Analysis & PDE

Volume 13 No. 1 2020

msp

Analysis & PDE
msp.org/apde

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

The subscription price for 2020 is US $340/year for the electronic version, and $550/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 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.

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 13, No. 1, 2020
dx.doi.org/10.2140/apde.2020.13.1

msp

ABSENCE OF CARTAN SUBALGEBRAS FOR RIGHT-ANGLED HECKE VON NEUMANN ALGEBRAS

MARTIJN CASPERS

For a right-angled Coxeter system (W, S) and $q > 0$, let $\mathcal{M}_q$ be the associated Hecke von Neumann algebra, which is generated by self-adjoint operators T_s, $s \in S$, satisfying the Hecke relation $(\sqrt{q}T_s - q)(\sqrt{q}T_s + 1) = 0$, as well as suitable commutation relations. Under the assumption that (W, S) is irreducible and $|S| \geq 3$ it was proved by Garncarek (*J. Funct. Anal.* **270**:3 (2016), 1202–1219) that $\mathcal{M}_q$ is a factor (of type II_1) for a range $q \in [\rho, \rho^{-1}]$ and otherwise $\mathcal{M}_q$ is the direct sum of a II_1-factor and $\mathbb{C}$.

In this paper we prove (under the same natural conditions as Garncarek) that $\mathcal{M}_q$ is noninjective, that it has the weak-$*$ completely contractive approximation property and that it has the Haagerup property. In the hyperbolic factorial case $\mathcal{M}_q$ is a strongly solid algebra and consequently $\mathcal{M}_q$ cannot have a Cartan subalgebra. In the general case $\mathcal{M}_q$ need not be strongly solid. However, we give examples of nonhyperbolic right-angled Coxeter groups such that $\mathcal{M}_q$ does not possess a Cartan subalgebra.

1. Introduction

Hecke algebras are one-parameter deformations of group algebras of a Coxeter group. They were the foundation for the theory of quantum groups [Jimbo 1986; Kassel 1995] and have remarkable applications in the theory of knot invariants, as was shown by V. Jones [1985]. A wide range of applications of Coxeter groups and their Hecke deformations can be found in [Davis 2008]. Dymara [2006] (see also [Davis 2008, Section 19]) introduced the von Neumann algebras generated by Hecke algebras. Many important results were then obtained (see also [Davis et al. 2007]) for these Hecke von Neumann algebras. This gave for example insight into the cohomology of associated constructions and its Betti numbers. In this paper we investigate the approximation properties of Hecke von Neumann algebras as well as their Cartan subalgebras (here we mean the notion of a Cartan subalgebra in the von Neumann algebraic sense, which we recall in Section 5, and not the Lie algebraic notion).

Let us recall the following definition. Let $q > 0$ and let W be a right-angled Coxeter group with generating set S (see Section 2). The associated Hecke algebra is a $*$-algebra generated by T_s, $s \in S$, which satisfies the relation

$$(\sqrt{q}\, T_s - q)(\sqrt{q}\, T_s + 1) = 0, \qquad T_s^* = T_s, \qquad \text{and} \qquad T_s T_t = T_t T_s$$

for $s, t \in S$ with $st = ts$. Hecke algebras carry a canonical faithful tracial vector state (the vacuum state) and therefore generate a von Neumann algebra $\mathcal{M}_q$ under its GNS construction. It was recently proved by Garncarek [2016] that if (W, S) is irreducible (see Section 2) and $|S| \geq 3$, the von Neumann

MSC2010: 47L10.
Keywords: Hecke von Neumann algebras, approximation properties, Cartan subalgebras.

algebra $\mathcal{M}_q$ is a factor in the case $q \in [\rho, \rho^{-1}]$, where ρ is the radius of convergence of the fundamental power series (2-2). If $q \notin [\rho, \rho^{-1}]$ then $\mathcal{M}_q$ is the direct sum of a II_1-factor and $\mathbb{C}$. For more general Coxeter groups/Hecke algebras (not necessarily being right-angled, or for multiparameters q) this result is unknown. It deserves to be emphasized that this in particular shows that the isomorphism class of $\mathcal{M}_q$ depends on q; an observation that was already made in the final remarks of [Davis 2008, Section 19].

The first aim of this paper is to determine approximation properties of $\mathcal{M}_q$ (assuming the same natural conditions as Garncarek). We first show that $\mathcal{M}_q$ is a noninjective von Neumann algebra and therefore falls outside Connes' classification of hyperfinite factors [1976]. Secondly we show that $\mathcal{M}_q$ has the weak-$*$ completely contractive approximation property (wk-$*$ CCAP). This means that there exists a net of completely contractive finite-rank maps on $\mathcal{M}_q$ that converges to the identity in the point σ-weak topology. In case $q = 1$ the algebra $\mathcal{M}_q$ is the group von Neumann algebra of a right-angled Coxeter group. In this case the result was known. For instance the CCAP follows from Reckwerdt's result [2015] and noninjectivity follows easily from identifying a copy of the free group inside W. Noninjectivity can also be proved for right-angled Coxeter groups through the techniques developed in [Bożejko and Speicher 1994]. Here we find the following:

Theorem A. *Let* $q > 0$:

(1) *Let* (W, S) *be an irreducible right-angled Coxeter system with* $|S| \geq 3$. *Then* $\mathcal{M}_q$ *is noninjective.*

(2) *For a general right-angled Coxeter system* (W, S) *the associated Hecke von Neumann algebra* $\mathcal{M}_q$ *has the wk-$*$ CCAP and the Haagerup property.*

The proofs of noninjectivity and the Haagerup property proceed by showing that Hecke von Neumann algebras are actually graph products [Caspers and Fima 2017] and then using general graph/free-product techniques involving important results of [Ueda 2011]. For the wk-$*$- CCAP we first obtain cb-estimates for radial multipliers and then use estimates of word-length projections (see Proposition 4.11) going back to [Haagerup 1978].

Our second aim is the study of Cartan subalgebras of the Hecke von Neumann algebra $\mathcal{M}_q$. Recall that a Cartan subalgebra of a II_1-factor is by definition a maximal abelian subalgebra whose normalizer generates the II_1-factor itself. Cartan subalgebras arise typically in crossed products of free ergodic probability measure preserving actions of discrete groups on a probability measure space.

Voiculescu [1996] was the first one to find factors (namely free group factors) that do not have a Cartan subalgebra. His proof relies on estimates for the free entropy dimension of the normalizer of an injective von Neumann algebra. Using a different approach, Ozawa and Popa [2010] were also able to find classes of von Neumann algebras that do not have a Cartan subalgebra (including the free group factors). Ozawa and Popa actually proved that these algebras have a stronger property that afterwards became known as strong solidity: the normalizer of a diffuse injective von Neumann subalgebra generates an injective von Neumann algebra again.

After these fundamental results by Ozawa and Popa, strong solidity was studied for many other von Neumann algebras. In particular Popa and Vaes [2014] (see also [Chifan and Sinclair 2013]) proved the absence of Cartan subalgebras for group factors of biexact groups that have the CBAP. Isono [2015] then put the results from [Popa and Vaes 2014] into a general von Neumann framework in order to prove absence of Cartan subalgebras for free orthogonal quantum groups. Isono proved that factors with the wk-$*$ CBAP that satisfy condition $(AO)^+$ are strongly solid. Using this strong solidity result by Isono we are able to prove the following.

Theorem B. *Let $q \in [\rho, \rho^{-1}]$ with ρ as in Theorem 2.2. Let (W, S) be an irreducible right-angled Coxeter system with $|S| \geq 3$. Assume that W is hyperbolic. Then the associated Hecke von Neumann algebra $\mathcal{M}_q$ is strongly solid.*

In turn as $\mathcal{M}_q$ is noninjective by Theorem A we are able to derive the result announced in the title of this paper for the hyperbolic case.

Corollary C. *Let $q \in [\rho, \rho^{-1}]$ with ρ as in Theorem 2.2. For an irreducible right-angled hyperbolic Coxeter system (W, S) with $|S| \geq 3$ the associated Hecke von Neumann algebra $\mathcal{M}_q$ does not have a Cartan subalgebra.*

General right-angled Hecke von Neumann algebras are not strongly solid; see Remark 5.6. Still we can prove in some cases that they do not possess a Cartan subalgebra. We do this by showing that if $\mathcal{M}_q$ were to have a Cartan subalgebra then under suitable conditions each of the three alternatives in [Vaes 2014, Theorem A] fails to be true, which leads to a contradiction.

Theorem D. *Let $q \in [\rho, \rho^{-1}]$. Let (W, S) be an irreducible right-angled Coxeter system with $|S| \geq 3$ for which the Coxeter graph satisfies the conditions of Theorem 6.7. Then the associated Hecke von Neumann algebra $\mathcal{M}_q$ does not have a Cartan subalgebra.*

Structure. In Section 2 we introduce Hecke von Neumann algebras and some basic algebraic properties. Lemma 2.7 is crucial for the results on strong solidity and the weak-$*$ CCAP. In Section 3 we obtain universal properties of Hecke von Neumann algebras and prove that they decompose as graph products. We collect the consequences for the Haagerup property and noninjectivity. In Section 4 we find approximation properties of $\mathcal{M}_q$ and conclude Theorem A. Section 5 proves the strong solidity result of Theorem B from which Corollary C shall easily follow. Finally Section 6 proves absence of Cartan subalgebras for the cases of Theorem D.

Convention. Let X be a set and let $A, B \subseteq X$. We will briefly write $A \backslash B$ for $A \backslash (A \cap B)$.

2. Notation and preliminaries

Standard results on operator spaces can be found in [Effros and Ruan 2000; Pisier 2003]. Standard references for von Neumann algebras are [Strătilă and Zsidó 1979; Takesaki 1979]. Recall that *ucp* stands for unital completely positive.

2A. *Coxeter groups.* A *Coxeter group* W is a group that is freely generated by a finite set S subject to relations

$$(st)^{m(s,t)} = 1$$

for some constant $m(s,t) \in \{1, 2, \dots, \infty\}$ with $m(s,t) = m(t,s) \geq 2,\ s \neq t$, and $m(s,s) = 1$. The constant $m(s,t) = \infty$ means that no relation is imposed, so that s, t are free variables. The Coxeter group W is called right-angled if either $m(s,t) = 2$ or $m(s,t) = \infty$ for all $s, t \in S,\ s \neq t$ and this is the only case we need in this paper. Therefore we assume from now on that W is a right-angled Coxeter group with generating set S. The pair (W, S) is also called a Coxeter system.

Let $\boldsymbol{w} \in W$ and suppose that $\boldsymbol{w} = w_1 \cdots w_n$ with $w_i \in S$. The representing expression $w_1 \cdots w_n$ is called reduced if whenever also $\boldsymbol{w} = w'_1 \cdots w'_m$ with $w'_i \in S$ then $n \leq m$; i.e., the expression is of minimal length. In that case we will write $|\boldsymbol{w}| = n$. Reduced expressions are not necessarily unique (only if $m(s,t) = \infty$ whenever $s \neq t$), but for each $\boldsymbol{w} \in W$ we may pick a reduced expression which we shall call minimal.

Convention. For $\boldsymbol{w} \in W$ we shall write w_i for the minimal representative $\boldsymbol{w} = w_1 \cdots w_n$.

To the pair (W, S) we associate a graph Γ with vertex set $V\Gamma = S$ and edge set $E\Gamma = \{(s,t) : m(s,t) = 2\}$. A subgraph Γ_0 of Γ is called *full* if the following property holds: for all $s, t \in V\Gamma_0$ with $(s,t) \in E\Gamma$ we have $(s,t) \in E\Gamma_0$.

A clique in Γ is a full subgraph in which every two vertices share an edge. We let $\mathsf{Cliq}(\Gamma)$ denote the set of cliques in Γ. To keep the notation consistent with the literature the empty graph is in $\mathsf{Cliq}(\Gamma)$ by convention (in this paper we shall sometimes exclude the empty graph from $\mathsf{Cliq}(\Gamma)$ explicitly or treat it as a special case to keep some of the arguments more transparent).

For $s \in S$ we set

$$\mathsf{Link}(s) = \{t \in S : m(s,t) = 2\},$$

so these are all vertices in Γ that have distance exactly 1 to s. For a subset $X \subseteq V\Gamma$ we set $\mathsf{Link}(X) = \bigcap_{s \in X} \mathsf{Link}(s)$. We sometimes regard $\mathsf{Link}(X)$ as a full subgraph of Γ.

Definition 2.1. A Coxeter system (W, S) is called *irreducible* if the complement of Γ is connected. Here the complement Γ^c of the graph Γ is the graph with the same vertex set $V\Gamma$ and for $v, w \in V\Gamma$ we have $(v, w) \in E\Gamma^c$ if and only if $(v, w) \notin E\Gamma$.

2B. *Hecke von Neumann algebras.* Let (W, S) be a right-angled Coxeter system. Let $q > 0$. By [Davis 2008, Proposition 19.1.1] there exists a unique unital $*$-algebra $\mathbb{C}_q(\Gamma)$ generated by a basis $\{\widetilde{T}_{\boldsymbol{w}} : \boldsymbol{w} \in W\}$ satisfying the following relations. For every $s \in S$ and $\boldsymbol{w} \in W$ we have

$$\widetilde{T}_s \widetilde{T}_{\boldsymbol{w}} = \begin{cases} \widetilde{T}_{s\boldsymbol{w}} & \text{if } |s\boldsymbol{w}| > |\boldsymbol{w}|, \\ q\widetilde{T}_{s\boldsymbol{w}} + (q-1)\widetilde{T}_{\boldsymbol{w}} & \text{otherwise,} \end{cases}$$

$$\widetilde{T}^*_{\boldsymbol{w}} = \widetilde{T}_{\boldsymbol{w}^{-1}}.$$

We define normalized elements $T_{\boldsymbol{w}} = q^{-|\boldsymbol{w}|/2}\widetilde{T}_{\boldsymbol{w}}$. Then for $\boldsymbol{w} \in W$ and $s \in S$,

$$T_s T_{\boldsymbol{w}} = \begin{cases} T_{s\boldsymbol{w}} & \text{if } |s\boldsymbol{w}| > |\boldsymbol{w}|, \\ T_{s\boldsymbol{w}} + pT_{\boldsymbol{w}} & \text{otherwise,} \end{cases} \tag{2-1}$$

where

$$p = \frac{q-1}{\sqrt{q}}.$$

There is a natural positive linear tracial map τ on $\mathbb{C}_q(W)$ satisfying $\tau(T_{\boldsymbol{w}})=0$, $\boldsymbol{w}\neq 1$, and $\tau(1)=1$. Let $L^2(\mathcal{M}_q)$ be the Hilbert space given by the closure of $\mathbb{C}_q(W)$ with respect to $\langle x, y\rangle=\tau(y^*x)$ and let $\mathcal{M}_q$ be the von Neumann algebra generated by $\mathbb{C}_q(W)$ acting on $L^2(\mathcal{M}_q)$. The map τ extends to a state on $\mathcal{M}_q$ and $L^2(\mathcal{M}_q)$ is its GNS space with cyclic vector $\Omega := T_e$. $\mathcal{M}_q$ is called the *Hecke von Neumann algebra* at parameter q associated to the right-angled Coxeter system (W, S).

Theorem 2.2 (see [Garncarek 2016]). *Let (W, S) be an irreducible right-angled Coxeter system and suppose that $|S|\geq 3$. Let ρ be the radius of convergence of the fundamental power series*

$$\sum_{k=0}^{\infty}|\{\boldsymbol{w}\in W : |\boldsymbol{w}|=k\}|z^k. \tag{2-2}$$

For every $q\in[\rho,\rho^{-1}]$ the von Neumann algebra $\mathcal{M}_q$ is a factor. For $q>0$ not in $[\rho,\rho^{-1}]$ the von Neumann algebra $\mathcal{M}_q$ is the direct sum of a factor and $\mathbb{C}$.

As $\mathcal{M}_q$ possesses a normal faithful tracial state the factors appearing in Theorem 2.2 are of type II_1.

For the analysis of $\mathcal{M}_q$ we shall in fact need $\mathcal{M}_1$, which is the group von Neumann algebra of the Coxeter group W. It can be represented on $L^2(\mathcal{M}_q)$. Indeed, let $T_{\boldsymbol{w}}^{(1)}$ denote the generators of $\mathcal{M}_1$ as in (2-1) and let $T_{\boldsymbol{w}}$ be the generators of $\mathcal{M}_q$. Define the unitary map[1]

$$U : L^2(\mathcal{M}_1)\to L^2(\mathcal{M}_q) : T_{\boldsymbol{w}}^{(1)}\Omega\to T_{\boldsymbol{w}}\Omega.$$

In this paper we shall always assume that $\mathcal{M}_1$ is represented on $L^2(\mathcal{M}_q)$ by the identification $\mathcal{M}_1\to\mathcal{B}(L^2(\mathcal{M}_q)) : x\mapsto UxU^*$. Note that this way

$$T_{\boldsymbol{v}}^{(1)}(T_{\boldsymbol{w}}\Omega)=T_{\boldsymbol{vw}}\Omega. \tag{2-3}$$

For $\boldsymbol{w}\in W$ we shall write $P_{\boldsymbol{w}}$ for the projection of $L^2(\mathcal{M}_q)$ onto the closure of the space spanned linearly by $\{T_{\boldsymbol{v}}\Omega : |\boldsymbol{w}^{-1}\boldsymbol{v}|=|\boldsymbol{v}|-|\boldsymbol{w}|\}$ (see Remark 2.3 below). For $\Gamma_0\in\mathsf{Cliq}(\Gamma)$ we shall write $P_{V\Gamma_0}$ for $P_{\boldsymbol{w}}$, where $\boldsymbol{w}\in W$ is the product of all vertex elements of Γ_0, and $|V\Gamma_0|$ for the number of elements in $V\Gamma_0$. Note that if $s, t\in V\Gamma_0$ then P_s and P_t commute and so $P_{V\Gamma_0}$ is well-defined. Similarly we shall write $P_{\boldsymbol{v}V\Gamma_0}$ for $P_{\boldsymbol{w}}$, where $\boldsymbol{w}\in W$ is the product of $\boldsymbol{v}$ with all vertex elements of Γ_0.

Remark 2.3 (creation and annihilation arguments). Note that for $\boldsymbol{w}, \boldsymbol{v}\in W$ saying that $|\boldsymbol{w}^{-1}\boldsymbol{v}|=|\boldsymbol{v}|-|\boldsymbol{w}|$ just means that the start of $\boldsymbol{v}$ contains the word $\boldsymbol{w}$. Throughout the paper we say that $s\in S$ acts by means of a creation operator on $\boldsymbol{v}\in W$ if $|s\boldsymbol{v}|=|\boldsymbol{v}|+1$. It acts as an annihilation operator if $|s\boldsymbol{v}|=|\boldsymbol{v}|-1$. Note that as W is right-angled we cannot have $|s\boldsymbol{v}|=|\boldsymbol{v}|$. For $\boldsymbol{v}, \boldsymbol{w}\in W$ we may always decompose $\boldsymbol{w}$ as $\boldsymbol{w}=\boldsymbol{w}'\boldsymbol{w}''$ such that $|\boldsymbol{w}|=|\boldsymbol{w}'|+|\boldsymbol{w}''|$, $|\boldsymbol{w}''\boldsymbol{v}|=|\boldsymbol{v}|-|\boldsymbol{w}''|$, and $|\boldsymbol{wv}|=|\boldsymbol{v}|-|\boldsymbol{w}''|+|\boldsymbol{w}'|$. That is, $\boldsymbol{w}$ first acts by means of annihilations of the letters of $\boldsymbol{w}''$ and then $\boldsymbol{w}'$ acts as a creation operator on $\boldsymbol{w}''\boldsymbol{v}$. We will use such arguments without further reference.

[1]Unitarity follows as the vectors $T_{\boldsymbol{w}}\Omega$ are orthonormal. Indeed $\langle T_{\boldsymbol{w}}\Omega, T_{\boldsymbol{v}}\Omega\rangle=\langle T_{\boldsymbol{v}}^*T_{\boldsymbol{w}}\Omega,\Omega\rangle$. If $\boldsymbol{v}^*\boldsymbol{w}$ is reducible this expression is 0. Otherwise there exists a letter w_1 at the starts of $\boldsymbol{v}$ and $\boldsymbol{w}$ such that $T_{\boldsymbol{v}}^*T_{\boldsymbol{w}}=T_{\boldsymbol{v}'}^*T_{\boldsymbol{w}'}+pT_{\boldsymbol{v}'}^*T_{w_1}T_{\boldsymbol{w}'}$, where $w_1\boldsymbol{w}'=\boldsymbol{w}$ and $w_1\boldsymbol{v}'=\boldsymbol{v}$ and $\boldsymbol{w}'$ and $\boldsymbol{v}'$ are of shorter length. The term $pT_{\boldsymbol{v}'}^*T_{w_1}T_{\boldsymbol{w}'}$ reduces further and can be written as a sum of operators $\sum_i T_{\boldsymbol{u}_i}$ but each $\boldsymbol{u}_i$ must contain the letter w_1 as else $\boldsymbol{w}$ and $\boldsymbol{v}$ would not be reducible. Therefore $\langle pT_{\boldsymbol{v}'}^*T_{w_1}T_{\boldsymbol{w}'}\Omega,\Omega\rangle=0$. So $\langle T_{\boldsymbol{v}}^*T_{\boldsymbol{w}}\Omega,\Omega\rangle=\langle T_{\boldsymbol{v}'}^*T_{\boldsymbol{w}'}\Omega,\Omega\rangle$. Continuing inductively we get $\langle T_{\boldsymbol{v}}^*T_{\boldsymbol{w}}\Omega,\Omega\rangle=\delta_{\boldsymbol{v},\boldsymbol{w}}$.

The following Lemma 2.5, together with Lemma 2.7, says that $T_{\boldsymbol{w}}$ decomposes in terms of a sum of operators that first act by annihilation (this is $T^{(1)}_{\boldsymbol{u}''}$) then a diagonal action (this is the projection $P_{\boldsymbol{u}V\Gamma_0}$) and finally by creation (this is $T^{(1)}_{\boldsymbol{u}'}$).

