. Math. Pures Appl.* (9) **102**:4 (2014), 782–811. MR Zbl

[Cazenave 2003] T. Cazenave, *Semilinear Schrödinger equations*, Courant Lecture Notes in Mathematics **10**, Courant Institute of Mathematical Sciences, New York, 2003. MR Zbl

[Ginibre and Velo 1985] J. Ginibre and G. Velo, "The global Cauchy problem for the nonlinear Schrödinger equation revisited", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **2**:4 (1985), 309–327. MR Zbl

[Hörmander 1960] L. Hörmander, "Estimates for translation invariant operators in L^p spaces", *Acta Math.* **104** (1960), 93–140. MR Zbl

[Keel and Tao 1998] M. Keel and T. Tao, "Endpoint Strichartz estimates", *Amer. J. Math.* **120**:5 (1998), 955–980. MR Zbl

[Linares and Ponce 2015] F. Linares and G. Ponce, *Introduction to nonlinear dispersive equations*, 2nd ed., Springer, 2015. MR Zbl

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series **43**, Princeton University Press, 1993. MR Zbl

[Strichartz 1977] R. S. Strichartz, "Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations", *Duke Math. J.* **44**:3 (1977), 705–714. MR Zbl

[Watson 1944] G. N. Watson, *A treatise on the theory of Bessel functions*, Cambridge University Press, 1944. MR Zbl

Received 3 Nov 2019. Revised 21 Jan 2020. Accepted 21 Feb 2020.

RAINER MANDEL: rainer.mandel@kit.edu
Karlsruhe Institute of Technology, Institute for Analysis, Karlsruhe, Germany

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the submission page.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles are usually in English or French, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not refer to bibliography keys. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and a Mathematics Subject Classification for the article, and, for each author, affiliation (if appropriate) and email address.

Format. Authors are encouraged to use LaTeX and the standard amsart class, but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should normally be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages — Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc. — allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with as many details as you can about how your graphics were generated.

Bundle your figure files into a single archive (using zip, tar, rar or other format of your choice) and upload on the link you been provided at acceptance time. Each figure should be captioned and numbered so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript with all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables.

White Space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PURE and APPLIED ANALYSIS

vol. 2 no. 2 2020