Definition 2.4. Let $\boldsymbol{w} \in W$. Let $A_{\boldsymbol{w}}$ be the set of triples $(\boldsymbol{w}', \Gamma_0, \boldsymbol{w}'')$ with $\boldsymbol{w}', \boldsymbol{w}'' \in W$ and $\Gamma_0 \in \mathsf{Cliq}(\Gamma)$ such that (1) $\boldsymbol{w} = \boldsymbol{w}' V\Gamma_0 \boldsymbol{w}''$, (2) $|\boldsymbol{w}| = |\boldsymbol{w}'| + |V\Gamma_0| + |\boldsymbol{w}''|$, (3) if $s \in S$ commutes with $V\Gamma_0$ then $|\boldsymbol{w}' s| > |\boldsymbol{w}'|$ (that is, letters commuting with $V\Gamma_0$ cannot occur at the end of $\boldsymbol{w}'$, and if they are present they should occur at the start of $\boldsymbol{w}''$).

Lemma 2.5. *For $(\boldsymbol{w}', \Gamma_0, \boldsymbol{w}'') \in A_{\boldsymbol{w}}$ there exist $\boldsymbol{u}, \boldsymbol{u}', \boldsymbol{u}'' \in W$ such that*

$$T^{(1)}_{\boldsymbol{w}'} P_{V\Gamma_0} T^{(1)}_{\boldsymbol{w}''} = T^{(1)}_{\boldsymbol{u}'} P_{\boldsymbol{u}V\Gamma_0} T^{(1)}_{\boldsymbol{u}''}, \tag{2-4}$$

and moreover if $s \in S$ is such that $|\boldsymbol{u}' s| < |\boldsymbol{u}'|$ then $|s\boldsymbol{u}''| > |\boldsymbol{u}''|$. We may assume that $\boldsymbol{u}' = \boldsymbol{w}'\boldsymbol{u}^{-1}$ and $\boldsymbol{u}'' = \boldsymbol{u}\boldsymbol{w}''$.

Proof. Let $\boldsymbol{u} \in W$ be the (unique) element of maximal length such that $|\boldsymbol{w}'\boldsymbol{u}^{-1}| = |\boldsymbol{w}'| - |\boldsymbol{u}|$ and $|\boldsymbol{u}\boldsymbol{w}''| = |\boldsymbol{w}''| - |\boldsymbol{u}|$. Set $\boldsymbol{u}' = \boldsymbol{w}'\boldsymbol{u}^{-1}$ and $\boldsymbol{u}'' = \boldsymbol{u}\boldsymbol{w}''$. It then remains to prove (2-4) as the rest of the properties are obvious or follow by maximality of $\boldsymbol{u}$. We must show that

$$P_{\boldsymbol{u}V\Gamma_0} = T^{(1)}_{\boldsymbol{u}} P_{V\Gamma_0} T^{(1)}_{\boldsymbol{u}^{-1}}.$$

Take $\boldsymbol{v} \in W$. If $|(\boldsymbol{u}V\Gamma_0)^{-1}\boldsymbol{v}| = |\boldsymbol{v}| - |\boldsymbol{u}V\Gamma_0|$ (i.e., $\boldsymbol{v}$ starts with $\boldsymbol{u}V\Gamma_0$) then $|(V\Gamma_0)^{-1}\boldsymbol{u}^{-1}\boldsymbol{v}| = |\boldsymbol{u}^{-1}\boldsymbol{v}| - |V\Gamma_0|$ (i.e., $\boldsymbol{u}^{-1}\boldsymbol{v}$ starts with $V\Gamma_0$). We shall prove that the converse holds. First, we claim that if $|(V\Gamma_0)^{-1}\boldsymbol{u}^{-1}\boldsymbol{v}| = |\boldsymbol{u}^{-1}\boldsymbol{v}| - |V\Gamma_0|$ then $|\boldsymbol{u}^{-1}\boldsymbol{v}| = |\boldsymbol{v}| - |\boldsymbol{u}^{-1}|$ (i.e., $\boldsymbol{v}$ starts with $\boldsymbol{u}$). Indeed, because if this were not the case then one of the letters in $\boldsymbol{u}$ would remain at the start of $\boldsymbol{u}^{-1}\boldsymbol{v}$. And as the letters of $\boldsymbol{u}$ do not commute with $V\Gamma_0$ this would mean that $|(V\Gamma_0)^{-1}\boldsymbol{u}^{-1}\boldsymbol{v}| \neq |\boldsymbol{u}^{-1}\boldsymbol{v}| - |V\Gamma_0|$, which is a contradiction. From the initial assumption $|(V\Gamma_0)^{-1}\boldsymbol{u}^{-1}\boldsymbol{v}| = |\boldsymbol{u}^{-1}\boldsymbol{v}| - |V\Gamma_0|$ ($\boldsymbol{u}^{-1}\boldsymbol{v}$ starts with $V\Gamma_0$) together with $|\boldsymbol{u}^{-1}\boldsymbol{v}| = |\boldsymbol{v}| - |\boldsymbol{u}^{-1}|$ ($\boldsymbol{v}$ starts with $\boldsymbol{u}$) we get that $|(\boldsymbol{u}V\Gamma_0)^{-1}\boldsymbol{v}| = |\boldsymbol{v}| - |\boldsymbol{u}V\Gamma_0|$.

The previous paragraph shows the first equality of

$$P_{\boldsymbol{u}V\Gamma_0}(T_{\boldsymbol{v}}\Omega) = T^{(1)}_{\boldsymbol{u}} P_{V\Gamma_0}(T_{\boldsymbol{u}^{-1}\boldsymbol{v}}\Omega) = T^{(1)}_{\boldsymbol{u}} P_{V\Gamma_0} T^{(1)}_{\boldsymbol{u}^{-1}}(T_{\boldsymbol{v}}\Omega). \qquad \square$$

Remark 2.6. In Lemma 2.5 the property that $|\boldsymbol{u}' s| < |\boldsymbol{u}'|$ implies that $|s\boldsymbol{u}''| > |\boldsymbol{u}''|$ is equivalent to $|\boldsymbol{u}'\boldsymbol{u}''| = |\boldsymbol{u}'| + |\boldsymbol{u}''|$. The words $\boldsymbol{u}'$ and $\boldsymbol{u}''$ in Lemma 2.5 are not unique: in the case $|s\boldsymbol{u}''| = |\boldsymbol{u}''| - 1$ and s commutes with $V\Gamma_0$, we may replace $(\boldsymbol{u}', \boldsymbol{u}'')$ by $(\boldsymbol{u}'s, s\boldsymbol{u}'')$.

Lemma 2.7. *We have*

$$T_{\boldsymbol{w}} = \sum_{(\boldsymbol{w}', \Gamma_0, \boldsymbol{w}'') \in A_{\boldsymbol{w}}} p^{|V\Gamma_0|} T^{(1)}_{\boldsymbol{w}'} P_{V\Gamma_0} T^{(1)}_{\boldsymbol{w}''}, \tag{2-5}$$

where $A_{\boldsymbol{w}}$ is given in Definition 2.4.

Proof. The proof proceeds by induction on the length of $\boldsymbol{w}$. If $|\boldsymbol{w}| = 1$ then $T_{\boldsymbol{w}} = T^{(1)}_{\boldsymbol{w}} + pP_{\boldsymbol{w}}$ by (2-1). Now suppose that (2-5) holds for all $\boldsymbol{w} \in W$ with $|\boldsymbol{w}| = n$. Let $\boldsymbol{v} \in W$ be such that $|\boldsymbol{v}| = n+1$. Decompose

$\boldsymbol{v}$ as $\boldsymbol{v} = s\boldsymbol{w},\ |\boldsymbol{w}| = n,\ s \in S$. Then,

$$T_{\boldsymbol{v}} = T_s T_{\boldsymbol{w}} = (T_s^{(1)} + pP_s)\Bigg(\sum_{(\boldsymbol{w}',\Gamma_0,\boldsymbol{w}'')\in A_{\boldsymbol{w}}} p^{|V\Gamma_0|} T_{\boldsymbol{w}'}^{(1)} P_{V\Gamma_0} T_{\boldsymbol{w}''}^{(1)}\Bigg)$$
$$= \sum_{(\boldsymbol{w}',\Gamma_0,\boldsymbol{w}'')\in A_{\boldsymbol{w}}} (p^{|V\Gamma_0|} T_{s\boldsymbol{w}'}^{(1)} P_{V\Gamma_0} T_{\boldsymbol{w}''}^{(1)} + p^{|V\Gamma_0|+1} P_s T_{\boldsymbol{w}'}^{(1)} P_{V\Gamma_0} T_{\boldsymbol{w}''}^{(1)}). \tag{2-6}$$

Now we need to make the following observations.

(1) If $s\boldsymbol{w}' = \boldsymbol{w}'s$ then $P_s T_{\boldsymbol{w}'}^{(1)} = T_{\boldsymbol{w}'}^{(1)} P_s$. So in that case,

$$P_s T_{\boldsymbol{w}'}^{(1)} P_{V\Gamma_0} T_{\boldsymbol{w}''}^{(1)} = T_{\boldsymbol{w}'}^{(1)} P_s P_{V\Gamma_0} T_{\boldsymbol{w}''}^{(1)}.$$

Moreover $P_s P_{V\Gamma_0}$ equals $P_{sV\Gamma_0}$ in the case s commutes with all elements of $V\Gamma_0$ and it equals 0 otherwise.

(2) In the case $s\boldsymbol{w}' \neq \boldsymbol{w}'s$ we claim that $P_s T_{\boldsymbol{w}'}^{(1)} P_{V\Gamma_0} T_{\boldsymbol{w}''}^{(1)} = 0$. To see this, rewrite $P_s T_{\boldsymbol{w}'}^{(1)} P_{V\Gamma_0} T_{\boldsymbol{w}''}^{(1)} = P_s T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}V\Gamma_0} T_{\boldsymbol{u}''}^{(1)}$ with $\boldsymbol{u}, \boldsymbol{u}', \boldsymbol{u}''$ as in Lemma 2.5. As $s\boldsymbol{w}' \neq \boldsymbol{w}'s$ we have $s\boldsymbol{u}' \neq \boldsymbol{u}'s$ and/or $s\boldsymbol{u} \neq \boldsymbol{u}s$ (because $\boldsymbol{w}' = \boldsymbol{u}'\boldsymbol{u}$ with $|\boldsymbol{w}'| = |\boldsymbol{u}'| + |\boldsymbol{u}|$; see Lemma 2.5).

(a) Assume $s\boldsymbol{u}' \neq \boldsymbol{u}'s$. For $\boldsymbol{v} \in W$ with $T_{\boldsymbol{u}''\boldsymbol{v}}\Omega$ in the range of $P_{\boldsymbol{u}V\Gamma_0}$,

$$P_s T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}V\Gamma_0} T_{\boldsymbol{u}''}^{(1)} (T_{\boldsymbol{v}}\Omega) = P_s T_{\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{v}}\Omega. \tag{2-7}$$

Furthermore, the assertions of Lemma 2.5 imply $|\boldsymbol{u}'\boldsymbol{u}V\Gamma_0| = |\boldsymbol{u}'| + |\boldsymbol{u}V\Gamma_0|$ and therefore (recalling that $T_{\boldsymbol{u}''\boldsymbol{v}}\Omega$ is in the range of $P_{\boldsymbol{u}V\Gamma_0}$) we get that $|\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{v}| = |\boldsymbol{u}''\boldsymbol{v}| + |\boldsymbol{u}'|$, which implies (because $s\boldsymbol{u}' \neq \boldsymbol{u}'s$ and $\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{v}$ starts with all letters of $\boldsymbol{u}'$) that (2-7) is 0. For $\boldsymbol{v} \in W$ with $T_{\boldsymbol{u}''\boldsymbol{v}}\Omega$ not in the range of $P_{\boldsymbol{u}V\Gamma_0}$ we have $T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}V\Gamma_0} T_{\boldsymbol{u}''}^{(1)} (T_{\boldsymbol{v}}\Omega) = 0$. In all we conclude $P_s T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}V\Gamma_0} T_{\boldsymbol{u}''}^{(1)} = 0$.

(b) Assume $s\boldsymbol{u}' = \boldsymbol{u}'s$ but $s\boldsymbol{u} \neq \boldsymbol{u}s$. Then $P_s T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}} = T_{\boldsymbol{u}'}^{(1)} P_s P_{\boldsymbol{u}} = 0$.

So in all (2-6) gives,

$$T_{\boldsymbol{v}} = \sum_{(\boldsymbol{w}',\Gamma_0,\boldsymbol{w}'')\in A_{\boldsymbol{w}}} p^{|V\Gamma_0|} T_{s\boldsymbol{w}'}^{(1)} P_{V\Gamma_0} T_{\boldsymbol{w}''}^{(1)} + \sum_{(\boldsymbol{w}',\Gamma_0,\boldsymbol{w}'')\in A_{\boldsymbol{w}},\, s\boldsymbol{w}'=\boldsymbol{w}'s,\, sV\Gamma_0=V\Gamma_0 s} p^{|V\Gamma_0|+1} T_{\boldsymbol{w}'}^{(1)} P_{sV\Gamma_0} T_{\boldsymbol{w}''}^{(1)},$$

and in turn an identification of all summands shows that the latter expression equals

$$\sum_{(\boldsymbol{v}',\Gamma_0,\boldsymbol{v}'')\in A_{s\boldsymbol{w}}} p^{|V\Gamma_0|} T_{\boldsymbol{v}'}^{(1)} P_{V\Gamma_0} T_{\boldsymbol{v}''}^{(1)}. \qquad \square$$

2C. *Group von Neumann algebras.* Let G be a discrete group with left regular representation $s \mapsto \lambda_s$ and group von Neumann algebra $\mathcal{L}(\mathsf{G}) = \{\lambda_s : s \in \mathsf{G}\}''$. We let $A(\mathsf{G})$ be the Fourier algebra consisting of functions $\varphi(s) = \langle \lambda_s \xi,\ \eta\rangle, \xi, \eta \in \ell^2(\mathsf{G})$. There is a pairing between $A(\mathsf{G})$ and $\mathcal{L}(\mathsf{G})$ which is given by $\langle \varphi, \lambda(f)\rangle = \int_{\mathsf{G}} f(s)\varphi(s)\, ds$ which turns $A(\mathsf{G})$ into an operator space that is completely isometrically identified with $\mathcal{L}(\mathsf{G})_*$. We let $M_{\mathcal{CB}}A(\mathsf{G})$ be the space of completely bounded Fourier multipliers of $A(\mathsf{G})$. For $m \in M_{\mathcal{CB}}A(\mathsf{G})$ we let $T_m : \mathcal{L}(\mathsf{G}) \to \mathcal{L}(\mathsf{G})$ be the normal completely bounded map determined by $\lambda(f) \mapsto \lambda(mf)$. The following theorem is due to Bożejko and Fendler [1984] (see also [Junge et al. 2009, Theorem 4.5]).

Theorem 2.8. *Let $m \in M_{\mathcal{CB}}A(\mathsf{G})$. There exists a unique normal completely bounded map $M_m : \mathcal{B}(\ell^2(\mathsf{G})) \to \mathcal{B}(\ell^2(\mathsf{G}))$ that is an $L^\infty(\mathsf{G})$-bimodule homomorphism and such that M_m restricts to $T_m : \lambda(f) \mapsto \lambda(mf)$ on $\mathcal{L}(\mathsf{G})$. Moreover, $\|M_m\|_{\mathcal{CB}} = \|T_m\|_{\mathcal{CB}} = \|m\|_{M_{\mathcal{CB}}A(\mathsf{G})}$.*

The map M_m is called the Herz–Schur multiplier.

3. Universal property and conditional expectations

In this section we establish universal properties for $\mathcal{M}_q$ and consequently show that $\mathcal{M}_q$ is noninjective and has the Haagerup property.

3A. *Universal properties.*

Theorem 3.1. *Let $q > 0$, put $p = (q-1)/\sqrt{q}$ and let (W, S) be a right-angled Coxeter system with associated Hecke von Neumann algebra $(\mathcal{M}_q, \tau)$. Suppose that $(\mathcal{N}, \tau_{\mathcal{N}})$ is a von Neumann algebra with GNS faithful state $\tau_{\mathcal{N}}$ that is generated by self-adjoint operators R_s, $s \in S$, that satisfy the relations $R_s R_t = R_t R_s$ whenever $m(s,t) = 2$, $R_s^2 = 1 + pR_s$, $s \in S$, and further $\tau_{\mathcal{N}}(R_{w_1} \cdots R_{w_n}) = 0$ for every nonempty reduced word $\boldsymbol{w} = w_1 \cdots w_n \in W$. Then there exists a unique normal $*$-homomorphism $\pi : \mathcal{M}_q \to \mathcal{N}$ such that $\pi(T_s) = R_s$. Moreover $\tau_{\mathcal{N}} \circ \pi = \tau$.*

Proof. The proof is routine; see [Caspers and Fima 2017, Proposition 2.12]. We sketch it here. Let $(L^2(\mathcal{N}), \pi_{\mathcal{N}}, \eta)$ be a GNS construction for $(\mathcal{N}, \tau_{\mathcal{N}})$. As $\tau_{\mathcal{N}}$ is GNS faithful we may assume that $\mathcal{N}$ is represented on $L^2(\mathcal{N})$ via $\pi_{\mathcal{N}}$. We define a linear map $V : L^2(\mathcal{M}_q) \to L^2(\mathcal{N})$ by $V\Omega = \eta$ and

$$V(T_{\boldsymbol{w}}\Omega) = R_{\boldsymbol{w}}\eta, \quad \text{where } \boldsymbol{w} \in W,$$

and $R_{\boldsymbol{w}} := R_{w_1} \cdots R_{w_n}$. One checks that V is isometric by showing that $\{R_{\boldsymbol{w}}\eta : \boldsymbol{w} \in W\}$ is an orthonormal system.[2] Putting $\pi(\,\cdot\,) = V(\,\cdot\,)V^*$ concludes the lemma. As $V\Omega = \eta$ we get $\tau_{\mathcal{N}} \circ \pi = \tau$. □

Remark 3.2. Note that the property $T_s^2 = 1 + pT_s$, $s \in S$, with $p = (q-1)/\sqrt{q}$, is equivalent to the usual Hecke relation $(\sqrt{q}\,T_s - q)(\sqrt{q}\,T_s + 1) = 0$ that appears in the literature.

We shall say that $(\widetilde{W}, \widetilde{S})$ is a Coxeter subsystem of (W, S) if $\widetilde{S} \subseteq S$ and $\tilde{m}(s,t) = m(s,t)$ for all $s, t \in \widetilde{S}$. Here $\tilde{m}$ is the function on $\widetilde{S} \times \widetilde{S}$ that determines the commutation relations for $\widetilde{W}$; see Section 2A.

Corollary 3.3. *Let $q > 0$. Let $(\widetilde{W}, \widetilde{S})$ be a Coxeter subsystem of a right-angled Coxeter system (W, S). Let $\widetilde{\mathcal{M}}_q$ and $\mathcal{M}_q$ be their respective Hecke von Neumann algebras. Then naturally $\widetilde{\mathcal{M}}_q$ is a von Neumann subalgebra of $\mathcal{M}_q$. In particular, there exists a trace-preserving normal conditional expectation $\mathcal{E} : \mathcal{M}_q \to \widetilde{\mathcal{M}}_q$.*

[2] The proof goes as follows. We may find unique coefficients $c_{\boldsymbol{v}}$ such that $T_{w'_n} \cdots T_{w'_1} T_{w_1} \cdots T_{w_n} = \sum_{\boldsymbol{v} \in W} c_{\boldsymbol{v}} T_{\boldsymbol{v}}$. We have $c_\varnothing = 1$ if $\boldsymbol{w} = \boldsymbol{w}'$ and $c_\varnothing = 0$ if $\boldsymbol{w} \neq \boldsymbol{w}'$ by comparing the trace of both sides of this expression. In fact the coefficients $c_{\boldsymbol{v}}$ may be found by using the commutation relations for T_s and the Hecke relation $T_s^2 = 1 + pT_s$ to "reduce" the left-hand side of this expression. As the same relations hold for the operators R_s (by the assumption of the lemma) we also get $R_{w'_n} \cdots R_{w'_1} R_{w_1} \cdots R_{w_n} = \sum_{\boldsymbol{v} \in W} c_{\boldsymbol{v}} R_{\boldsymbol{v}}$. So, $\langle R_{\boldsymbol{w}}\eta, R_{\boldsymbol{w}'}\eta \rangle = \tau_{\mathcal{N}}(R_{\boldsymbol{w}'}^* R_{\boldsymbol{w}}) = \tau_{\mathcal{N}}(R_{w'_n} \cdots R_{w'_1} R_{w_1} \cdots R_{w_n}) = \tau_{\mathcal{N}}\big(\sum_{\boldsymbol{v} \in W} c_{\boldsymbol{v}} R_{\boldsymbol{v}}\big) = c_\varnothing$. This proves that indeed V is isometric.

Proof. Theorem 3.1 implies that $\widetilde{\mathcal{M}}_q$ is a von Neumann subalgebra of $\mathcal{M}_q$ and the canonical trace of $\mathcal{M}_q$ agrees with the one on $\widetilde{\mathcal{M}}_q$. Therefore $\widetilde{\mathcal{M}}_q$ admits a trace-preserving normal conditional expectation value; see [Takesaki 2003, Theorem IX.4.2]. □

Consider the Hecke von Neumann algebra $\mathcal{M}_q$ for the case that S is a one-point set, $q > 0$, and $p = (q-1)/\sqrt{q}$. In that case we have $W = \{e, s\}$ and $L^2(\mathcal{M}_q)$ has a canonical basis Ω and $T_s\Omega$. With respect to this basis T_s takes the form $\left(\begin{smallmatrix} 0 & 1 \\ 1 & p \end{smallmatrix}\right)$ and one sees (using for example the relation $T_s^2 = 1 + pT_s$) that $\mathcal{M}_q = \mathbb{C}\mathrm{Id}_2 \oplus \mathbb{C}T_s$; i.e., it is 2-dimensional. The following corollary uses the graph product, for which we refer to [Caspers and Fima 2017]. It is a generalization of the free product by adding a commutation relation to vertex algebras that share an edge; the free product is then given by a graph product over a graph with no edges. In [Caspers and Fima 2017] the symbol $*$ was used for graph products. We use the notation $\star$ instead to distinguish them from free (amalgamated) products.

Corollary 3.4. *Let (W, S) be an arbitrary right-angled Coxeter system and let $q > 0$. Let Γ be the graph associated to (W, S) as before. For $s \in S$ let $\mathcal{M}_q(s)$ be the 2-dimensional Hecke von Neumann subalgebra corresponding to the one-point set $\{s\}$. Then we have a graph product decomposition $\mathcal{M}_q = \star_{s \in V\Gamma} \mathcal{M}_q(s)$.*

Proof. Let $T_s \in \mathcal{M}_q,\ s \in S$, be the operators as introduced in Section 2B. Let $\widetilde{T}_s,\ s \in S$, be the operator T_s but then considered in the algebra $\mathcal{M}_q(s)$, which in turn is contained in $\star_{s \in V\Gamma} \mathcal{M}_q(s)$ with conditional expectation. Now the map $T_s \mapsto \widetilde{T}_s$ determines an isomorphism by Theorem 3.1 and the universal property of the graph product given by [Caspers and Fima 2017, Proposition 2.12]. □

3B. *Noninjectivity.*

Definition 3.5. A von Neumann algebra $\mathcal{M} \subseteq \mathcal{B}(\mathcal{H})$ is called injective if there exists a conditional expectation $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{M}$.

Theorem 3.6. *Let (W, S) be an irreducible right-angled Coxeter system with $|S| \geq 3$. Then $\mathcal{M}_q$ is noninjective.*

Proof. It suffices to prove that $\mathcal{M}_q$ contains an expected noninjective von Neumann subalgebra. Now any irreducible Coxeter system (W, S) contains a Coxeter subsystem $(\widetilde{W}, \widetilde{S})$ either of the form $\widetilde{S} = \{r, s, t\}$ with $\tilde{m}(r, s) = \tilde{m}(r, t) = \tilde{m}(s, t) = \infty$ or $\widetilde{S} = \{r, s, t\}$ with $\tilde{m}(r, s) = \tilde{m}(r, t) = \infty$ and $\tilde{m}(s, t) = 2$. So it satisfies to prove noninjectivity for these systems. In both cases, for q fixed, set $\mathcal{M}$ to be the Hecke von Neumann algebra of the Coxeter system consisting of just $\{r\}$. $\mathcal{M}$ has dimension 2. Set $\mathcal{N}$ to be the Hecke von Neumann algebra of the Coxeter system $\{s, t\}$, which is infinite-dimensional in the case $m(s, t) = \infty$ and 4-dimensional if $m(s, t) = 2$ (being the tensor product of two 2-dimensional algebras). Then $\mathcal{M}_q$ is isomorphic to the free product $\mathcal{M} * \mathcal{N}$ over the canonical traces by Corollary 3.4 and [Caspers and Fima 2017, Remark 3.23]. As $\dim(\mathcal{M}) + \dim(\mathcal{N}) \geq 5$ it follows that $\mathcal{M}_q$ is noninjective from [Ueda 2011, Theorem 4.1] (see comment (5) in Remark 4.2 of that work). □

3C. *Haagerup property.*

We first construct radial multipliers.

Proposition 3.7. *Let (W, S) be a right-angled Coxeter group with Hecke von Neumann algebra $\mathcal{M}_q$, $q > 0$. For every $0 < r < 1$ there exists a normal unital completely positive map $\Phi_r : \mathcal{M}_q \to \mathcal{M}_q$ that is determined by $\Phi_r(T_{\mathbf{w}}) = r^{|\mathbf{w}|} T_{\mathbf{w}}$.*

Proof. As in Corollary 3.4 we identify $\mathcal{M}_q$ with the graph product $*_{s\in V\Gamma}^{\Gamma}(\mathcal{M}_q(s),\tau_s)$, where τ_s is the tracial state on $\mathcal{M}_q(s)$. Consider the map $\Phi_{r,s}:\mathcal{M}_q(s)\to\mathcal{M}_q(s)$ determined by $1\mapsto 1,\ T_s\mapsto rT_s$. This map is unital and completely positive: indeed consider the matrices

$$A:=\begin{pmatrix}\sqrt{1-r} & 0\\ 0 & 0\end{pmatrix},\quad B=\begin{pmatrix}0 & \sqrt{1-r}\\ 0 & 0\end{pmatrix},\quad C=\begin{pmatrix}\sqrt{r} & 0\\ 0 & \sqrt{r}\end{pmatrix}.$$

Then $\Phi_{r,s}$ agrees with $x\mapsto A^*xA+B^*xB+C^*xC$ as before Corollary 3.4 we already noted that $T_s=\left(\begin{smallmatrix}0 & 1\\ 1 & p\end{smallmatrix}\right)$. Furthermore $\Phi_{r,s}$ preserves the trace τ_s as τ_s is the vector state associated with $(1,0)^t$. Therefore we may apply [Caspers and Fima 2017, Proposition 2.30] and obtain the graph product ucp map $\Phi_r:=\star_{s\in V\Gamma}\Phi_{r,s}$, which proves the proposition. ☐

Definition 3.8. Recall that a von Neumann algebra $\mathcal{M}$ with normal faithful tracial state τ has the *Haagerup property* if there exists a net Φ_i of τ-preserving ucp maps $\mathcal{M}\to\mathcal{M}$ such that $T_i: x\Omega_\tau\mapsto\Phi_i(x)\Omega_\tau$ is compact and converges to 1 strongly.

Theorem 3.9. *For any Coxeter system (W,S) and any $q>0$ the von Neumann algebra $\mathcal{M}_q$ has the Haagerup property.*

Proof. If S is finite Proposition 3.7 directly shows that $\mathcal{M}_q$ has the Haagerup property by letting $r\nearrow 1$. Then the general case follows by an inductive limit argument on finite Coxeter subsystems using the conditional expectations from Corollary 3.3. ☐

4. Completely contractive approximation property

We show that for a right-angled Coxeter system (W,S) the Hecke von Neumann algebra $\mathcal{M}_q$ has the wk-$*$ CCAP; see Definition 4.12. The proof follows a — by now standard — strategy of [Haagerup 1978] by considering radial multipliers first and then showing that word-length cut-downs have a complete bound that is at most polynomial in the word length.

4A. ***Creation/annihilation arguments.*** Here we present some combinatorial arguments that we need in Section 4B. We have chosen to separate these from the proofs of Section 4B so that the reader could skip them at first sight.

We introduce the following notation. Let $\boldsymbol{x},\boldsymbol{w}\in W$. We shall write $\boldsymbol{w}\le\boldsymbol{x}$ to indicate that $|\boldsymbol{w}^{-1}\boldsymbol{x}|=|\boldsymbol{x}|-|\boldsymbol{w}|$. Then $\boldsymbol{w}<\boldsymbol{x}$ is defined naturally. So $\boldsymbol{w}\le\boldsymbol{x}$ means that $\boldsymbol{w}$ is obtained from $\boldsymbol{x}$ by cutting off a tail. An element $\boldsymbol{v}\in W$ is called a *clique word* in the case its letters form a clique. For Λ a clique in W and $\boldsymbol{v}\in W$ we define $\boldsymbol{v}(2,\varnothing)$ as the maximal[3] clique Γ_0 such that $|\boldsymbol{v}V\Gamma_0|=|\boldsymbol{v}|-|V\Gamma_0|$. Then we take the decomposition $\boldsymbol{v}=\boldsymbol{v}(1,\Lambda)\boldsymbol{v}(2,\Lambda)$ with $|\boldsymbol{v}|=|\boldsymbol{v}(1,\Lambda)|+|\boldsymbol{v}(2,\Lambda)|$ and $\boldsymbol{v}(2,\Lambda)=\boldsymbol{v}(2,\varnothing)\backslash\Lambda$ (which uniquely determines $\boldsymbol{v}(1,\Lambda)$). For $\boldsymbol{g}\le\boldsymbol{x}$ we let $\Lambda_{\boldsymbol{g},\boldsymbol{x}}$ be $(\boldsymbol{x}^{-1}\boldsymbol{g})(2,\varnothing)$. In other words $\Lambda_{\boldsymbol{g},\boldsymbol{x}}$ is the maximal clique that appears at the start of $\boldsymbol{g}^{-1}\boldsymbol{x}$. We let $C(\boldsymbol{g},\boldsymbol{x})$ be the collection of $\boldsymbol{w}\in W$ with $\boldsymbol{g}\le\boldsymbol{w}\le\boldsymbol{g}\Lambda_{\boldsymbol{g},\boldsymbol{x}}$. Note that $C(\boldsymbol{g},\boldsymbol{x})$ contains at least $\boldsymbol{g}$ and $\boldsymbol{g}\Lambda_{\boldsymbol{g},\boldsymbol{x}}$ (and the latter elements can be equal). We write $C(\boldsymbol{g},+)$ for $\bigcup_{\boldsymbol{g}\le\boldsymbol{x}}C(\boldsymbol{g},\boldsymbol{x})$.

[3] Suppose that Γ_0 and Γ_1 are cliques such that $|\boldsymbol{v}V\Gamma_i|=|\boldsymbol{v}|-|V\Gamma_i|$, $i=0,1$. Then the letters $V\Gamma_0$ and $V\Gamma_1$ must commute. So the union $\Gamma_2=\Gamma_0\cup\Gamma_1$ is a clique with $|\boldsymbol{v}V\Gamma_2|=|\boldsymbol{v}|-|V\Gamma_2|$.

Example 4.1. Consider the Coxeter system (W, S) with $S = \{r, s, t\}$ in which $m(r, s) = 2$ and $m(r, t) = m(s, t) = \infty$. Consider $\mathbf{v} = trs$. Then $\mathbf{v}(1, \varnothing) = t$, $\mathbf{v}(2, \varnothing) = rs$, $\mathbf{v}(1, r) = tr$, and $\mathbf{v}(2, r) = s$. Also $\Lambda_{t,trst} = \{t, tr, ts, trs\}$.

Lemma 4.2. *Let $\mathbf{x}, \mathbf{w} \in W$. Let $\mathbf{w} = \mathbf{w}'\mathbf{w}''$ be the decomposition with $|\mathbf{w}| = |\mathbf{w}'| + |\mathbf{w}''|$ such that $|\mathbf{w}''\mathbf{x}| = |\mathbf{x}| - |\mathbf{w}''|$ and $|\mathbf{w}\mathbf{x}| = |\mathbf{x}| - |\mathbf{w}''| + |\mathbf{w}'|$. Take $(\mathbf{w}'')^{-1} \leq \mathbf{g} \leq \mathbf{x}$. Then, for $\mathbf{v} \in C(\mathbf{g}, \mathbf{x})$,*

$$(\mathbf{w}\mathbf{v})(2, (\mathbf{w}\mathbf{g})(2, \varnothing)\backslash \mathbf{g}(2, \varnothing)) = \mathbf{v}(2, \mathbf{g}(2, \varnothing)\backslash(\mathbf{w}\mathbf{g})(2, \varnothing)) \tag{4-1}$$

and

$$|(\mathbf{w}\mathbf{v})(1, (\mathbf{w}\mathbf{g})(2, \varnothing)\backslash \mathbf{g}(2, \varnothing))| = |\mathbf{v}(1, \mathbf{g}(2, \varnothing)\backslash \mathbf{w}\mathbf{g}(2, \varnothing))| - |\mathbf{w}''| + |\mathbf{w}'|. \tag{4-2}$$

Proof. Let $\mathbf{v} \in C(\mathbf{g}, \mathbf{x})$. The clique $\mathbf{v}(2, \varnothing)$ consists of the clique $\mathbf{g}^{-1}\mathbf{v}$ plus all letters in $\mathbf{g}(2, \varnothing)$ that commute with $\mathbf{g}^{-1}\mathbf{v}$. Therefore $\mathbf{v}(2, \mathbf{g}(2, \varnothing)\backslash(\mathbf{w}\mathbf{g})(2, \varnothing))$ is the clique consisting of $\mathbf{g}^{-1}\mathbf{v}$ plus all letters in $(\mathbf{w}\mathbf{g})(2, \varnothing) \cap \mathbf{g}(2, \varnothing)$ that commute with $\mathbf{g}^{-1}\mathbf{v}$. On the other hand $(\mathbf{w}\mathbf{v})(2, \varnothing)$ consists of the clique $\mathbf{g}^{-1}\mathbf{v}$ together with all letters in $(\mathbf{w}\mathbf{g})(2, \varnothing)$ that commute with $\mathbf{g}^{-1}\mathbf{v}$. Then $(\mathbf{w}\mathbf{v})(2, (\mathbf{w}\mathbf{g})(2, \varnothing)\backslash \mathbf{g}(2, \varnothing))$ equals $\mathbf{g}^{-1}\mathbf{v}$ together with all elements in $(\mathbf{w}\mathbf{g})(2, \varnothing) \cap \mathbf{g}(2, \varnothing)$ that commute with $\mathbf{g}^{-1}\mathbf{v}$. So we conclude (4-1). Therefore,

$$\begin{aligned} |(\mathbf{w}\mathbf{v})(1, (\mathbf{w}\mathbf{g})(2, \varnothing)\backslash \mathbf{g}(2, \varnothing))| &= |\mathbf{w}\mathbf{v}| - |(\mathbf{w}\mathbf{v})(2, (\mathbf{w}\mathbf{g})(2, \varnothing)\backslash \mathbf{g}(2, \varnothing))| \\ &= |\mathbf{v}| - |\mathbf{w}''| + |\mathbf{w}'| - |\mathbf{v}(2, \mathbf{g}(2, \varnothing)\backslash(\mathbf{w}\mathbf{g})(2, \varnothing))| \\ &= |\mathbf{v}(1, \mathbf{g}(2, \varnothing)\backslash \mathbf{w}\mathbf{g}(2, \varnothing))| - |\mathbf{w}''| + |\mathbf{w}'|, \end{aligned} \tag{4-3}$$

completing the proof. □

Lemma 4.3. *Let $\mathbf{x}, \mathbf{w} \in W$ and decompose $\mathbf{w}$ as $\mathbf{w} = \mathbf{w}'\mathbf{w}''$ such that $|\mathbf{w}| = |\mathbf{w}'| + |\mathbf{w}''|$, $|\mathbf{w}''\mathbf{x}| = |\mathbf{x}| - |\mathbf{w}''|$, and $|\mathbf{w}\mathbf{x}| = |\mathbf{x}| - |\mathbf{w}''| + |\mathbf{w}'|$. Let $(\mathbf{w}'')^{-1} \leq \mathbf{g} \leq \mathbf{x}$. Then*:

(1) $\mathbf{g}(2, \varnothing)\backslash(\mathbf{w}\mathbf{g})(2, \varnothing) = \mathbf{g}(2, \varnothing)\backslash(\mathbf{w}''\mathbf{g})(2, \varnothing)$.

(2) *For $\mathbf{v} \in C(\mathbf{g}, \mathbf{x})$ we have*

$$\mathbf{v}(2, \mathbf{v}(2, \varnothing)\backslash(\mathbf{w}''\mathbf{v})(2, \varnothing)) = \mathbf{v}(2, \mathbf{g}(2, \varnothing)\backslash(\mathbf{w}''\mathbf{g})(2, \varnothing)). \tag{4-4}$$

Proof. (1) Because $(\mathbf{w}'')^{-1} \leq \mathbf{g} \leq \mathbf{x}$ we also have $|\mathbf{w}''\mathbf{g}| = |\mathbf{g}| - |\mathbf{w}''|$ and $|\mathbf{w}\mathbf{g}| = |\mathbf{g}| - |\mathbf{w}''| + |\mathbf{w}'|$. So $\mathbf{w}'$ creates letters in $\mathbf{w}''\mathbf{g}$ so that $\mathbf{g}(2, \varnothing)\backslash(\mathbf{w}\mathbf{g})(2, \varnothing) = \mathbf{g}(2, \varnothing)\backslash(\mathbf{w}''\mathbf{g})(2, \varnothing)$.

(2) Let A be the set of letters in $\mathbf{g}(2, \varnothing)$ that commute with $\mathbf{g}^{-1}\mathbf{v}$. The clique $\mathbf{v}(2, \varnothing)$ consists of $\mathbf{g}^{-1}\mathbf{v} \cup A$. This means that $\mathbf{v}(2, \mathbf{v}(2, \varnothing)\backslash(\mathbf{w}''\mathbf{v})(2, \varnothing))$ consists of $\mathbf{g}^{-1}\mathbf{v} \cup A$ intersected with $(\mathbf{w}''\mathbf{v})(2, \varnothing)$. The intersection of $(\mathbf{w}''\mathbf{v})(2, \varnothing)$ with $\mathbf{g}^{-1}\mathbf{v}$ is $\mathbf{g}^{-1}\mathbf{v}$ so that $\mathbf{v}(2, \mathbf{v}(2, \varnothing)\backslash(\mathbf{w}''\mathbf{v})(2, \varnothing)) = \mathbf{g}^{-1}\mathbf{v} \cup (A \cap (\mathbf{w}''\mathbf{v})(2, \varnothing))$. On the other hand $\mathbf{v}(2, \mathbf{g}(2, \varnothing)\backslash(\mathbf{w}''\mathbf{g})(2, \varnothing))$ equals $\mathbf{g}^{-1}\mathbf{v} \cup (A \cap (\mathbf{w}''\mathbf{g})(2, \varnothing))$ and as $\mathbf{g}(2, \varnothing) \cap (\mathbf{w}''\mathbf{g})(2, \varnothing) = \mathbf{g}(2, \varnothing) \cap (\mathbf{w}''\mathbf{v})(2, \varnothing)$ clearly $(A \cap (\mathbf{w}''\mathbf{v})(2, \varnothing)) = (A \cap (\mathbf{w}''\mathbf{g})(2, \varnothing))$. This proves (4-4). □

Although Coxeter groups generally do not have polynomial growth (nor are they hyperbolic) we still have the polynomial estimate of the following Lemma 4.4. We do not believe that the degree of the polynomial bound we obtain in Lemma 4.4 is optimal, but it suffices for our purposes and it admits a short proof.

Lemma 4.4. *Let W be a right-angled Coxeter group with finite graph Γ. Let $\boldsymbol{x} \in W$. For $a \in \mathbb{N}$ define*

$$\kappa_{\boldsymbol{x}}(a) = |\{\boldsymbol{w} \leq \boldsymbol{x} : |\boldsymbol{w}| = a\}|.$$

Then $\kappa_{\boldsymbol{x}}(a) \leq Ca^{|V\Gamma|-2}$. Moreover, the constant C can be taken uniformly in $\boldsymbol{x}$.

Proof. To carry out the proof we shall actually count a more refined number. We write $\Lambda \leq \Gamma$ to indicate that Λ is a complete subgraph of Γ. We say that $\boldsymbol{w}$ is a $(\leq \Lambda)$-word if its letters (in reduced form) are all in $V\Lambda$ (they do not need to exhaust all of $V\Lambda$); we say that $\boldsymbol{w}$ is a Λ-word if its letters are exactly $V\Lambda$. Then define

$$\kappa_{\boldsymbol{x}}^{\Lambda}(a) = |\{\boldsymbol{v} \leq \boldsymbol{x} : |\boldsymbol{v}| = a \text{ and } \boldsymbol{v} \text{ is a } (\leq \Lambda)\text{-word}\}|. \tag{4-5}$$

Let c and k_0 be constants such that for $a \in \{0, 1\}$ we have for all $\varnothing \neq \Lambda < \Gamma$ that $\kappa_{\boldsymbol{x}}^{\Lambda}(a) \leq c(a+k_0)^{|V\Lambda|-2}$ and further for all $a \in \mathbb{N}$ and all nonempty complete subgraphs Λ of Γ we have $2^{|V\Lambda|}ca \leq (a+k_0)^2$. We prove by induction on $a \in \mathbb{N}$ that for all $\varnothing \neq \Lambda < \Gamma$ we have $\kappa_{\boldsymbol{x}}^{\Lambda}(a) \leq c(a+k_0)^{|V\Lambda|-2}$.

Inductive step. Pick some fixed $\boldsymbol{w} < \boldsymbol{x}$ with $|\boldsymbol{w}| = a$ and $\boldsymbol{w}$ a Λ-word. Now if $\boldsymbol{v} < \boldsymbol{x}$ with $|\boldsymbol{v}| = a$ then let $\boldsymbol{v}_0$ be an element of maximal length such that both $\boldsymbol{v}_0 \leq \boldsymbol{v}$ and $\boldsymbol{v}_0 \leq \boldsymbol{w}$ (we leave in the middle if $\boldsymbol{v}_0$ is unique).

Let $s \in S$ be a letter that appears at the start of $\boldsymbol{v}_0^{-1}\boldsymbol{w}$. We claim that the letter s must commute with $\boldsymbol{v}_0^{-1}\boldsymbol{v}$. Indeed, first observe that as $\boldsymbol{v}_0$ has maximal length s cannot appear at the start of $\boldsymbol{v}_0^{-1}\boldsymbol{v}$. Further, write $\boldsymbol{x} = \boldsymbol{v}_0(\boldsymbol{v}_0^{-1}\boldsymbol{v})(\boldsymbol{v}^{-1}\boldsymbol{x})$ and $\boldsymbol{x} = \boldsymbol{v}_0(\boldsymbol{v}_0^{-1}\boldsymbol{w})(\boldsymbol{w}^{-1}\boldsymbol{x})$. So,

$$(\boldsymbol{v}_0^{-1}\boldsymbol{w})(\boldsymbol{w}^{-1}\boldsymbol{x}) = (\boldsymbol{v}_0^{-1}\boldsymbol{v})(\boldsymbol{v}^{-1}\boldsymbol{x}). \tag{4-6}$$

Now s appears at the start of $(\boldsymbol{v}_0^{-1}\boldsymbol{w})$ and hence this letter must occur somewhere in the expression $(\boldsymbol{v}_0^{-1}\boldsymbol{v})(\boldsymbol{v}^{-1}\boldsymbol{x})$ as well. Consider the first occurrence of s in $(\boldsymbol{v}_0^{-1}\boldsymbol{v})(\boldsymbol{v}^{-1}\boldsymbol{x})$. All the letters before it must then commute with s as otherwise the equality (4-6), saying that s is at the start, is violated (see the normal form theorem [Green 1990, Theorem 3.9]). But then s does not occur on $\boldsymbol{v}_0^{-1}\boldsymbol{v}$ as then it is automatically at its start. So the first time s occurs in $(\boldsymbol{v}_0^{-1}\boldsymbol{v})(\boldsymbol{v}^{-1}\boldsymbol{x})$ is in the part $(\boldsymbol{v}^{-1}\boldsymbol{x})$ and so it commutes with all elements in $(\boldsymbol{v}_0^{-1}\boldsymbol{v})$.

So if $\boldsymbol{v}_0^{-1}\boldsymbol{w}$ is a Λ-word then $\boldsymbol{v}_0^{-1}\boldsymbol{v}$ is a $\mathsf{Link}(\Lambda)$-word (recall $\mathsf{Link}(\Lambda) = \bigcap_{s \in V\Lambda} \mathsf{Link}(s)$); in fact it must be a $(\mathsf{Link}(\Lambda) \cap \Lambda)$-word as we only deal with words with letters in Λ. Moreover $\boldsymbol{v}_0^{-1}\boldsymbol{v}$ must appear at the start of $\boldsymbol{w}^{-1}\boldsymbol{x}$. So every word in the set we count in (4-5) is obtained from $\boldsymbol{w}$ by cutting off a tail (this is $\boldsymbol{v}_0^{-1}\boldsymbol{w}$) and then adding a tail of the same size with commuting letters (this is $\boldsymbol{v}_0^{-1}\boldsymbol{v}$). This certainly gives the inequality

$$\kappa_{\boldsymbol{x}}^{\Lambda}(a) \leq \sum_{\Lambda' \leq \Lambda} \sum_{\substack{\boldsymbol{v} \leq \boldsymbol{w} \\ \boldsymbol{v}^{-1}\boldsymbol{w} \text{ is a } \Lambda'\text{-word}}} \kappa_{\boldsymbol{w}^{-1}\boldsymbol{x}}^{\mathsf{Link}(\Lambda')}(|\boldsymbol{v}^{-1}\boldsymbol{w}|).$$

Note that the number of $\boldsymbol{v} \in W$ with $\boldsymbol{v} < \boldsymbol{w}$, $|\boldsymbol{v}| = l$ and $\boldsymbol{v}$ a Λ' -word is smaller than or equal to $\kappa_{\boldsymbol{w}^{-1}}^{\Lambda'}(|\boldsymbol{w}| - l)$. In the case $l = 0$ we have $\kappa_{\boldsymbol{w}^{-1}}^{\Lambda'}(|\boldsymbol{w}| - l) = 1$ (elementary) and in the case $l > 0$ we can apply our induction hypothesis to get $\kappa^{\Lambda'}_{\boldsymbol{w}^{-1}}(|\boldsymbol{w}|-l) \le c(a-l+k_0)^{|V\Lambda'|-2}$. Therefore we get

$$\begin{aligned}\kappa^{\Lambda}_{\boldsymbol{x}}(a) &\le \sum_{\Lambda'<\Lambda} \sum_{\substack{\boldsymbol{v}<\boldsymbol{w}\\ \boldsymbol{v}^{-1}\boldsymbol{w} \text{ is a } \Lambda'\text{-word}}} c(a+k_0)^{|\mathsf{Link}(\Lambda')\cap V\Lambda|-2}\\ &\le \sum_{\Lambda'<\Lambda} \sum_{l=0}^{a} c^2(a-l+k_0)^{|V\Lambda'|-2}(a+k_0)^{|\mathsf{Link}(\Lambda')\cap V\Lambda|-2}\\ &\le \sum_{\Lambda'<\Lambda} \sum_{l=0}^{a} c^2(a+k_0)^{|V\Lambda'|-2}(a+k_0)^{|\mathsf{Link}(\Lambda')\cap V\Lambda|-2}.\end{aligned}$$

Since the intersection of each $V\Lambda'$ and $\mathsf{Link}(\Lambda')\cap V\Lambda$ is empty we find

$$\kappa^{\Lambda}_{\boldsymbol{x}}(a) \le \sum_{\Lambda'<\Lambda} \sum_{l=0}^{a} c^2(a+k_0)^{|V\Lambda|-4} \le 2^{|V\Gamma|}c^2(a+1)(a+k_0)^{|V\Lambda|-4} \le c(a+k_0)^{|V\Lambda|-2}.$$

The last line follows from the choice of c and k_0. □

4B. ***Word-length projections.*** The aim of this section is to prove that $T_{\boldsymbol{w}} \mapsto \delta(|\boldsymbol{w}| \le n)T_{\boldsymbol{w}}$ gives a complete bounded multiplier of $\mathcal{M}_q$ with complete bound growing at most polynomially in n. Firstly we simplify notation a little bit.

Remark 4.5. We may identify $\ell^2(W)$ with basis $\delta_{\boldsymbol{x}}$, $\boldsymbol{x} \in W$, with $L^2(\mathcal{M}_q)$ with basis $T_{\boldsymbol{x}}\Omega$. This way $T^{(1)}_{\boldsymbol{w}}$ acts on $\ell^2(W)$ by means of the left regular representation.

We borrow the following construction from [Ozawa 2008]. We let $B_f(W)$ be the set of finite subsets of W. For $A \in B_f(W)$ we define $\tilde{\xi}^{\pm}_A$ to be the vectors in $\ell^2(B_f(W))$ given by

$$\tilde{\xi}^{+}_A(\omega) = \begin{cases} 1 & \text{if } \omega \subseteq A,\\ 0 & \text{otherwise,}\end{cases} \qquad \tilde{\xi}^{-}_A(\omega) = \begin{cases} (-1)^{|\omega|} & \text{if } \omega \subseteq A,\\ 0 & \text{otherwise.}\end{cases}$$

Using the binomial formula (see Lemma 4 of [Ozawa 2008]), we have $\|\tilde{\xi}^{\pm}_A\|^2 = 2^{|A|}$ and

$$\langle \tilde{\xi}^{+}_A, \tilde{\xi}^{-}_B\rangle = \begin{cases} 0 & A\cap B \ne \varnothing,\\ 1 & \text{otherwise.}\end{cases}$$

We let

$$\mathcal{R} = \operatorname{span}\{P_{\boldsymbol{w}} : \boldsymbol{w} \in W\}. \tag{4-7}$$

Let $Q_{\boldsymbol{w}}$ be the operator

$$Q_{\boldsymbol{w}}\delta_{\boldsymbol{x}} = \delta(\boldsymbol{w}=\boldsymbol{x})\delta_{\boldsymbol{x}};$$

i.e., $Q_{\boldsymbol{w}}$ is the Dirac delta function at $\boldsymbol{w}$ seen as a multiplication operator.

Lemma 4.6. *For $\boldsymbol{w} \in W$ we have $Q_{\boldsymbol{w}} = \sum_{\boldsymbol{v}\in C(\boldsymbol{w},+)}(-1)^{|\boldsymbol{w}^{-1}\boldsymbol{v}|}P_{\boldsymbol{v}}$.*

Proof. Firstly, $Q_{\boldsymbol{w}}(\boldsymbol{w}) = 1 = P_{\boldsymbol{w}}(\boldsymbol{w}) = \big(\sum_{\boldsymbol{v}\in C(\boldsymbol{w},+)}(-1)^{|\boldsymbol{w}^{-1}\boldsymbol{v}|}P_{\boldsymbol{v}}\big)(\boldsymbol{w})$. Let $\boldsymbol{x} \in W$. If $\boldsymbol{w} \not\le \boldsymbol{x}$ we get $Q_{\boldsymbol{w}}(\boldsymbol{x}) = 0 = \big(\sum_{\boldsymbol{v}\in C(\boldsymbol{w},+)}(-1)^{|\boldsymbol{w}^{-1}\boldsymbol{v}|}P_{\boldsymbol{v}}\big)(\boldsymbol{x})$. In the case $\boldsymbol{w} < \boldsymbol{x}$ we find

$$\Big(\sum_{\boldsymbol{v}\in C(\boldsymbol{w},+)}(-1)^{|\boldsymbol{w}^{-1}\boldsymbol{v}|}P_{\boldsymbol{v}}\Big)(\boldsymbol{x}) = \sum_{\boldsymbol{v}\in C(\boldsymbol{w},\boldsymbol{x})}(-1)^{|\boldsymbol{w}^{-1}\boldsymbol{v}|}, \tag{4-8}$$

and this expression equals 0 by the binomial formula. Indeed, let $\Lambda_{\boldsymbol{w},\boldsymbol{x}}$ be the maximal clique appearing at the start of $\boldsymbol{w}^{-1}\boldsymbol{x}$ (see Section 4A). The number of words smaller than $\Lambda_{\boldsymbol{w},\boldsymbol{x}}$ of length l is $|\Lambda_{\boldsymbol{w},\boldsymbol{x}}|$ choose l. So (4-8) equals

$$\sum_{l=0}^{|\Lambda_{\boldsymbol{w},\boldsymbol{x}}|} \sum_{\boldsymbol{v}\in C(\boldsymbol{w},\boldsymbol{x}),|\boldsymbol{w}^{-1}\boldsymbol{v}|=l} (-1)^{|\boldsymbol{w}^{-1}\boldsymbol{v}|} = \sum_{l=0}^{|\Lambda_{\boldsymbol{w},\boldsymbol{x}}|} \binom{|\Lambda_{\boldsymbol{w},\boldsymbol{x}}|}{l} (-1)^{|\boldsymbol{w}^{-1}\boldsymbol{v}|} = 0. \qquad \square$$

Now let $\mathcal{A}_q$ be the $*$-algebra generated by the operators $T_{\boldsymbol{w}}$, $\boldsymbol{w} \in W$. So $\mathcal{M}_q$ is the σ-weak closure of $\mathcal{A}_q$. We define

$$\Psi_{\leq n} : \mathcal{A}_q \to \mathcal{M}_q : T_{\boldsymbol{w}} \mapsto \begin{cases} T_{\boldsymbol{w}}, & |\boldsymbol{w}| \leq n, \\ 0, & \text{otherwise.} \end{cases}$$

We also set $\Psi_n = \Psi_{\leq n} - \Psi_{\leq(n-1)}$. The crucial part which we need to prove is that $\Psi_{\leq n}$ is completely bounded with a complete bound that can be upper-estimated in n polynomially. In order to do so we first introduce three auxiliary maps.

Auxiliary map 1: Recall that $\mathcal{M}_1$ is just the group von Neumann algebra of the right-angled Coxeter group W. For $k \in \mathbb{N}$ define the multiplier $\mathcal{A}_1 \to \mathcal{A}_1$

$$\rho_k(T_{\boldsymbol{w}}^{(1)}) = \delta(|\boldsymbol{w}| = k) T_{\boldsymbol{w}}^{(1)}.$$

This map is completely bounded as the range is finite-dimensional. We may extend ρ_k to a σ-weakly continuous map $\mathcal{M}_1 \to \mathcal{M}_1$ (for convenience of the reader we provided details of this extension trick through double duality in Theorem 4.13). By the Bożejko–Fendler theorem, Theorem 2.8, we may extend ρ_k uniquely to a σ-weakly continuous $\ell^\infty(W)$-bimodule map $\mathcal{B}(\ell^2(W)) \to \mathcal{B}(\ell^2(W))$ with the same completely bounded norm. Using Lemma 2.7 we see that

$$\Psi_{\leq n} = \sum_{k=0}^{n} \rho_k \circ \Psi_{\leq n}.$$

We emphasize at this point that in our proofs we shall not need a growth estimate for $\|\rho_k\|_{\mathcal{CB}}$ in terms of k. It is known however by [Reckwerdt 2015] that $\|\rho_k\|_{\mathcal{CB}}$ admits a polynomial bound in k. In the hyperbolic case this would already follow from [Ozawa 2008, Theorem 1(2)].

Only in the hyperbolic case it is known by [Ozawa 2008, Theorem 1(2)] that this map is completely bounded and moreover $\|\rho_k\|_{\mathcal{CB}} \leq C(k+1)$ for some constant C independent of k.

Auxiliary map 2: Let $\mathbb{T}$ be the unit circle in $\mathbb{C}$. For $z \in \mathbb{T}$ we define a unitary map,

$$A_z : \ell^2(W) \to \ell^2(W) : \delta_{\boldsymbol{w}} \mapsto z^{|\boldsymbol{w}|}\delta_{\boldsymbol{w}}.$$

We set for $i \in \mathbb{Z}$,

$$\Phi_i : \mathcal{B}(\ell^2(W)) \to \mathcal{B}(\ell^2(W)) : x \mapsto \int_{\mathbb{T}} z^{-i} A_z^* x A_z \, dz,$$

where the measure is the normalized Lebesgue measure on $\mathbb{T}$. Intuitively Φ_i cuts out the operators that create i more letters than they annihilate (where a negative creation is an annihilation). Using Lemma 2.7

we see that

$$\Psi_{\leq n} = \sum_{i=-n}^{n} \Phi_i \circ \Psi_{\leq n}.$$

Auxiliary map 3: Assume that Γ is finite. For $a \in \mathbb{N}$ we define Stinespring dilations

$$U_a^{\pm} : \ell^2(W) \to \ell^2(W) \otimes \ell^2(W) \otimes \ell^2(W) \otimes \ell^2(B_f(W)) \tag{4-9}$$

by mapping $\delta_{\boldsymbol{x}}$ to (see Section 4A for notation)

$$\sum_{\boldsymbol{g} \leq \boldsymbol{x}} \sum_{\Lambda \leq \boldsymbol{g}(2,\varnothing)} \beta^{\pm}_{\boldsymbol{g},\boldsymbol{x},\Lambda,a} \delta_{\boldsymbol{g}} \otimes \delta_{\boldsymbol{g}^{-1}\boldsymbol{x}} \otimes \delta_{\boldsymbol{g}(2,\Lambda)} \otimes \tilde{\xi}^{\pm}_{\Lambda}.$$

Here

$$\beta^{+}_{\boldsymbol{g},\boldsymbol{x},\Lambda,a} = \sum_{\boldsymbol{v} \in C(\boldsymbol{g},\boldsymbol{x})} (-1)^{|\boldsymbol{g}^{-1}\boldsymbol{v}|} F_{\Lambda,a}(\boldsymbol{v}), \tag{4-10}$$

where $F_{\Lambda,a}(\boldsymbol{v}) = 1$ if

$$2|\boldsymbol{v}(1,\Lambda)| + |\boldsymbol{v}(2,\Lambda)| \leq a,$$

and else $F_{\Lambda,a}(\boldsymbol{v}) = 0$. We let $\beta^{-}_{\boldsymbol{g},\boldsymbol{x},\Lambda,a} = 1$ if $\beta^{+}_{\boldsymbol{g},\boldsymbol{x},\Lambda,a} \neq 0$ and $\beta^{-}_{\boldsymbol{g},\boldsymbol{x},\Lambda,a} = 0$ otherwise. Then set,

$$\sigma_{a,b}(x) = (U_a^{-})^*(x \otimes 1 \otimes 1 \otimes 1)U_b^{+}. \tag{4-11}$$

The map $U_a^{\pm}$ is bounded with polynomial bound in a by the following lemma.

Lemma 4.7. *If Γ is finite, the map $U_a^{\pm}$ is bounded. Moreover, there exists a polynomial P such that $\|U_a^{\pm}\| \leq P(a)$.*

Proof. It follows by a comparison of the first two tensor legs in the definition of $U_a^{\pm}$ that the images of $\delta_{\boldsymbol{x}}$, $\boldsymbol{x} \in W$, are orthogonal vectors. Therefore it suffices to show that $\sup_{\boldsymbol{x} \in W} \|U_a^{\pm}\delta_{\boldsymbol{x}}\|$ is bounded polynomially. Now let $C = \sum_{\Lambda \in \mathrm{Cliq}(\Gamma)} 2^{|\Lambda|/2}$. Then

$$\begin{aligned}\|U_a\delta_{\boldsymbol{x}}\| &= \Big\| \sum_{\boldsymbol{g} \leq \boldsymbol{x}} \sum_{\Lambda \leq \boldsymbol{g}(2,\varnothing)} \beta^{\pm}_{\boldsymbol{g},\boldsymbol{x},\Lambda,a} \delta_{\boldsymbol{g}} \otimes \delta_{\boldsymbol{g}^{-1}\boldsymbol{x}} \otimes \delta_{\boldsymbol{g}(2,\Lambda)} \otimes \tilde{\xi}^{\pm}_{\Lambda} \Big\| \\ &\leq \sum_{\boldsymbol{g} \leq \boldsymbol{x}} \sum_{\Lambda \leq \boldsymbol{g}(2,\varnothing)} |\beta^{\pm}_{\boldsymbol{g},\boldsymbol{x},\Lambda,a}| 2^{\frac{1}{2}|\Lambda|} \leq C \sum_{\boldsymbol{g} \leq \boldsymbol{x}} \max_{\Lambda \in \mathrm{Cliq}(\Gamma)} |\beta^{\pm}_{\boldsymbol{g},\boldsymbol{x},\Lambda,a}|. \end{aligned} \tag{4-12}$$

In the case

$$a \leq 2|\boldsymbol{g}(1,\Lambda)| + |\boldsymbol{g}(2,\Lambda)|, \tag{4-13}$$

$\beta^{\pm}_{\boldsymbol{g},\boldsymbol{x},\Lambda,a} = 0$ by definition. Inequality (4-13) will certainly hold when $a \leq |\boldsymbol{g}|$. Let M be the maximum length of a clique in $\mathrm{Cliq}(\Gamma)$. Then if

$$2|\boldsymbol{g}(1,\Lambda)| + |\boldsymbol{g}(2,\Lambda)| \leq a - 2M - 1, \tag{4-14}$$

we find that $\beta^{\pm}_{\boldsymbol{g},\boldsymbol{x},\Lambda,a} = 0$ by the binomial formula as for every $\boldsymbol{v} \in C(\boldsymbol{g},\boldsymbol{x})$ we have $F_{\Lambda,a}(\boldsymbol{v}) = 1$. Inequality (4-14) will certainly hold if $2|\boldsymbol{g}| \leq a - 2M - 1$. So (4-12) can be estimated by C times the number of

$\boldsymbol{g} \leq \boldsymbol{x}$ with
$$\tfrac{1}{2}(a-2M-1) \leq |\boldsymbol{g}| \leq a.$$
But the number such $\boldsymbol{g}$'s grows polynomially in a; see Lemma 4.4. □

Lemma 4.8. *Let $\boldsymbol{x} \in W$. Let $\boldsymbol{u}', \boldsymbol{u}'' \in W$ be such that $|\boldsymbol{u}''\boldsymbol{x}| = |\boldsymbol{x}| - |\boldsymbol{u}''|$, $|\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x}| = |\boldsymbol{x}| - |\boldsymbol{u}''| + |\boldsymbol{u}'|$. Let $\boldsymbol{v} \in W$ be such that $(\boldsymbol{u}'')^{-1} \leq \boldsymbol{v} \leq \boldsymbol{x}$. Then,*
$$\sum_{\boldsymbol{v} \leq \boldsymbol{g} \leq \boldsymbol{x}} \beta^{+}_{\boldsymbol{g},\boldsymbol{x},\boldsymbol{g}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing),a} \beta^{-}_{\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g},\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x},(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing)\backslash\boldsymbol{g}(2,\varnothing),a-2|\boldsymbol{u}'|+2|\boldsymbol{u}''|}$$
$$= \begin{cases} 1 & \text{if } 2|\boldsymbol{v}(1, \boldsymbol{v}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{v})(2,\varnothing))| + |\boldsymbol{v}(2, \boldsymbol{v}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{v})(2,\varnothing))| \leq a, \\ 0 & \text{otherwise}. \end{cases} \tag{4-15}$$

Proof. By (4-1) and (4-2) for $\boldsymbol{v} \leq \boldsymbol{g}$ we get
$$\begin{aligned}\beta^{+}_{\boldsymbol{g},\boldsymbol{x},\boldsymbol{g}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing),a} &= \sum_{\boldsymbol{w} \in C(\boldsymbol{g},\boldsymbol{x})} (-1)^{|\boldsymbol{g}^{-1}\boldsymbol{v}|} F_{\boldsymbol{g}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing),a}(\boldsymbol{w}) \\ &= \sum_{\boldsymbol{w} \in C(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g},\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x})} (-1)^{|\boldsymbol{g}^{-1}\boldsymbol{v}|} F_{(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing)\backslash\boldsymbol{g}(2,\varnothing),a-2|\boldsymbol{u}''|+2|\boldsymbol{u}'|}(\boldsymbol{w}) \\ &= \beta^{+}_{\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g},\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x},(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing)\backslash\boldsymbol{g}(2,\varnothing),a-2|\boldsymbol{u}'|+2|\boldsymbol{u}''|}.\end{aligned}$$
Therefore also,
$$\beta^{-}_{\boldsymbol{g},\boldsymbol{x},\boldsymbol{g}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing),a} = \beta^{-}_{\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g},\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x},(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing)\backslash\boldsymbol{g}(2,\varnothing),a-2|\boldsymbol{u}'|+2|\boldsymbol{u}''|}.$$
We thus have that the left-hand side of (4-15) equals
$$\sum_{\boldsymbol{v} \leq \boldsymbol{g} \leq \boldsymbol{x}} \beta^{+}_{\boldsymbol{g},\boldsymbol{x},\boldsymbol{g}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing),a} \beta^{-}_{\boldsymbol{g},\boldsymbol{x},\boldsymbol{g}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing),a} = \sum_{\boldsymbol{v} \leq \boldsymbol{g} \leq \boldsymbol{x}} \beta^{+}_{\boldsymbol{g},\boldsymbol{x},\boldsymbol{g}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing),a}.$$
To compute this sum, recall that $\mathcal{R}$ was defined in (4-7), and define the mapping
$$\kappa_a : \mathcal{R} \to \mathcal{R} : P_{\boldsymbol{w}} \mapsto F_{\boldsymbol{w}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{w})(2,\varnothing),a}(\boldsymbol{w}) P_{\boldsymbol{w}}.$$
Then, using Lemma 4.6, the definition of κ_a, Lemma 4.3, and the definition (4-10),
$$\begin{aligned}\kappa_a(Q_{\boldsymbol{g}})(\boldsymbol{x}) &= \kappa_a\Bigl(\sum_{\boldsymbol{w} \in C(\boldsymbol{g},\boldsymbol{x})} (-1)^{|\boldsymbol{g}^{-1}\boldsymbol{w}|} P_{\boldsymbol{w}}\Bigr)(\boldsymbol{x}) = \sum_{\boldsymbol{w} \in C(\boldsymbol{g},\boldsymbol{x})} (-1)^{|\boldsymbol{g}^{-1}\boldsymbol{w}|} F_{\boldsymbol{w}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{w})(2,\varnothing),a}(\boldsymbol{w}) \\ &= \beta^{+}_{\boldsymbol{g},\boldsymbol{x},\boldsymbol{g}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing),a}.\end{aligned}$$
As $\sum_{\boldsymbol{v} \leq \boldsymbol{g} \leq \boldsymbol{x}} Q_{\boldsymbol{g}}$ can be written as $P_{\boldsymbol{v}}$ plus projections in $\mathcal{R}$ that are not supported at $\boldsymbol{x}$ we see therefore that
$$\sum_{\boldsymbol{v} \leq \boldsymbol{g} \leq \boldsymbol{x}} \beta^{+}_{\boldsymbol{g},\boldsymbol{x},\boldsymbol{g}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing),a} = \sum_{\boldsymbol{v} \leq \boldsymbol{g} \leq \boldsymbol{x}} \kappa_a(Q_{\boldsymbol{g}})(\boldsymbol{x}) = \kappa_a(P_{\boldsymbol{v}})(\boldsymbol{x}).$$
This expression equals 1 if $F_{\boldsymbol{v}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{v})(2,\varnothing),a}(\boldsymbol{v}) = 1$ and 0 otherwise, which corresponds exactly to the statement of the lemma. □

Lemma 4.9. *Assume that Γ is finite so that (4-11) is defined boundedly. We have for $n \in \mathbb{N}$ that $\Psi_{\leq n} = \sum_{i=-n}^{n} \sigma_{n-i,n+i} \circ \Phi_i \circ \Psi_{\leq n}$.*

Proof. Let $T_{\boldsymbol{w}} \in \mathcal{M}_q$ with $|\boldsymbol{w}| \leq n$. We need to show that

$$T_{\boldsymbol{w}} = \sum_{k=0}^{n} \sum_{i=-n}^{n} \sigma_{n-i,n+i} \circ \Phi_i \circ \rho_k(T_{\boldsymbol{w}}).$$

We split $T_{\boldsymbol{w}}$ by Lemma 2.7,

$$T_{\boldsymbol{w}} = \sum_{(\boldsymbol{w}',\Gamma_0,\boldsymbol{w}'')\in A_{\boldsymbol{w}}} T_{\boldsymbol{w}'}^{(1)} P_{V\Gamma_0} T_{\boldsymbol{w}''}^{(1)},$$

and show that

$$\sum_{k=0}^{n} \sum_{i=-n}^{n} \sigma_{n-i,n+i} \circ \Phi_i \circ \rho_k$$

applied to each of these summands acts as the identity. Let us consider a summand $T_{\boldsymbol{w}'}^{(1)} P_{V\Gamma_0} T_{\boldsymbol{w}''}^{(1)}$ with $(\boldsymbol{w}', \Gamma_0, \boldsymbol{w}'') \in A_{\boldsymbol{w}}$. Let $\boldsymbol{u}, \boldsymbol{u}', \boldsymbol{u}''$ be as in Lemma 2.5 so that $T_{\boldsymbol{w}'}^{(1)} P_{V\Gamma_0} T_{\boldsymbol{w}''}^{(1)} = T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}V\Gamma_0} T_{\boldsymbol{u}''}^{(1)}$. We have

$$\rho_k(T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}V\Gamma_0} T_{\boldsymbol{u}''}^{(1)}) = \begin{cases} T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}V\Gamma_0} T_{\boldsymbol{u}''}^{(1)} & \text{if } k = |\boldsymbol{u}'| + |\boldsymbol{u}''|, \\ 0 & \text{otherwise.} \end{cases}$$

So the only nonzero summand is $k = |\boldsymbol{u}'| + |\boldsymbol{u}''|$ so that it remains to show that for $\boldsymbol{x}, \boldsymbol{y} \in W$

$$\left\langle \sum_{i=-n}^{n} \sigma_{n-i,n+i} \circ \Phi_i(T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}V\Gamma_0} T_{\boldsymbol{u}''}^{(1)})\delta_{\boldsymbol{x}}, \delta_{\boldsymbol{y}} \right\rangle = \langle T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}V\Gamma_0} T_{\boldsymbol{u}''}^{(1)}\delta_{\boldsymbol{x}}, \delta_{\boldsymbol{y}} \rangle. \tag{4-16}$$

If the right-hand side is nonzero then we must have $\boldsymbol{y} = \boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x}$. Furthermore, recall that there is a choice for $\boldsymbol{u}', \boldsymbol{u}''$ and we may choose them (depending on $\boldsymbol{x}$) such that $|\boldsymbol{u}''\boldsymbol{x}| = |\boldsymbol{x}| - |\boldsymbol{u}''|$ and $|\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x}| = |\boldsymbol{x}| - |\boldsymbol{u}''| + |\boldsymbol{u}'|$. After making this choice the right-hand side is nonzero in the case $(\boldsymbol{u}'')^{-1}\boldsymbol{u}V\Gamma_0 \leq \boldsymbol{x}$, in which case the expression equals 1.

Now consider the left-hand side of (4-16),

$$\begin{aligned} &\langle (\Phi_i(T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}V\Gamma_0} T_{\boldsymbol{u}''}^{(1)}) \otimes 1 \otimes 1 \otimes 1) U_{n-i}^{+}\delta_{\boldsymbol{x}}, U_{n+i}^{-}\delta_{\boldsymbol{y}} \rangle \\ &\quad = \Big\langle \sum_{\boldsymbol{g}\leq \boldsymbol{x}} \sum_{\Lambda \leq \boldsymbol{g}(2,\varnothing)} \beta^{+}_{\boldsymbol{g},\boldsymbol{x},\Lambda,n-i} \Phi_i(T_{\boldsymbol{u}'}^{(1)} P_{\boldsymbol{u}V\Gamma_0} T_{\boldsymbol{u}''}^{(1)})\delta_{\boldsymbol{g}} \otimes \delta_{\boldsymbol{g}^{-1}\boldsymbol{x}} \otimes \delta_{\boldsymbol{g}(2,\Lambda)} \otimes \tilde{\xi}^{+}_{\Lambda}, \\ &\qquad\qquad \sum_{\boldsymbol{h}\leq \boldsymbol{y}} \sum_{\Lambda' \leq \boldsymbol{h}(2,\varnothing)} \beta^{-}_{\boldsymbol{h},\boldsymbol{y},\Lambda',n+i} \delta_{\boldsymbol{h}} \otimes \delta_{\boldsymbol{h}^{-1}\boldsymbol{y}} \otimes \delta_{\boldsymbol{h}(2,\Lambda')} \otimes \tilde{\xi}^{-}_{\Lambda} \Big\rangle. \end{aligned} \tag{4-17}$$

Comparing the first two tensor legs of this equation we derive the following. The only summands that are nonzero are the ones where $\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g} = \boldsymbol{h}$ and at the same time $\boldsymbol{g}^{-1}\boldsymbol{x} = \boldsymbol{h}^{-1}\boldsymbol{y}$. In particular we must have $\boldsymbol{y} = \boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x}$ and there is a choice for $\boldsymbol{u}', \boldsymbol{u}''$ (same choice as above) such that in fact $|\boldsymbol{u}''\boldsymbol{x}| = |\boldsymbol{x}| - |\boldsymbol{u}''|$ and $|\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x}| = |\boldsymbol{x}| - |\boldsymbol{u}''| + |\boldsymbol{u}'|$. We also see that we must have $(\boldsymbol{u}'')^{-1}\boldsymbol{u}V\Gamma_0 \leq \boldsymbol{x}$ for this expression to be nonzero. Taking into account Φ_i we see that (4-17) is nonzero only if $i = |\boldsymbol{u}''| - |\boldsymbol{u}'|$.

Next we note that by comparing the last two tensor legs, if a summand in (4-17) is nonzero then we have $\boldsymbol{g}(2, \Lambda) = \boldsymbol{h}(2, \Lambda')$ and $\Lambda \cap \Lambda' = \varnothing$. Recall that $\boldsymbol{h} = \boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g}$. But then Λ must equal the letters in $\boldsymbol{g}(2, \varnothing)$ that are not anymore in $(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2, \varnothing)$ and Λ' must equal the letters in $(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2, \varnothing)$ that are not anymore in $\boldsymbol{g}(2, \varnothing)$. This precisely means that $\Lambda = \boldsymbol{g}(2, \varnothing) \backslash (\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2, \varnothing)$ and $\Lambda' = (\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g}) \backslash \boldsymbol{g}(2, \varnothing)$.

In all, we find that

$$
\begin{aligned}
(4\text{-}17) &= \langle (T^{(1)}_{\boldsymbol{u}'} P_{\boldsymbol{u}V\Gamma_0} T^{(1)}_{\boldsymbol{u}''} \otimes 1 \otimes 1 \otimes 1) U^{+}_{n-i}\delta_{\boldsymbol{x}}, U^{-}_{n+i}\delta_{\boldsymbol{y}} \rangle \\
&= \Big\langle \sum_{\boldsymbol{g}\leq \boldsymbol{x}} \sum_{\Lambda \leq \boldsymbol{g}(2,\varnothing)} \beta^{+}_{\boldsymbol{g},\boldsymbol{x},\Lambda,n-i} T^{(1)}_{\boldsymbol{u}'} P_{\boldsymbol{u}V\Gamma_0} T^{(1)}_{\boldsymbol{u}''} \delta_{\boldsymbol{g}} \otimes \delta_{\boldsymbol{g}^{-1}\boldsymbol{x}} \otimes \delta_{\boldsymbol{g}(2,\Lambda)} \otimes \tilde{\xi}^{+}_{\Lambda}, \\
&\qquad\qquad \sum_{\boldsymbol{h}\leq \boldsymbol{y}} \sum_{\Lambda' \leq \boldsymbol{h}(2,\varnothing)} \beta^{-}_{\boldsymbol{h},\boldsymbol{y},\Lambda',n+i} \delta_{\boldsymbol{h}^{-1}\boldsymbol{x}} \otimes \delta_{\boldsymbol{h}} \otimes \delta_{\boldsymbol{h}(2,\Lambda')} \otimes \tilde{\xi}^{-}_{\Lambda} \Big\rangle \\
&= \sum_{(\boldsymbol{u}'')^{-1}\boldsymbol{u}V\Gamma_0 \leq \boldsymbol{g} \leq \boldsymbol{x}} \beta^{+}_{\boldsymbol{g},\boldsymbol{x},\boldsymbol{g}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing),n-i} \beta^{-}_{\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g},\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x},(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing)\backslash \boldsymbol{g}(2,\varnothing),n+i}.
\end{aligned}
$$

We claim that this expression is 1 by verifying Lemma 4.8. Indeed set $\boldsymbol{w} := (\boldsymbol{u}'')^{-1}\boldsymbol{u}V\Gamma_0$. First suppose that $\boldsymbol{u}$ is the empty word. Then

$$\boldsymbol{w}(2, \boldsymbol{w}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{w})(2,\varnothing)) = V\Gamma_0$$

and so

$$\boldsymbol{w}(1, \boldsymbol{w}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{w})(2,\varnothing)) = (\boldsymbol{u}'')^{-1}.$$

If $\boldsymbol{u}$ is not the empty word, then let $s \in W$ be a final letter of $\boldsymbol{u}$ (i.e., $|\boldsymbol{u}s| = |\boldsymbol{u}| - 1$). Then s cannot commute with $V\Gamma_0$ as this would violate the equation $T^{(1)}_{\boldsymbol{u}'} P_{\boldsymbol{u}V\Gamma_0} T^{(1)}_{\boldsymbol{u}''} = T^{(1)}_{\boldsymbol{w}'} P_{V\Gamma_0} T^{(1)}_{\boldsymbol{w}''}$. Therefore again,

$$\boldsymbol{w}(2, \boldsymbol{w}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{w})(2,\varnothing)) = \boldsymbol{w}(2,\varnothing) = V\Gamma_0$$

and so

$$\boldsymbol{w}(1, \boldsymbol{w}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{w})(2,\varnothing)) = (\boldsymbol{u}'')^{-1}\boldsymbol{u}.$$

Further our constructions give that $|\boldsymbol{u}''| = (k-i)/2$ and $2|\boldsymbol{u}| + |V\Gamma_0| = |\boldsymbol{w}| - |\boldsymbol{u}'| - |\boldsymbol{u}''| = |\boldsymbol{w}| - k$. So we have

$$
\begin{aligned}
&2|\boldsymbol{w}(1, \boldsymbol{w}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{w})(2,\varnothing))| + |\boldsymbol{w}(2, \boldsymbol{w}(2,\varnothing)\backslash(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{w})(2,\varnothing))| \\
&\qquad = 2|(\boldsymbol{u}'')^{-1}| + 2|\boldsymbol{u}| + |V\Gamma_0| = 2\frac{k-i}{2} + (|\boldsymbol{w}| - k) \\
&\qquad = |\boldsymbol{w}| - i \leq n - i, \qquad (4\text{-}18)
\end{aligned}
$$

so that by Lemma 4.8 we see that (4-17) is 1. So we conclude that (4-16) holds. □

Lemma 4.10. *Assume that* Γ *is finite so that* (4-11) *is defined boundedly. We have for* $n \in \mathbb{N}$, $-n \leq i \leq n$,

$$\sigma_{n-i,n+i} \circ \Phi_i \circ \Psi_{\leq n} = \sigma_{n-i,n+i} \circ \Phi_i.$$

Proof. The proof pretty much parallels the proof of Lemma 4.9. We need to show that the right-hand side applied to $T_{\boldsymbol{w}}$ with $|\boldsymbol{w}| > n$ equals 0. Therefore we may look at the summands $T^{(1)}_{\boldsymbol{w}'} P_{V\Gamma_0} T^{(1)}_{\boldsymbol{w}''}$ with $(\boldsymbol{w}', \Gamma_0, \boldsymbol{w}'') \in A_{\boldsymbol{w}}$ which can be further decomposed as $T^{(1)}_{\boldsymbol{u}'} P_{\boldsymbol{u}V\Gamma_0} T^{(1)}_{\boldsymbol{u}''}$ with $\boldsymbol{u}, \boldsymbol{u}', \boldsymbol{u}''$ as in Lemma 2.5. It suffices then to show that for all choices of k the following expression is 0:

$$\langle \sigma_{n-i,n+i} \circ \Phi_i \circ \rho_k (T^{(1)}_{\boldsymbol{u}'} P_{\boldsymbol{u}V\Gamma_0} T^{(1)}_{\boldsymbol{u}''}) \delta_{\boldsymbol{x}}, \delta_{\boldsymbol{y}} \rangle. \qquad (4\text{-}19)$$

Firstly, this expression is 0 in the case $|\boldsymbol{u}'| + |\boldsymbol{u}''| \neq k$. So assume $|\boldsymbol{u}'| + |\boldsymbol{u}''| = k$. Then,

$$(4\text{-}19) = \langle \sigma_{n-i,n+i} \circ \Phi_i (T^{(1)}_{\boldsymbol{u}'} P_{\boldsymbol{u}V\Gamma_0} T^{(1)}_{\boldsymbol{u}''}) \delta_{\boldsymbol{x}}, \delta_{\boldsymbol{y}} \rangle.$$

As in the proof of Lemma 4.9 the expression (4-19) equals 0 unless $\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x} = \boldsymbol{y}$ and $(\boldsymbol{u}'')^{-1}\boldsymbol{u}V\Gamma_0 \leq \boldsymbol{x}$ with $\boldsymbol{u}'', \boldsymbol{u}'$ chosen in such a way that $|\boldsymbol{u}''\boldsymbol{x}| = |\boldsymbol{x}| - |\boldsymbol{u}''|$ and $|\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x}| = |\boldsymbol{x}| - |\boldsymbol{u}''| + |\boldsymbol{u}'|$. In that case $i = |\boldsymbol{u}'| - |\boldsymbol{u}''|$. As in (4-17),

$$\begin{aligned}(4\text{-}19) &= \langle (T^{(1)}_{\boldsymbol{u}'} P_{\boldsymbol{u}V\Gamma_0} T^{(1)}_{\boldsymbol{u}''} \otimes 1 \otimes 1 \otimes 1) U^{+}_{n-i}\delta_{\boldsymbol{x}}, U^{-}_{n+i}\delta_{\boldsymbol{y}}\rangle \\ &= \sum_{(\boldsymbol{u}'')^{-1}\boldsymbol{u}V\Gamma_0 \leq \boldsymbol{g} \leq \boldsymbol{x}} \beta_{\boldsymbol{g},\boldsymbol{x},\boldsymbol{g}(2,\varnothing)\setminus(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing),n-i}\, \beta_{\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g},\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{x},(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{g})(2,\varnothing)\setminus \boldsymbol{g}(2,\varnothing),n+i}.\end{aligned} \tag{4-20}$$

As for $\boldsymbol{w} := (\boldsymbol{u}'')^{-1}\boldsymbol{u}V\Gamma_0$ we have again by the same reasoning as in/before (4-18) that

$$2|\boldsymbol{w}(1, \boldsymbol{w}(2,\varnothing)\setminus(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{w})(2,\varnothing))| + |\boldsymbol{w}(2, \boldsymbol{w}(2,\varnothing)\setminus(\boldsymbol{u}'\boldsymbol{u}''\boldsymbol{w})(2,\varnothing))| = |\boldsymbol{w}| - i > n - i.$$

The expression (4-20) is 0 by Lemma 4.8. □

Proposition 4.11. *We have $\|\Psi_{\leq n}\|_{\mathcal{CB}} \leq P(n)$ for some polynomial P.*

Proof. By Lemmas 4.9 and 4.10 we have

$$\Psi_{\leq n} = \sum_{i=-n}^{n} \sigma_{n-i,n+i} \circ \Phi_i \circ \Psi_{\leq n} = \sum_{i=-n}^{n} \sigma_{n-i,n+i} \circ \Phi_i,$$

and the right-hand side is completely bounded with polynomial bound in n; indeed the bound of $\sigma_{n-i,n+i}$ is polynomial in n by its very definition and Lemma 4.7. □

Definition 4.12. A von Neumann algebra $\mathcal{M}$ has the weak-$*$ completely bounded approximation property (wk-$*$ CBAP) if there exists a net of normal finite-rank maps $\Phi_i : \mathcal{M} \to \mathcal{M}$ such that $\Phi_i(x) \to x$ in the σ-weak topology and moreover $\sup_i \|\Phi_i\|_{\mathcal{CB}} < \infty$. If the maps Φ_i can be chosen so that $\limsup_i \|\Phi_i\|_{\mathcal{CB}} \leq 1$ then $\mathcal{M}$ is said to have the weak-$*$ completely contractive approximation property (wk-$*$ CCAP).

Theorem 4.13. *Let (W, S) be a right-angled Coxeter system and let $q > 0$. The Hecke von Neumann algebra $\mathcal{M}_q$ has the wk-$*$ CCAP.*

Proof. By an inductive limit argument and Corollary 3.3 we may assume that Γ is finite. The proof goes back to [Haagerup 1978]. Consider the completely bounded map $\Psi_{\leq n} \circ \Phi_r : \mathcal{A}_q \to \mathcal{M}_q$. Clearly as $n \to \infty$ and $r \nearrow 1$ this map converges to the identity in the point σ-weak topology. Let $\epsilon > 0$. We have

$$\|\Psi_{\leq n} \circ \Phi_r\|_{\mathcal{CB}} \leq \|(\Psi_{\leq n} - \mathrm{Id}) \circ \Phi_r\|_{\mathcal{CB}} + \|\Phi_r\|_{\mathcal{CB}} \leq \left(\sum_{i=n+1}^{\infty} r^n \|\Psi_n\|_{\mathcal{CB}} \right) + \|\Phi_r\|_{\mathcal{CB}},$$

which shows using Propositions 4.11 and 3.7 that we may let $r \nearrow 1$ and then choose $n := n_r$ converging to ∞ such that $\|\Psi_{\leq n_r} \circ \Phi_r\|_{\mathcal{CB}} \leq 1 + \epsilon$ for some constant.

The map Φ_r is normal. Also $\Psi_{\leq n}$ is normal by a standard argument: indeed using duality and Kaplansky's density theorem one sees that Ψ_n maps $L^1(\mathcal{M}_q) \to L^1(\mathcal{M}_q)$ boundedly. Then taking the dual of this map yields that $\Psi_n : \mathcal{M}_q \to \mathcal{M}_q$ is a normal map. We may extend $\Psi_{\leq n} \circ \Phi_r$ to a normal map $\mathcal{M}_q \to \mathcal{M}_q$. Then using a standard approximation argument yields the result. □

Remark 4.14. In case our right-angled Coxeter group is free (i.e., $m(s,t)=\infty$ for all $s \neq t$) it is possible to adapt the arguments of [Ricard and Xu 2006] in order to obtain word-length cut-downs with polynomial bound. This argument — purely based on bookkeeping of creations/annihilations — seems unrepairable in the general case. In the case $q=1$, for a general right-angled Coxeter group, word-length cut-downs were obtained in [Reckwerdt 2015] by using actions on CAT(0)-spaces. The connection with the general Hecke case is unclear.

5. Strong solidity in the hyperbolic case

We prove that in the factorial case (see Theorem 2.2) $\mathcal{M}_q$ is a strongly solid von Neumann algebra in the case the Coxeter group is hyperbolic.

5A. *Preliminaries on strongly solid algebras.* The *normalizer* of a von Neumann subalgebra $\mathcal{P}$ of $\mathcal{M}$ is defined as $\{u \in \mathcal{U}(\mathcal{M}) : u\mathcal{P}u^* = \mathcal{P}\}$. We define $\mathsf{Nor}_{\mathcal{P}}(\mathcal{M})$ as the von Neumann algebra generated by the normalizer of $\mathcal{P}$ in $\mathcal{M}$. A von Neumann algebra is called *diffuse* if it does not contain minimal projections.

Definition 5.1. A finite von Neumann algebra $\mathcal{M}$ is *strongly solid* if for any diffuse injective von Neumann subalgebra $\mathcal{P} \subseteq \mathcal{M}$ the von Neumann algebra $\mathsf{Nor}_{\mathcal{M}}(\mathcal{P})$ is again injective.

Ozawa and Popa [2010] proved that free group factors are strongly solid and consequently they could prove that these are II_1-factors that have no Cartan subalgebras (as was proved in [Voiculescu 1996] earlier by a completely different method). A general source for strongly solid von Neumann algebras are group von Neumann algebras of groups that have the weak-$*$ completely bounded approximation property and are biexact (see [Chifan and Sinclair 2013; Chifan et al. 2013; Popa and Vaes 2014]; we also refer to these sources for the definition of biexactness). The following definition and subsequent theorem were then introduced and proved in [Isono 2015]. For standard forms of von Neumann algebras we refer to [Takesaki 2003].

Definition 5.2. Let $\mathcal{M} \subseteq \mathcal{B}(\mathcal{H})$ be a von Neumann algebra represented on the standard Hilbert space $\mathcal{H}$ with modular conjugation J. We say that $\mathcal{M}$ satisfies condition $(\mathrm{AO})^+$ if there exists a unital C^*-subalgebra $\mathsf{A} \subseteq \mathcal{M}$ that is σ-weakly dense in $\mathcal{M}$ and which satisfies the following two conditions:

(1) A is locally reflexive.

(2) There exists a ucp map $\theta : \mathsf{A} \otimes_{\min} J\mathsf{A}J \to \mathcal{B}(\mathcal{H})$ such that $\theta(a \otimes b) - ab$ is a compact operator on $\mathcal{H}$.

Theorem 5.3 [Isono 2015]. *Let $\mathcal{M}$ be a II_1-factor with separable predual. Suppose that $\mathcal{M}$ satisfies condition $(AO)^+$ and has the weak-$*$ completely bounded approximation property. Then $\mathcal{M}$ is strongly solid.*

A maximal abelian von Neumann subalgebra $\mathcal{P} \subseteq \mathcal{M}$ of a II_1-factor $\mathcal{M}$ is called a *Cartan subalgebra* if $\mathsf{Nor}_{\mathcal{M}}(\mathcal{P}) = \mathcal{M}$. It is then obvious that if $\mathcal{M}$ is a noninjective strongly solid II_1-factor, then $\mathcal{M}$ cannot contain a Cartan subalgebra. Therefore we will now prove that the Hecke von Neumann algebra $\mathcal{M}_q$ in the factorial, hyperbolic case satisfies condition $(\mathrm{AO})^+$.

5B. ***Crossed products.*** Let A be a C^*-algebra that is represented on a Hilbert space $\mathcal{H}$. Let $\alpha : \mathsf{G} \curvearrowright \mathsf{A}$ be a continuous action of a discrete group G on A. The reduced crossed product $\mathsf{A} \rtimes_r \mathsf{G}$ is the C^*-algebra of operators acting on $\mathcal{H} \otimes \ell^2(\mathsf{G})$ generated by operators

$$u_g := \sum_{h \in \mathsf{G}} 1 \otimes e_{gh,h}, \quad g \in \mathsf{G}, \qquad \text{and} \qquad \pi(x) := \sum_{h \in \mathsf{G}} h^{-1} \cdot x \otimes e_{h,h}, \quad x \in \mathsf{A}. \tag{5-1}$$

Here the convergence of the sums should be understood in the strong topology. There is also a universal crossed product $\mathsf{A} \rtimes_u \mathsf{G}$ for which we refer to [Brown and Ozawa 2008] (in the case we need it, it turns out to equal the reduced crossed product).

5C. ***Gromov boundary and condition (AO)$^+$.*** Let again (W, S) be a Coxeter system which we assume to be hyperbolic (see [Brown and Ozawa 2008, Section 5.3]). Let Λ be the associated Cayley tree. A geodesic ray starting at a point $\boldsymbol{w} \in \Lambda$ is a sequence $(\boldsymbol{w}, \boldsymbol{w}v_1, \boldsymbol{w}v_1v_2, \dots)$ such that $|\boldsymbol{w}v_1 \cdots v_n| = |\boldsymbol{w}| + n$. We typically write $\omega = (\omega(0), \omega(1), \dots)$ for a geodesic ray. Let ∂W be the Gromov boundary of W which is the collection of all geodesic rays starting at the identity of W modulo the equivalence $\omega_1 \simeq \omega_2$ if and only if $\lim_{x,y\to\infty} \mathrm{dist}(\omega_1(x), \omega_2(y)) = 0$. $W \cup \partial W$ may be topologized as in [Brown and Ozawa 2008, Section 5.3].

Let $W \curvearrowright W$ be the action by means of left translation. The action extends continuously to $W \cup \partial W$ and then restricts to an action $W \curvearrowright \partial W$. We may pull back this action to obtain $W \curvearrowright C(\partial W)$. As in this section we assumed that W is a hyperbolic group, the action $W \curvearrowright \partial W$ is well known to be amenable [Brown and Ozawa 2008], which implies that $C(\partial W) \rtimes_u W = C(\partial W) \rtimes_r W$, and furthermore this crossed product is a nuclear C^*-algebra. Let $f \in C(\partial W)$, let $\tilde{f}_1, \tilde{f}_2 \in C(W \cup \partial W)$ be two continuous extensions of f, and let f_1 and f_2 be their respective restrictions to W. Then $f_1 - f_2 \in C_0(W)$. That is, the multiplication map $f_1 - f_2$ acting on $\ell^2(W)$ determines a compact operator. So the assignment $f \mapsto f_1$ is a well-defined $*$-homomorphism $C(\partial W) \to \mathcal{B}(\ell^2(W))/\mathcal{K}$, where $\mathcal{K}$ are the compact operators on $\ell^2(W)$. It is easy to check that this map is W-equivariant and thus we obtain a $*$-homomorphism:

$$\pi_1 : C(\partial W) \rtimes_u W \to \mathcal{B}(\ell^2(W))/\mathcal{K}. \tag{5-2}$$

Let again $W \curvearrowright W$ be the action by means of left translation which may be pulled back to obtain an action $W \curvearrowright \ell^\infty(W)$. Let

$$\rho : \ell^\infty(W) \rtimes_r W \to \mathcal{B}(\ell^2(W))$$

be the σ-weakly continuous $*$-isomorphism determined by $\rho : u_{\boldsymbol{w}} \mapsto T_{\boldsymbol{w}}^{(1)}$ and $\rho : \pi(x) \mapsto x$ (see [Vaes 2001, Theorem 5.3]). In fact ρ is an injective map (this follows immediately from [De Commer 2011, Theorem 2.1] as the operator G in this theorem equals the multiplicative unitary/structure operator [Takesaki 2003, p. 68]). Let $C_\infty(W)$ be the C^*-algebra generated by the projections $P_{\boldsymbol{w}}$, $\boldsymbol{w} \in W$. Take $f \in C_\infty(W)$ and let $\tilde{f}$ be the continuous extension of f to $W \cup \partial W$. The map $f \mapsto \tilde{f}|_{\partial W}$ determines a $*$-homomorphism $\sigma : C_\infty(W) \to C(\partial W)$ that is W-equivariant. Therefore it extends to the crossed product map

$$\sigma \rtimes_r \mathrm{Id} : C_\infty(W) \rtimes_r W \to C(\partial W) \rtimes_r W.$$

Theorem 5.4. *Let (W, S) be a right-angled hyperbolic Coxeter group and let $q \in [\rho, \rho^{-1}]$; see Theorem 2.2. The von Neumann algebra $\mathcal{M}_q$ satisfies condition (AO)$^+$.*

Proof. We let A_q be the unital C*-subalgebra of $\mathcal{M}_q$ generated by operators $T_{\boldsymbol{w}}$, $\boldsymbol{w} \in W$. It is easy to see that A_q is preserved by the multipliers that we constructed in order to prove that $\mathcal{M}_q$ had the wk-$*$ CBAP; see Section 4 (indeed these were compositions of radial multipliers — see Proposition 3.7 — and word-length projections — see Proposition 4.11). Therefore A_q has the CBAP; hence by the remarks before [Haagerup and Kraus 1994, Theorem 2.2] it is exact. Therefore A_q is locally reflexive [Brown and Ozawa 2008; Pisier 2003, Chapter 18].

It remains to prove condition (2) of Definition 5.2. By Lemma 2.7 we see that A_q is contained in the C*-subalgebra of $\mathcal{B}(\ell^2(W))$ generated by the operators $P_{\boldsymbol{w}}, T_{\boldsymbol{w}}^{(1)}$ with $\boldsymbol{w} \in W$. So $\rho^{-1}(\mathsf{A}_q)$ is contained in $C_\infty(W) \rtimes_r W$ and therefore we may set

$$\gamma : \mathsf{A}_q \to C(\partial W) \rtimes_r W \quad \text{as } \gamma = (\sigma \rtimes_r \mathrm{Id}) \circ \rho^{-1}.$$

The mapping $\pi_2 : J\mathsf{A}_q J \to \mathcal{B}(\ell^2(W))/\mathcal{K} : b \mapsto b$ is a $*$-homomorphism and its image commutes with the image of π_1 of (5-2) (as was argued in [Higson and Guentner 2004, Lemma 6.2.8]). By the definition of the maximal tensor product there exists a $*$-homomorphism

$$(\pi_1 \otimes \pi_2) : (C(\partial W) \rtimes_u W) \otimes_{\max} J\mathsf{A}_q J \to \mathcal{B}(\ell^2(W))/\mathcal{K} : a \otimes JbJ \mapsto \pi_1(a)JbJ.$$

We may now consider the following composition of $*$-homomorphisms:

$$\begin{array}{ccc} \mathsf{A}_q \otimes_{\min} J\mathsf{A}_q J & \xrightarrow{\gamma\otimes \mathrm{id}} & (C(\partial W) \rtimes_r W) \otimes_{\min} J\mathsf{A}_q J \\ & & \downarrow \simeq \\ \mathcal{B}(\ell^2(W))/\mathcal{K} & \xleftarrow[\pi_1\otimes\pi_2]{} & (C(\partial W) \rtimes_u W) \otimes_{\max} J\mathsf{A}_q J. \end{array} \tag{5-3}$$

By construction this map is given by

$$a \otimes JbJ \mapsto aJbJ + \mathcal{K}, \quad \text{where } a, b \in \mathsf{A}_q. \tag{5-4}$$

The map π_1 is nuclear because we already observed that $C(\partial W) \rtimes_u W$ is nuclear. Also π_2 is nuclear as it equals $J(\,\cdot\,)J \circ \pi_1 \circ \gamma \circ J(\,\cdot\,)J$. It therefore follows that the mapping $\pi_1 \otimes \pi_2 : (C(\partial W) \rtimes_r W) \otimes_{\min} JA_qJ \to \mathcal{B}(\ell^2(W))/\mathcal{K}$ in diagram (5-3) is nuclear and we may apply the Choi–Effros lifting theorem [1976] in order to obtain a ucp lift $\theta : (C(\partial W) \rtimes_r W) \otimes_{\min} JA_qJ \to \mathcal{B}(\ell^2(W))$. Then $\theta \circ (\gamma \otimes \mathrm{Id})$ together with (5-4) witness the result. □

Corollary 5.5. *Let (W, S) be an irreducible hyperbolic Coxeter system with $|S| \geq 3$ and $q \in [\rho, \rho^{-1}]$. Then the Hecke von Neumann algebra $\mathcal{M}_q$ has no Cartan subalgebra.*

Proof. This is a consequence of Theorem 5.3 together with Theorems 3.6, 4.13, and 5.4. □

Remark 5.6. In case W is not hyperbolic, it is not necessarily true that the group von Neumann algebra $\mathcal{M}_1$ is strongly solid. The easiest case is when Γ is $K_{2,3}$, the complete bipartite graph with $2 + 3$ vertices. Then the graph product $W = *_{K_{2,3}}\mathbb{Z}_2 = (\mathbb{Z}_2 * \mathbb{Z}_2) \times (\mathbb{Z}_2 * \mathbb{Z}_2 * \mathbb{Z}_2)$ contains a copy of $\mathbb{Z} \times \mathbb{F}_2$. Thus $\mathcal{M}_1$ cannot be strongly solid as it contains the group von Neumann algebra of $\mathbb{Z} \times \mathbb{F}_2$. Note that $K_{2,3}$ is not an irreducible graph but the same argument applies if one adds one point with no edges to $K_{2,3}$.

6. Absence of Cartan subalgebras

As we saw in Remark 5.6 the absence of Cartan subalgebras for general right-angled Hecke von Neumann algebras cannot be proved through strong solidity. In this section we obtain absence of Cartan subalgebras for some additional Hecke von Neumann algebras through an analysis of amalgamated free products in conjunction with [Vaes 2014, Theorem A] (see also [Ioana 2015] for related results). We need some terminology first.

Definition 6.1. Let $\mathcal{N}, \mathcal{P} \subseteq \mathcal{M}$ be finite von Neumann algebras. We say that $\mathcal{N}$ is injective (or amenable) relative to $\mathcal{P}$ if there is a completely positive map Φ from the basic construction $\langle \mathcal{M}, e_{\mathcal{P}} \rangle$ onto $\mathcal{N}$ such that $\Phi|_{\mathcal{M}}$ is the conditional expectation of $\mathcal{M}$ onto $\mathcal{N}$. Here $e_{\mathcal{P}}$ is the Jones projection, i.e., the conditional expectation of $\mathcal{M}$ to $\mathcal{P}$ on the L^2-level.

The following Theorem 6.2 uses Popa's intertwining by bimodules technique. For us it suffices that for finite (separable) von Neumann algebras $\mathcal{N}, \mathcal{P} \subseteq \mathcal{M}$ we say that $\mathcal{N} \prec_{\mathcal{M}} \mathcal{P}$ if there exists no sequence of unitaries w_k in $\mathcal{N}$ such that for all $x, y \in \mathcal{M}$ we have $\|\mathcal{E}_{\mathcal{P}}(x w_k y)\|_2 \to 0$. The following theorem is a somewhat less general version of [Vaes 2014, Theorem A].

Theorem 6.2. *Let $\mathcal{N}_i$, $i = 1, 2$, be finite von Neumann algebras with common von Neumann subalgebra $\mathcal{B}$. Let $\mathcal{N} = \mathcal{N}_1 *_{\mathcal{B}} \mathcal{N}_2$ be the (tracial) amalgamated free product. Let $\mathcal{A} \subseteq \mathcal{N}$ be a von Neumann subalgebra that is injective relative to one of the $\mathcal{N}_i$, $i = 1, 2$. Then at least one of the following statements holds true:*

(1) $\mathcal{A} \prec_{\mathcal{N}} \mathcal{B}$.

(2) *There exists i such that* $\mathrm{Nor}_{\mathcal{N}}(\mathcal{A}) \prec_{\mathcal{N}} \mathcal{N}_i$.

(3) $\mathrm{Nor}_{\mathcal{N}}(\mathcal{A})$ *is injective relative to $\mathcal{B}$.*

Recall that for a graph Γ and $r \in V\Gamma$ we have $\mathrm{Link}(r) = \{s \in V\Gamma : (r, s) \in E\Gamma\}$ and $\mathrm{Star}(r) = \mathrm{Link}(r) \cup \{r\}$. We include the following lemma to show that part of the condition in Theorem 6.7 can always be achieved.

Lemma 6.3. *Every irreducible graph Γ with $|V\Gamma| \geq 3$ contains a vertex $r \in V\Gamma$ such that $V\Gamma - \mathrm{Star}(r)$ contains at least two points.*

Proof. Pick some random point $r \in V\Gamma$. We cannot have $\mathrm{Star}(r) = V\Gamma$ because then Γ would not be irreducible. So there is at least one point $w \in V\Gamma - \mathrm{Star}(r)$. If there is another point in $V\Gamma - \mathrm{Star}(r)$ then we are done, so we assume that w is the only point in $V\Gamma - \mathrm{Star}(r)$. This implies that $\mathrm{Link}(r)$ is nonempty. $\mathrm{Star}(w)$ does not contain r as $w \notin \mathrm{Star}(r)$. Also there must be at least one point $u \in \mathrm{Link}(r)$ (which was nonempty!) that is not connected to w because if this is not the case then every two elements in $\mathrm{Link}(r)$ and $\{r, w\}$ would be connected so that Γ is not irreducible. In all we proved that w has the property that $V\Gamma - \mathrm{Star}(w)$ contains at least two elements, namely r and u. □

We recall the following definitions from [Caspers and Fima 2017].

Definition 6.4. Let Γ be a graph and let $w = w_1 \cdots w_n$ be a word with letters in $V\Gamma$. Suppose that $w_i = w_j$. We say that the i-th and j-th letters of w are separated if there is a k with $i < k < j$ such that $w_k \notin \mathrm{Star}(w_i)$. If every two (equal) letters in w are separated then w is called *reduced*.

Definition 6.5. Let Γ be a graph and for $s \in V\Gamma$ let $\mathcal{M}(s)$ be a von Neumann algebra with normal faithful tracial state τ_s. Let $\mathcal{M}(s)^\circ = \{a \in \mathcal{M}(s) : \tau_s(a) = 0\}$. Let $a = a_1 \cdots a_n$ with $a_i \in \mathcal{M}(s_i)^\circ$ be an operator in the graph product von Neumann algebra $\star_{s \in V\Gamma}\mathcal{M}(s)$. Then a is called *reduced* if the word $s_1 \cdots s_n$ is reduced. The word $s_1 \cdots s_n$ is then called the *type* of a. We also say that two operators a_i and a_j of the same type $s \in V\Gamma$ are separated if there exists $i < k < j$ such that the type of a_k is not in $\mathsf{Star}(s)$.

Definition 6.6. An inclusion of tracial von Neumann algebras $\mathcal{B} \subseteq \mathcal{N}$ is called mixing if for every sequence b_n in $\mathcal{B}$ with $\|b_n\| \leq 1$ and $b_n \to 0$ weakly we have that $\|\mathcal{E}_\mathcal{B}(xb_ny)\|_2 \to 0$ for all $x, y \in \mathcal{N} \ominus \mathcal{B}$.

For the proof of the following theorem we need a condition assuming the existence of a specific point $r \in S$. The condition is chosen such that in Claim 2 of the proof of Theorem 6.7 we get a mixing inclusion of von Neumann algebras. This gives examples of Hecke von Neumann algebras of nonhyperbolic Coxeter groups that do not possess Cartan subalgebras. Indeed examples can easily be constructed; for example if there exists a point $r \in S$ such that $\mathsf{Link}(r)$ is the graph of a nonhyperbolic Coxeter group and if there are few edges between $\mathsf{Link}(r)$ and $V\Gamma - \mathsf{Star}(r)$ (i.e., such that the condition below is satisfied). Though we believe that the theorem should hold without this condition we were unable to find a complete proof.

Theorem 6.7. *Let (W, S) be an irreducible right-angled Coxeter group with $|S| \geq 3$. Let $q \in [\rho, \rho^{-1}]$. Assume that there is an element $r \in S$ such that:*

- *$V\Gamma - \mathsf{Star}(r)$ contains at least two points.*
- *For every $s, t \in \mathsf{Link}(r)$ such that $(s, t) \notin E\Gamma$ we have that*

$$\mathsf{Link}(s) \cap \mathsf{Link}(t) \cap (V\Gamma - \mathsf{Star}(r)) = \varnothing.$$

Then the Hecke-von Neumann algebra $\mathcal{M}_q$ does not have a Cartan subalgebra.

Proof. Let $\Gamma = (V\Gamma, E\Gamma)$ be the graph of (W, S). By Corollary 3.4 we get a graph product decomposition $\mathcal{M}_q = \star_{s \in V\Gamma}\mathcal{M}_q(s)$ with $\mathcal{M}_q(s)$ the Hecke-von Neumann algebra associated with the Coxeter subsystem generated by just s (so it is 2-dimensional by Section 3). Choose $r \in V\Gamma$ satisfying the conditions of the statement of the theorem. Put

$$\mathcal{N}_1 = \star_{s \in \mathsf{Star}(r)}\mathcal{M}_q(s), \quad \mathcal{N}_2 = \star_{s \in V\Gamma - \{r\}}\mathcal{M}_q(s), \quad \text{and} \quad \mathcal{B} = \star_{s \in \mathsf{Link}(r)}\mathcal{M}_q(s).$$

Here $\mathsf{Link}(r)$, $\mathsf{Star}(r)$, and $V\Gamma - \{r\}$ are all viewed as full subgraphs of Γ, i.e., a subgraph for which two vertices share an edge if and only if they share an edge in Γ. Simply write $\mathcal{M}$ for $\mathcal{M}_q$. By [Caspers and Fima 2017, Theorem 2.26] we get

$$\mathcal{M} = \mathcal{N}_1 *_\mathcal{B} \mathcal{N}_2.$$

Now suppose that $\mathcal{A} \subseteq \mathcal{M}$ is a Cartan subalgebra. We are going to derive a contradiction by showing that any of the three alternatives of Theorem 6.2 is absurd.

Claim 1: We cannot have $\mathsf{Nor}_\mathcal{M}(\mathcal{A}) \prec_\mathcal{M} \mathcal{N}_i$ for either $i = 1, 2$.

Proof of the claim. As $\mathcal{A}$ is assumed to be Cartan we need to prove that $\mathcal{M} \not\prec_\mathcal{M} \mathcal{N}_i$. Let $t \in V\Gamma - \mathsf{Star}(r)$. Then the subalgebra of $\mathcal{M}$ generated by $\mathcal{M}_q(r)$ and $\mathcal{M}_q(t)$ is the tracial free product $\mathcal{M}_q(r) * \mathcal{M}_q(t)$.

Take unitaries $u \in \mathcal{M}_q(r)$ and $v \in \mathcal{M}_q(t)$ with trace 0. Put $w_k = (uv)^k$, which then is a unitary in $\mathcal{M}_q(r) * \mathcal{M}_q(t)$ with trace 0.

We need to show that for all $x, y \in \mathcal{M}$ we have $\|\mathcal{E}_{\mathcal{N}_i}(xw_k y)\|_2 \to 0$. Recall that $\mathcal{M}_q(s)^\circ$ is the space of elements $z \in \mathcal{M}_q(s)$ with trace 0. By a density argument we may and will assume that $x = x_1 \cdots x_k$ and $y = y_1 \cdots y_l$ are reduced operators with $x_i, y_i \in \mathcal{M}_q(s)^\circ$ for some s (see Definition 6.5 or [Caspers and Fima 2017, Definition 2.10] for the notion of reduced operators). Take a decomposition $x = x'a$, where $x' = x_1 \cdots x_m$ and $a = x_{m+1} \cdots x_k$, with $x_{m+1}, \dots, x_k \in \mathcal{M}_q(r)^\circ \cup \mathcal{M}_q(t)^\circ$. We may assume that this decomposition is taken in such a way that the length of a is maximal; in other words, the end of the expression x' has (after possible commutations) no factors x_i that come from $\mathcal{M}_q(r)^\circ$ and $\mathcal{M}_q(t)^\circ$. We take a similar decomposition for y. We may write $y = by'$ with $y' = y_{n+1} \cdots y_l$ and $b = y_1 \cdots y_n$ with y_i, $1 \le i \le n$, elements of either $\mathcal{M}_q(r)^\circ$ or $\mathcal{M}_q(t)^\circ$. Again we may assume that this decomposition is maximal meaning that (after possible commutations) the expression y' does not have factors at the start that come from either $\mathcal{M}_q(r)^\circ$ or $\mathcal{M}_q(t)^\circ$.

Now write $xw_k y = x'(aw_k b)y'$. For k big (in fact $k \ge m+n+1$ suffices) we get that $aw_k b$ is not contained in $\mathcal{N}_i$ for $i = 1, 2$. Indeed a and b can never cancel all the occurrences of u and v in $w_k = (uv)^k$ so that $aw_k b \in \mathcal{M}_q(r) * \mathcal{M}_q(t) \ominus (\mathcal{M}_q(r) \cup \mathcal{M}_q(t))$. So $xw_k y = x'(aw_k b)y' \notin \mathcal{N}_i$ for either $i = 1, 2$. Therefore $\|\mathcal{E}_{\mathcal{N}_i}(xw_k y)\|_2 \to 0$ as $k \to \infty$.

Claim 2: We do not have $\mathcal{A} \prec_{\mathcal{M}} \mathcal{B}$.

Proof of the claim. Firstly we check that the inclusion $\mathcal{B} \subseteq \mathcal{N}_2$ is mixing. Let b_n be a sequence in $\mathcal{B}$ with $\|b_n\| \le 1$ such that $b_n \to 0$ weakly. Take $x, y \in \mathcal{N}_2 \ominus \mathcal{B}$. By linearity and density we may assume that both x and y are reduced operators. In particular write a reduced expression $x = x_1 \cdots x_n$ with $x_i \in \mathcal{M}_q(s_i)^\circ$ for some $s_i \in V\Gamma - \{r\}$ and $1 \le i \le n$. Since x is not in $\mathcal{B}$ let x_{i_0} be such that $s_{i_0} \notin \mathsf{Link}(r)$. Let $V\Lambda$ be the set of all vertices in $\mathsf{Link}(r)$ that share an edge with s_{i_0}. Let Λ be the full subgraph of Γ with edge set $V\Lambda$. Then Λ must be complete (i.e., every two vertices share an edge) because otherwise this would contradict the assumptions on r. This means that $\widetilde{\mathcal{B}} := \star_{s \in V\Lambda} \mathcal{M}_q(s) = \bigotimes_{s \in V\Gamma} \mathcal{M}_q(s)$ is finite-dimensional, as $\mathcal{M}_q(s)$ is 2-dimensional; see Section 3. This in turn implies that $\|\mathcal{E}_{\widetilde{\mathcal{B}}}(b_n)\|_2 \to 0$ (indeed, b_n is bounded and converges to 0 weakly, hence σ-weakly; so $\mathcal{E}_{\widetilde{\mathcal{B}}}(b_n) \to 0$ σ-weakly and hence in the $\|\cdot\|_2$-norm, by finite dimensionality). Now we have

$$\mathcal{E}_{\mathcal{B}}(xb_n y) = \mathcal{E}_{\mathcal{B}}(x(b_n - \mathcal{E}_{\widetilde{\mathcal{B}}}(b_n))y) + \mathcal{E}_{\mathcal{B}}(x\mathcal{E}_{\widetilde{\mathcal{B}}}(b_n)y),$$

where the second summand converges to 0 in the $\|\cdot\|_2$-norm as $n \to \infty$. Further $\mathcal{E}_{\mathcal{B}}(x(b_n - \mathcal{E}_{\widetilde{\mathcal{B}}}(b_n))y) = 0$ for every n as the operator $x_{s_{i_0}}$ is separated from any other operator of type s_{i_0}. So this shows that

$$\|\mathcal{E}_{\mathcal{B}}(xb_n y)\|_2 = \|\mathcal{E}_{\mathcal{B}}(x\mathcal{E}_{\widetilde{\mathcal{B}}}(b_n)y)\|_2 \to 0.$$

This concludes our claim that the inclusion $\mathcal{B} \subseteq \mathcal{N}_2$ is mixing.

If $\mathcal{A} \prec_{\mathcal{M}} \mathcal{B}$ then we certainly have $\mathcal{A} \prec_{\mathcal{M}} \mathcal{N}_2$. But then by [Ioana 2015, Lemma 9.4] and the previous paragraph which shows that the inclusion $\mathcal{N}_i \subseteq \mathcal{M}$ is mixing, we get that also $\mathsf{Nor}_{\mathcal{M}}(\mathcal{A}) \prec_{\mathcal{M}} \mathcal{N}_2$. However this is impossible by Claim 1.

Claim 3: $\mathcal{M}$ is not relatively injective with respect to $\mathcal{B}$.

Proof of the claim. Recall our choice of $r \in V\Gamma$ at the start of the proof. Let t_1, t_2 be two different points in $V\Gamma - \mathsf{Star}(r)$. Let Λ be the full subgraph of Γ with vertex set $\{r, t_1, t_2\}$. Let $\mathcal{N} = \star_{s \in V\Lambda} \mathcal{M}_q(s)$. Note that $\mathcal{N} \cap \mathcal{B} = \mathbb{C}$. Suppose that $\mathcal{M}$ were to be relatively injective with respect to $\mathcal{B}$. Then there exists a (possibly nonnormal) conditional expectation $\Phi : \langle \mathcal{M}, e_{\mathcal{B}} \rangle \to \mathcal{M}$. We shall prove that this implies that $\mathcal{N}$ is injective.

Let A be the set of all reduced words $\boldsymbol{w}$ with letters in $V\Gamma$ that do not end on letters in $\mathsf{Link}(r)$ and that do not start with letters in $\{r, t_1, t_2\}$, meaning that for each $s \in \mathsf{Link}(r)$ the word $\boldsymbol{w}s$ is reduced and for each $s \in \{r, t_1, t_2\}$ the word $s\boldsymbol{w}$ is reduced. For each word $\boldsymbol{w} \in W$ let $X_{\boldsymbol{w}}$ be a maximal set of reduced operators in $\mathcal{M}$ of type $\boldsymbol{w}$ that form an orthonormal system in $L^2(\mathcal{M})$. Let $x \in X_{\boldsymbol{w}},\ x' \in X_{\boldsymbol{w}'}$ with $\boldsymbol{w}, \boldsymbol{w}' \in A$ and $x \neq x'$. The spaces spanned by $\mathcal{N} x \mathcal{B}$ and $\mathcal{N} x' \mathcal{B}$ are orthogonal in $L^2(\mathcal{M})$ and invariant subspaces for $\mathcal{N}$. Moreover, the projection[4] of $L^2(\mathcal{M})$ onto $\overline{\mathrm{span}\ \mathcal{N} x \mathcal{B}}^{\|\cdot\|_2}$ is given by

$$p_x = \sum_{i \in I} n_i x e_{\mathcal{B}} x^* n_i^*,$$

where we have chosen $n_i,\ i \in I$, to be elements of $\mathcal{N}$ that form an orthonormal basis of $L^2(\mathcal{N})$. In particular $p_x \in \langle \mathcal{M}, e_{\mathcal{B}} \rangle$. We have that the projections $p_x,\ x \in X_{\boldsymbol{w}},\ \boldsymbol{w} \in A$, commute with $\mathcal{N}$ and they sum up to 1 as

$$L^2(\mathcal{M}) = \bigoplus_{\boldsymbol{w} \in A, x \in X_{\boldsymbol{w}}} \overline{\mathrm{span}\ \mathcal{N} x \mathcal{B}}^{\|\cdot\|_2}.$$

For $\boldsymbol{w} \in A,\ x \in X_{\boldsymbol{w}}$ set

$$p'_x = x e_{\mathcal{B}} x^*.$$

Similarly, p'_x is the projection onto $\overline{\mathrm{span}\ x\mathcal{B}}^{\|\cdot\|_2}$ and $p'_x \leq p_x$. We claim that the von Neumann algebra generated by $p_x \mathcal{N} p_x$ and p'_x is homogeneous of type I. In order to do so note that there is a unitary[5] map

$$U_x : \overline{\mathrm{span}\ \mathcal{N} x \mathcal{B}}^{\|\cdot\|_2} \to L^2(\mathcal{N}) \otimes L^2(\mathcal{B}) : nxb \mapsto n \otimes b.$$

We have $U_x n U_x^* = n \otimes \mathrm{Id}_{L^2(\mathcal{B})}$ and $U_x p'_x U_x^* = p_\Omega \otimes \mathrm{Id}_{L^2(\mathcal{B})}$, where p_Ω is the projection onto $\Omega := 1_{\mathcal{N}} \in L^2(\mathcal{N})$. So that the von Neumann algebra $U_x \langle p_x \mathcal{N} p_x, p'_x \rangle U_x^*$ is isomorphic to $\mathcal{B}(L^2(\mathcal{N})) \otimes \mathrm{Id}_{L^2(\mathcal{B})}$, which is homogeneous of type I.

Now consider $\Psi : \langle \mathcal{M}, e_{\mathcal{B}} \rangle \to^{\Phi} \mathcal{M} \to^{\mathcal{E}_{\mathcal{N}}} \mathcal{N}$. This is a conditional expectation for the inclusion $\mathcal{N} \to \langle \mathcal{M}, e_{\mathcal{B}} \rangle$. Let $\mathcal{P}$ be the subalgebra of $\langle \mathcal{M}, e_{\mathcal{B}} \rangle$ that is generated by all $p_x \mathcal{N} p_x$ and p'_x with $x \in X_{\boldsymbol{w}}$,

[4] Indeed p_x is a projection: clearly $p_x^* = p_x$. Further, by assumption on $x = x_1 \cdots x_k$ we have for $n \in \mathcal{N}$ that nx is a reduced operator. Take $b \in \mathcal{B}$ of trace 0. The word $n_i x b$ is then reduced. In order to determine the conditional expectation $\mathcal{E}_{\mathcal{B}}$ of $x^* n_i^* n_j x b$ one needs to write $x^* n_i^* n_j x b$ as a sum of reduced operators and delete all terms that are not in $\mathcal{B}$. But the only such terms are the ones where n_i^* annihilates n_j and where each x_i^* annihilates x_i. That is, $\mathcal{E}_{\mathcal{B}}(x^* n_i^* n_j x b) = \tau(n_i^* n_j)\tau(x^* x) b = \delta_{i,j} b$. Similarly, in order to determine $\mathcal{E}_{\mathcal{B}}(x^* n_i^* n_j x b)$ one writes $x^* n_i^* n_j x$ as a reduced expression and filters all operators that are in $\mathcal{B}$. Using that x does not end on letters in $\mathcal{B}$, this can only happen if n_i^* annihilates the letter n_j and each x_i^* annihilates x_i. That is, $\mathcal{E}_{\mathcal{B}}(x^* n_i^* n_j x) = \tau(n_i^* n_j)\tau(x^* x) = \delta_{i,j}$. So we conclude $\mathcal{E}_{\mathcal{B}}(x^* n_i^* n_j x b) = \delta_{i,j} b$ for any $b \in \mathcal{B}$. This gives $e_{\mathcal{B}} x^* n_i^* n_j x e_{\mathcal{B}} = \delta_{i,j} e_{\mathcal{B}}$. Then $p_x^2 = \sum_{i,j \in I} n_i x e_{\mathcal{B}} x^* n_i^* n_j x e_{\mathcal{B}} x^* n_j^* = \sum_{i \in I} n_i x e_{\mathcal{B}} x^* n_i^* = p_x$. The image of p_x is clearly contained in $\overline{\mathrm{span}\ \mathcal{N} x \mathcal{B}}^{\|\cdot\|_2}$. Finally for a vector $nxb,\ n \in \mathcal{N},\ b \in \mathcal{B}$, we have $p_x(nxb) = \sum_{i \in I} n_i x e_{\mathcal{B}} x^* n_i^* n x b = \sum_{i \in I} n_i x \tau(x^* x) \tau(n_i^* n) b = \sum_{i \in I} n_i x \tau(n_i^* n) b = nxb$.

[5] Indeed U_x is unitary as $\| \sum_i n_i x b_i \|_2^2 = \sum_{i,j} \tau(b_j^* x^* n_j^* n_i x b_i) = \sum_{i,j} \tau(n_j^* n_i)\tau(b_j^* b_i) = \| \sum_i n_i \otimes b_i \|_2^2$, where the second equality uses that $n_i x b_i$ is reduced by definition of x and that $\tau(x^* x) = 1$ as x had norm 1 in $L^2(\mathcal{M})$.

$\boldsymbol{w} \in A$. The previous paragraph shows that $\mathcal{P} = \bigoplus_{x \in X_{\boldsymbol{w}}, \boldsymbol{w} \in A} \langle p_x \mathcal{N} p_x, p'_x \rangle$ is homogeneous of type I. Restricting Ψ to $\mathcal{P}$ gives a conditional expectation for the inclusion $\mathcal{N} \to \mathcal{P}$ (recall that $\mathcal{N}$ is contained in $\mathcal{P}$ as the projections p_x sum up to 1). Hence $\mathcal{N}$ is an expected subalgebra of a homogeneous type-I algebra. As homogeneous type-I algebras are expected subalgebras of a type-I factor we conclude that $\mathcal{N}$ is injective.

Remainder of the proof. Now Theorem 6.2 implies that either (1) $\mathrm{Nor}_{\mathcal{M}}(\mathcal{A}) \prec_{\mathcal{M}} \mathcal{N}_i$ for either $i = 1$ or $i = 2$; (2) $\mathcal{A} \prec_{\mathcal{M}} \mathcal{B}$; (3) $\mathcal{M}$ is injective relative to $\mathcal{B}$. The three claims above rule out all of these possibilities, showing that $\mathcal{M}$ does not possess a Cartan subalgebra. □

Acknowledgements

The author wishes to express his gratitude to the following people: Sergey Neshveyev, Łukasz Garncarek and Adam Skalski for enlightening discussions on Hecke von Neumann algebras, and the anonymous referees for several comments that led to significant improvements of the paper.

References

[Bożejko and Fendler 1984] M. Bożejko and G. Fendler, "Herz–Schur multipliers and completely bounded multipliers of the Fourier algebra of a locally compact group", *Boll. Un. Mat. Ital. A* (6) **3**:2 (1984), 297–302. MR Zbl

[Bożejko and Speicher 1994] M. Bożejko and R. Speicher, "Completely positive maps on Coxeter groups, deformed commutation relations, and operator spaces", *Math. Ann.* **300**:1 (1994), 97–120. MR Zbl

[Brown and Ozawa 2008] N. P. Brown and N. Ozawa, *C^*-algebras and finite-dimensional approximations*, Graduate Studies in Math. **88**, Amer. Math. Soc., Providence, RI, 2008. MR Zbl

[Caspers and Fima 2017] M. Caspers and P. Fima, "Graph products of operator algebras", *J. Noncommut. Geom.* **11**:1 (2017), 367–411. MR Zbl

[Chifan and Sinclair 2013] I. Chifan and T. Sinclair, "On the structural theory of II_1 factors of negatively curved groups", *Ann. Sci. Éc. Norm. Sup.* (4) **46**:1 (2013), 1–33. MR Zbl

[Chifan et al. 2013] I. Chifan, T. Sinclair, and B. Udrea, "On the structural theory of II_1 factors of negatively curved groups, II: Actions by product groups", *Adv. Math.* **245** (2013), 208–236. MR Zbl

[Choi and Effros 1976] M. D. Choi and E. G. Effros, "The completely positive lifting problem for C^*-algebras", *Ann. of Math.* (2) **104**:3 (1976), 585–609. MR Zbl

[Connes 1976] A. Connes, "Classification of injective factors: cases II_1, II_∞, III_λ, $\lambda \neq 1$", *Ann. of Math.* (2) **104**:1 (1976), 73–115. MR Zbl

[Davis 2008] M. W. Davis, *The geometry and topology of Coxeter groups*, London Math. Soc. Monographs Series **32**, Princeton Univ. Press, 2008. MR Zbl

[Davis et al. 2007] M. W. Davis, J. Dymara, T. Januszkiewicz, and B. Okun, "Weighted L^2-cohomology of Coxeter groups", *Geom. Topol.* **11** (2007), 47–138. MR Zbl

[De Commer 2011] K. De Commer, "Galois objects and cocycle twisting for locally compact quantum groups", *J. Operator Theory* **66**:1 (2011), 59–106. MR Zbl

[Dymara 2006] J. Dymara, "Thin buildings", *Geom. Topol.* **10** (2006), 667–694. MR Zbl

[Effros and Ruan 2000] E. G. Effros and Z.-J. Ruan, *Operator spaces*, London Math. Soc. Monographs New Series **23**, Oxford Univ. Press, 2000. MR Zbl

[Garncarek 2016] Ł. Garncarek, "Factoriality of Hecke–von Neumann algebras of right-angled Coxeter groups", *J. Funct. Anal.* **270**:3 (2016), 1202–1219. MR Zbl

[Green 1990] E. R. Green, *Graph products of groups*, Ph.D. thesis, University of Leeds, 1990, Available at http://tinyurl.com/phdgreen.

[Haagerup 1978] U. Haagerup, "An example of a nonnuclear C^*-algebra, which has the metric approximation property", *Invent. Math.* **50**:3 (1978), 279–293. MR Zbl

[Haagerup and Kraus 1994] U. Haagerup and J. Kraus, "Approximation properties for group C^*-algebras and group von Neumann algebras", *Trans. Amer. Math. Soc.* **344**:2 (1994), 667–699. MR Zbl

[Higson and Guentner 2004] N. Higson and E. Guentner, "Group C^*-algebras and K-theory", pp. 137–251 in *Noncommutative geometry* (Martina Franca, Italy, 2000), edited by S. Doplicher and R. Longo, Lecture Notes in Math. **1831**, Springer, 2004. MR Zbl

[Ioana 2015] A. Ioana, "Cartan subalgebras of amalgamated free product II_1 factors", *Ann. Sci. Éc. Norm. Sup.* (4) **48**:1 (2015), 71–130. MR Zbl

[Isono 2015] Y. Isono, "Examples of factors which have no Cartan subalgebras", *Trans. Amer. Math. Soc.* **367**:11 (2015), 7917–7937. MR Zbl

[Jimbo 1986] M. Jimbo, "A q-analogue of $U(\mathfrak{gl}(N+1))$, Hecke algebra, and the Yang–Baxter equation", *Lett. Math. Phys.* **11**:3 (1986), 247–252. MR Zbl

[Jones 1985] V. F. R. Jones, "A polynomial invariant for knots via von Neumann algebras", *Bull. Amer. Math. Soc.* (*N.S.*) **12**:1 (1985), 103–111. MR Zbl

[Junge et al. 2009] M. Junge, M. Neufang, and Z.-J. Ruan, "A representation theorem for locally compact quantum groups", *Int. J. Math.* **20**:3 (2009), 377–400. MR Zbl

[Kassel 1995] C. Kassel, *Quantum groups*, Graduate Texts in Math. **155**, Springer, 1995. MR Zbl

[Ozawa 2008] N. Ozawa, "Weak amenability of hyperbolic groups", *Groups Geom. Dyn.* **2**:2 (2008), 271–280. MR Zbl

[Ozawa and Popa 2010] N. Ozawa and S. Popa, "On a class of II_1 factors with at most one Cartan subalgebra", *Ann. of Math.* (2) **172**:1 (2010), 713–749. MR Zbl

[Pisier 2003] G. Pisier, *Introduction to operator space theory*, London Math. Soc. Lecture Note Series **294**, Cambridge Univ. Press, 2003. MR Zbl

[Popa and Vaes 2014] S. Popa and S. Vaes, "Unique Cartan decomposition for II_1 factors arising from arbitrary actions of hyperbolic groups", *J. Reine Angew. Math.* **694** (2014), 215–239. MR Zbl

[Reckwerdt 2015] E. Reckwerdt, "Weak amenability is stable under graph products", preprint, 2015. arXiv

[Ricard and Xu 2006] É. Ricard and Q. Xu, "Khintchine type inequalities for reduced free products and applications", *J. Reine Angew. Math.* **599** (2006), 27–59. MR Zbl

[Strătilă and Zsidó 1979] Ş. Strătilă and L. Zsidó, *Lectures on von Neumann algebras*, Editura Academiei, Bucharest, 1979. MR Zbl

[Takesaki 1979] M. Takesaki, *Theory of operator algebras, I*, Springer, 1979. MR Zbl

[Takesaki 2003] M. Takesaki, *Theory of operator algebras, II*, Encycl. Math. Sci. **125**, Springer, 2003. MR Zbl

[Ueda 2011] Y. Ueda, "Factoriality, type classification and fullness for free product von Neumann algebras", *Adv. Math.* **228**:5 (2011), 2647–2671. MR Zbl

[Vaes 2001] S. Vaes, "The unitary implementation of a locally compact quantum group action", *J. Funct. Anal.* **180**:2 (2001), 426–480. MR Zbl

[Vaes 2014] S. Vaes, "Normalizers inside amalgamated free product von Neumann algebras", *Publ. Res. Inst. Math. Sci.* **50**:4 (2014), 695–721. MR Zbl

[Voiculescu 1996] D. Voiculescu, "The analogues of entropy and of Fisher's information measure in free probability theory, III: The absence of Cartan subalgebras", *Geom. Funct. Anal.* **6**:1 (1996), 172–199. MR Zbl

Received 10 Nov 2016. Revised 2 May 2018. Accepted 12 Feb 2019.

MARTIJN CASPERS: m.p.t.caspers@uu.nl
Utrecht University, Utrecht, Netherlands

ANALYSIS AND PDE
Vol. 13, No. 1, 2020
dx.doi.org/10.2140/apde.2020.13.29

A VECTOR FIELD METHOD FOR RADIATING BLACK HOLE SPACETIMES

JESÚS OLIVER AND JACOB STERBENZ

We develop a commuting vector field method for a general class of radiating spacetimes. The metrics we consider are modeled on those constructed from global nonlinear stability problems in general relativity, and our method provides sharp peeling estimates for solutions to both linear and (null form) nonlinear scalar fields.

1.	Introduction	29
2.	Formulas for commutators and multipliers	40
3.	Algebraic formulas involving Bondi coordinates	42
4.	Asymptotic estimates involving Bondi coordinates	43
5.	Proof of the weighted L^2 estimates for $k=0$	59
6.	Estimates for commutators	64
7.	L^∞ estimates	73
8.	Estimates for nonlinear problems	78
Appendix A.	Coordinates	83
Appendix B.	Local energy decay	87
Appendix C.	Hardy and trace inequalities	89
References		91

1. Introduction

In this paper we develop a sharp variant of Klainerman's vector field method for solutions to scalar wave equations on a generic class of asymptotically flat spacetimes. These are taken to be certain long-range perturbations of Minkowski space which enjoy a standard local energy decay assumption.

The first main consideration here is to place the minimal conditions on our metrics at null infinity which are compatible with gravitational radiation, but at the same time are also strong enough to provide full peeling estimates for scalar waves. Our conditions turn out to be natural even if one is only interested in stationary long-range perturbations of Minkowski space, because they highlight certain peeling properties of Lorentzian metrics at null infinity which appear to be necessary in order to produce estimates on the order of the classical Morawetz conformal energy.

The second main consideration here is to produce a collection of norms which are natural for studying nonlinear stability problems, at least when the quadratic part of nonlinearity enjoys a certain generalized null condition. In fact, we will produce a range of norms with weights that are also capable of handling

MSC2010: 35L10.
Keywords: wave equation, local energy decay, vector field method.

systems satisfying weaker "null conditions", so our setup should also be useful for a class of quasilinear stability problems where one assumes certain bounds on Ricci curvature; but we leave this to a subsequent work.

First we discuss the basic assumptions and results which are contained in this paper. Additional remarks and references with then follow.

1A. *Basic notation and metric assumptions.* When stating inequalities we will use $A \lesssim B$ to mean $A \leq CB$ for some fixed $C > 0$ independent of A, B. We also employ the index notation $A \lesssim_k B$ when $C = C(k)$ depends on some auxiliary parameter k (although we will not always use such notation when a dependency exists).

The setting for the paper is the following: We fix some compact $\mathcal{K} \subset \mathbb{R}^3$ with smooth boundary such that $\mathbb{R}^3 \setminus \mathcal{K}$ is connected. On the manifold $\mathcal{M} = [0, \infty) \times (\mathbb{R}^3 \setminus \mathcal{K})$ we suppose there is given a smooth Lorentzian metric $g_{\alpha\beta}$. We write (t, x) for rectangular coordinates $0 \leq t < \infty$ and $x \in \mathbb{R}^3 \setminus \mathcal{K}$. Note that we may assume $\mathcal{K} = \varnothing$. In the sequel we will also assume that the level sets $t = \text{const}$ are uniformly spacelike in the sense that $-C < g^{00} < -c$. In this paper $r = |x|$ with the Euclidean norm and $\langle x \rangle = (1 + |x|^2)^{\frac{1}{2}}$.

With respect to the coordinates (t, x) and Euclidean measure $dt\, dx$ we have L^p norms restricted to time slabs in $\mathcal{M}$ of the form $[0, T] \times (\mathbb{R}^3 \setminus \mathcal{K})$. These will be denoted by $L^p[0, T]$. On the time slice $t = \text{const}$ we denote by L^p_x the corresponding restricted norm, and we denote by $L^p_t L^q_x[0, T]$ the mixed norms. It will also be useful to employ various inhomogeneous Besov versions of these spaces. For example we define $\|\phi\|^p_{\ell^p_r L^q[0,T]} = \sum_{j \geq 0} \|\chi_j \phi\|^p_{L^q[0,T]}$, where χ_j is a spatial cutoff on scale $\langle x \rangle \approx 2^j$. Similar notation is used for dyadic summations with respect to the time variable t, and other auxiliary variables as well (such as the optical functions described below). We define fixed-time and spacetime Sobolev spaces which will be denoted by H^s_x and $H^s[0, T]$, with the convention that in both cases we consider *all* (t, x)-derivatives ∂^I for multiindex $|I| \leq s$.

In this work we make two basic assumptions about our spacetimes $(\mathcal{M}, g)$. The first is an asymptotic condition on the metric:

Definition 1.1 (asymptotically flat radiating spacetimes). Suppose $(\mathcal{M}, g)$ is given as above. Then we say $g_{\alpha\beta}$ is "outgoing radiating" if the following hold:

(I) (weak "optical function") On $\mathcal{M}$ there is defined a smooth $u = u(t, x)$ such that (u, x) forms a uniform set of coordinates in the sense that $C^{-1} < u_t < C$, and u has the symbol bounds

$$\|(\tau_- \partial_t)^i (\tau_x \tau_0 \partial_x)^J (\partial_t u - 1, \partial_i u + \omega^i)\|_{\ell^1_r L^\infty[0,\infty)} < \infty \quad \text{for all } (i, J) \in \mathbb{N} \times \mathbb{N}^3, \tag{1}$$

where $\tau_x = \langle x \rangle$, $\tau_- = \langle u \rangle$, $\tau_+ = \langle (t, x) \rangle$, $\tau_0 = \tau_- \tau_+^{-1}$, and where $\omega^i = x^i \tau_x^{-1}$. Henceforth we let $\partial = (\partial_t, \partial_x)$ denote the (t, x)-coordinate derivatives, and likewise $\partial^b = (\partial^b_u, \partial^b_x)$ will denote the (u, x)-coordinate derivatives.

(II) (outgoing radiation condition) First define symbol classes $\mathcal{Z}^k$ in terms of the seminorms (restricted to $t \in [0, \infty)$)

$$|||q|||_{k,N} := \sum_{i+|J| \leq N} \|\tau_0^{-k} (\tau_- \partial^b_u)^i (\tau_x \partial^b_x)^J q\|_{\ell^1_r L^\infty} + \sum_{i+|J| \leq N} \|\tau_0^{-k} (\tau_- \partial^b_u)^i (\tau_x \partial^b_x)^J q\|_{\ell^1_u \ell^1_r L^\infty(\frac{1}{2}t < r < 2t)}. \tag{2}$$

Then for the inverse metric $g^{\alpha\beta}$ in (u,x)-coordinates one has the symbol bounds

$$g^{\alpha\beta} - h^{\alpha\beta} \in \mathcal{Z}^0, \quad \sqrt{|g|}g^{iu} + \omega^i \in \mathcal{Z}^{\frac{1}{2}}, \quad g^{ui} - \omega^i\omega_j g^{uj} \in \mathcal{Z}^1, \quad g^{uu} \in \mathcal{Z}^2, \tag{3}$$

where

$$h^{uu} = 0, \quad h^{ui} = -\omega^i, \quad h^{ij} = \delta^{ij}. \tag{4}$$

A number of further remarks are in order concerning the previous definition. Proofs are provided in Appendix A.

Remark 1.2. An immediate consequence of (1) is that one has a uniformly bounded change of frame between (∂_t, ∂_x) and $(\partial_u^b, \partial_x^b)$. Specifically

$$\partial_\alpha^b = e_\alpha^\beta \partial_\beta, \quad \sum_{|I|\le k} |\partial^I e_\alpha^\beta| \lesssim_k 1, \qquad \partial_\alpha = f_\alpha^\beta \partial_\beta^b, \quad \sum_{|I|\le k} |(\partial^b)^I f_\alpha^\beta| \lesssim_k 1. \tag{5}$$

In particular one may use either (∂_t, ∂_x) or $(\partial_u^b, \partial_x^b)$ to define the Sobolev spaces $H^s[0,T]$ and H_x^s, and for this purpose we will use these frames interchangeably in the sequel. Note however that the frames (∂_t, ∂_x) and $(\partial_u^b, \partial_x^b)$ are not interchangeable in all contexts. See for instance Remark 1.12 below.

Remark 1.3. The condition (I) implies that $\tau_+^{-1}(u + \tau_x - t) = o_r(1)$, and together with (II) we have

$$\|(\tau_-\partial_t)^i (\tau_x\tau_0\partial_x)^J (g-\eta)\|_{\ell_r^1 L^\infty[0,\infty)} < \infty \quad \text{for all } (i,J) \in \mathbb{N}\times\mathbb{N}^3,$$

where $\eta = \operatorname{diag}(-1,1,1,1)$ is the Minkowski metric in (t,x)-coordinates. In particular g is weakly asymptotically flat in the sense that $\partial^J(g-\eta) = o_r(1)$ for all $J \in \mathbb{N}^4$, and away from the region $\langle t-r\rangle \ll \langle t+r\rangle$ this can be strengthened to $(t\partial_t)^i(\tau_x\partial)^J(g-\eta) = o_r(1)$ for all $(i,J) \in \mathbb{N}\times\mathbb{N}^4$.

Remark 1.4. Definition 1.1 allows us to replace u by $\tilde{u} = \chi u + (1-\chi)(t-\tau_x)$, where χ is a cutoff supported where $\langle t-r\rangle \ll \langle t+r\rangle$ with bounds $|(\tau_+\partial)^J\chi| \lesssim 1$. Thus, in the sequel we shall always assume that $u = t - \tau_x$ away from the "wave zone" $\langle t-r\rangle \ll \langle t+r\rangle$.

Remark 1.5. For metrics which are quasistationary and quasispherical in the sense that $g = g_0 + g_1$, where g_0 is spherical in (t,x)-coordinates and

$$\|\ln^2(1+\tau_x)(t\partial_t)^i(\tau_x\partial)^J(g_0-\eta)\|_{\ell_r^1 L^\infty[0,\infty)} + \|\tau_x(t\partial_t)^i(\tau_x\partial)^J g_1\|_{\ell_r^1 L^\infty[0,\infty)} < \infty \quad \text{for all } (i,J)\in\mathbb{N}\times\mathbb{N}^4,$$

it can be shown the conditions of Definition 1.1 hold. In particular this guarantees that our conditions hold for certain stationary long-range perturbations of the Kerr family of spacetimes. In general we leave as an open question to find natural (e.g., geometric) conditions on metrics which satisfy only

$$\|\tau_x^a(t\partial_t)^i(\tau_x\partial)^J(g-\eta)\|_{\ell_r^1 L^\infty[0,\infty)} < \infty$$

for some $0 \le a < 1$, which would guarantee the existence of an optical function such as in Definition 1.1.

Remark 1.6. The peeling conditions (3) represent sharp streamlined versions of the asymptotic estimates obtained for geometric quantities in the nonlinear stability of Minkowski space (see [Christodoulou and Klainerman 1993; Bieri 2010]), as well as asymptotically flat dynamical black hole spacetimes (see [Dafermos et al. 2013]). While it would take a bit of work to translate the connection/curvature bounds of

[Christodoulou and Klainerman 1993; Bieri 2010; Dafermos et al. 2013] into coordinate conditions such as (3), we expect that the latter are in fact far more general. We leave this aspect of our work to further investigations.

1B. ***The local energy decay assumption.*** The second main assumption of this paper concerns the behavior of solutions to the inhomogeneous scalar wave equation on $(\mathcal{M}, g)$. Following [Marzuola et al. 2010; Tataru and Tohaneanu 2011; Tataru 2013] we define the local energy decay norms.

Definition 1.7 ("classical" local energy decay norms). First set

$$\|\phi\|_{\mathrm{LE}[0,T]} = \|\tau_x^{-\frac{1}{2}}\phi\|_{\ell_r^\infty L^2[0,T]}, \qquad \|F\|_{\mathrm{LE}^*[0,T]} = \|\tau_x^{\frac{1}{2}} F\|_{\ell_r^1 L^2[0,T]}.$$

Next, for a fixed $R_0 \geq 1$ sufficiently large ($r < R_0$ may be assumed to contain $\mathcal{K}$), and integer $s \geq 0$, we define

$$\|\phi\|_{\mathrm{LE}^s_{\mathrm{class}}[0,T]} = \sum_{|J|\leq s} (\|\tau_x^{-1}\partial^J\phi\|_{\mathrm{LE}[0,T]} + \|\partial\partial^J\phi\|_{\mathrm{LE}[0,T]}), \tag{6a}$$

$$\|\phi\|_{\mathrm{WLE}^s_{\mathrm{class}}[0,T]} = \sum_{|J|\leq s} (\|\tau_x^{-1}\partial^J\phi\|_{\mathrm{LE}[0,T]} + \|\partial\partial^J\phi\|_{\mathrm{LE}(r>R_0)[0,T]}), \tag{6b}$$

$$\|F\|_{\mathrm{LE}^{*,s}[0,T]} = \sum_{|J|\leq s} \|\partial^J F\|_{\mathrm{LE}^*[0,T]}, \tag{6c}$$

$$\|F\|_{\mathrm{WLE}^{*,s}[0,T]} = \sum_{|J|\leq s} (\|\partial^J F\|_{\mathrm{LE}^*[0,T]} + \|\partial\partial^J F\|_{L^2(r<R_0)[0,T]}). \tag{6d}$$

In terms of these spaces the second main assumption of the paper is:

Assumption 1.8 (weak and stationary energy boundedness/decay estimates). For R_0 as in Definition 1.7 above and any $s \geq 0$ there hold the estimates

$$\sup_{0\leq t\leq T} \|\partial\phi(t)\|_{H_x^s} + \|\phi\|_{\mathrm{WLE}^s_{\mathrm{class}}[0,T]} \lesssim_s \|\partial\phi(0)\|_{H_x^s} + \|\Box_g\phi\|_{(\mathrm{WLE}^{*,s}+L_t^1 H_x^s)[0,T]}, \tag{7a}$$

$$\sup_{0\leq t\leq T} \|\partial\phi(t)\|_{H_x^s} + \|\phi\|_{\mathrm{LE}^s_{\mathrm{class}}[0,T]} \lesssim_s \|\partial\phi(0)\|_{H_x^s} + \|(\phi, \partial_t\phi)\|_{L^2\times H^s(r<R_0)[0,T]} + \|\Box_g\phi\|_{\mathrm{LE}^{*,s}[0,T]}. \tag{7b}$$

In practice it is useful to have a bound which does not include the low-frequency error on the right-hand side of (7b). This is found by concatenating (7a)–(7b), which gives

$$\sup_{0\leq t\leq T} \|\partial\phi(t)\|_{H_x^s} + \|\phi\|_{\mathrm{LE}^s_{\mathrm{class}}[0,T]} \lesssim_s \|\partial\phi(0)\|_{H_x^s} + \|\partial_t\phi\|_{H^s(r<R_0)[0,T]} + \|\Box_g\phi\|_{\mathrm{WLE}^{*,s}[0,T]}. \tag{8}$$

Before continuing we make a number of remarks about these assumptions.

Remark 1.9. Estimates of the form (7a) have a long history for both asymptotically flat nontrapping spacetimes, and for the Kerr family of metrics. For the case of black holes we refer the reader to [Blue and Sterbenz 2006; Dafermos and Rodnianski 2009; Marzuola et al. 2010; Tataru and Tohaneanu 2011; Dafermos et al. 2016] for more detailed discussions. In the sharp form needed for the present paper the estimate (7a) was proved by Marzuola, Metcalfe, Tataru, and Tohaneanu [Marzuola et al. 2010] for Schwarzschild space and Tataru and Tohaneanu [2011] for Kerr with $|a| \ll M$. More recently a slightly less precise version of (7a) was established for the full subextremal range of Kerr by Dafermos, Rodnianski, and Shlapentokh-Rothman [Dafermos et al. 2016]. It is expected that (7a) holds for a wide class of asymptotically flat spacetimes which satisfy certain natural structural conditions similar to those of the Kerr metric, as well as certain natural spectral assumptions. We refer the reader to [Metcalfe et al. 2017] for a definitive account in the case of nontrapping spacetimes.

Remark 1.10. The second estimate (7b) is essentially the "stationary local energy decay" estimate of [Metcalfe et al. 2012]. It can be shown that (7b) follows from estimates similar to (7a) when ∂_t is timelike on a set $\mathcal{T}$ where one loses regularity in the local energy decay norms (6). See Appendix B for a proof. We remark that such timelike conditions hold for the Kerr family of metrics when the angular momentum satisfies $0 \le |a| < a_0$ for some $0 < a_0 < M$. On the other hand one should still expect (7b) to hold for the full subextremal range $0 \le |a| < M$ of Kerr thanks to the (microlocal) nonvanishing of the symbol of ∂_t on the trapped set.

1C. ***Norms and vector fields.*** We now introduce the weighted local energy decay norms that will play a leading role in the remainder of the paper.

Definition 1.11 (null energy decay norms). First we define a null energy seminorm with the same scaling as LE:

$$\|\phi\|_{\mathrm{NLE}[0,T]} = \|\tau_-^{-\frac{1}{2}}\phi\|_{\ell_u^\infty L^2(\frac{1}{2}t<r<2t)[0,T]}, \qquad \|F\|_{\mathrm{NLE}^*[0,T]} = \|\tau_-^{\frac{1}{2}} F\|_{\ell_u^1 L^2(\frac{1}{2}t<r<2t)[0,T]}.$$

Next, we have the generalization of $\mathrm{LE}^s_{\mathrm{class}}$ and $\mathrm{WLE}^s_{\mathrm{class}}$ which include the seminorms

$$\|\phi\|_{\mathrm{LE}^s[0,T]} = \|\phi\|_{\mathrm{LE}^s_{\mathrm{class}}[0,T]} + \sum_{|J|\le s} \|(\partial_x^b(\partial^b)^J\phi,\, \tau_x^{-1}\partial^J\phi)\|_{\mathrm{NLE}(r>R_0)[0,T]}, \tag{9a}$$

$$\|\phi\|_{\mathrm{WLE}^s[0,T]} = \|\phi\|_{\mathrm{WLE}^s_{\mathrm{class}}[0,T]} + \sum_{|J|\le s} \|(\partial_x^b(\partial^b)^J\phi,\, \tau_x^{-1}\partial^J\phi)\|_{\mathrm{NLE}(r>R_0)[0,T]}. \tag{9b}$$

Remark 1.12. Notice we have specifically used Bondi coordinate derivatives $(\partial_u^b, \partial_x^b)$ inside the gradient portion of the right-hand side of (9). This appears to be necessary because condition (1) does not guarantee a good peeling estimate for $(\partial_x^b)^I\, \partial u$. In particular we don't assume improved control of the commutators $[\partial_x^b, \partial_\alpha]$ beyond the bounds (5).

Remark 1.13. We remark that our definition of NLE is essentially a sharp (dyadic) version of Alinhac's "ghost weight" energy [2001]. The authors wish to thank the anonymous referee for pointing out this connection. Note also that the NLE terms in (9a) and (9b) simply represent a weaker (averaged) version of the natural outgoing null energy on the hypersurfaces $u = \mathrm{const}$.

Next, we define LE-type norms with uniform weights in time. These will play a central role in establishing peeling estimates for solutions to the wave equation.

Definition 1.14 (weighted LE and conformal energy norms). For functions ϕ which solve the wave equation and $0 \leq a \leq 1$ we set

$$\|\phi(t)\|_{E^a} = \|\tau_+^a \tau_0^{\max\{a,\frac{1}{2}\}} \partial\phi(t)\|_{L_x^2} + \|\tau_+^a(\partial_x^b\phi(t),\, \tau_x^{-1}\phi(t))\|_{L_x^2}, \tag{10a}$$

$$\|\phi\|_{S^a[0,T]} = \|\tau_+^a \tau_0^{\max\{a,\frac{1}{2}\}} \partial\phi\|_{\mathrm{LE}[0,T]} + \|\tau_+^a(\partial_x^b\phi,\, \tau_x^{-1}\phi)\|_{\mathrm{LE}[0,T]} + \|\tau_+^a \tau_x^{-1}\partial_r^b(\tau_x\phi)\|_{\mathrm{NLE}[0,T]}, \tag{10b}$$

$$\|\phi\|_{S^{1,\infty}[0,T]} = \|\phi\|_{\ell_t^\infty S^1[0,T]} + \|\tau_+ \tau_x^{-1}\partial_r^b(\tau_x\phi)\|_{\mathrm{NLE}[0,T]}. \tag{10c}$$

For source terms we fix a parameter R_0 as in Definition 1.7 and set

$$\|F\|_{N^a[0,T]} = \|\tau_+^a \tau_0^{\frac{1}{2}} F\|_{\mathrm{LE}^*[0,T]} + \|\tau_+^a \partial F\|_{L^2(r<R_0)[0,T]} + \|\tau_+^a F\|_{\mathrm{NLE}^*[0,T]}, \tag{11a}$$

$$\|F\|_{N^{1,1}[0,T]} = \|\tau_+ \tau_0^{\frac{1}{2}} F\|_{\ell_t^1 \mathrm{LE}^*[0,T]} + \|\tau_+ \partial F\|_{\ell_t^1 L^2(r<R_0)[0,T]} + \|\tau_+ F\|_{\mathrm{NLE}^*[0,T]}. \tag{11b}$$

Finally, we construct higher-order versions of all the above norms.

Definition 1.15 (modified vector fields). First define the approximate Lie algebras

$$\mathbb{L}_0 = \{\partial_u^b,\, \partial_i^b - \omega^i \partial_u^b\}, \quad \mathbb{L} = \{S,\, \Omega_{ij}\} \cup \mathbb{L}_0, \quad \text{where } S = u\partial_u^b + r\partial_r^b,\ \Omega_{ij} = x^i\partial_j^b - x^j\partial_i^b. \tag{12}$$

Note that all members of $\mathbb{L}$ commute modulo $\mathbb{L}$ with the exception of

$$[\partial_i^b - \omega^i\partial_u^b,\, S] = \partial_i^b - \omega^i\partial_u^b + \tau_x^{-2}\omega^i\partial_u^b. \tag{13}$$

For a function ϕ we write

$$\phi^{(k)} = (\phi,\, \Gamma^{I_1}\phi,\, \Gamma^{I_2}\phi,\, \dots),$$

where the right-hand side is an array of all products Γ^I of vector fields in $\mathbb{L}$ up to length $|I| \leq k$. If $\|\cdot\|$ is any norm we write

$$\|\phi^{(k)}\| = \sum_{|I|\leq k} \|\Gamma^I\phi\|, \qquad \Gamma \in \mathbb{L}.$$

We use a similar notation for pointwise identities; for example $|\partial\phi^{(k)}| = \sum_{|I|\leq k} |\partial\Gamma^I\phi|$ etc. In the case of norms we use a subscript notation to denote higher-order derivatives by vector fields:

$$\|\phi(t)\|_{E_k^a} = \|\phi^{(k)}(t)\|_{E^a}, \quad \|\phi\|_{S_k^a[0,T]} = \|\phi^{(k)}\|_{S^a[0,T]}, \quad \|F\|_{N_k^a[0,T]} = \|F^{(k)}\|_{N^a[0,T]},$$

and similarly for LE-type norms. In addition we set $\|\phi\|_{H_k^s} = \|\phi^{(k)}\|_{H^s}$ and $\|\phi(t)\|_{H_{x,k}^s} = \|\phi^{(k)}(t)\|_{H_x^s}$.

1D. *Main results, I: Linear estimates.*

The main result of the paper can now be stated as follows.

Theorem 1.16 (weighted local energy decay estimates). *Assuming estimates (7a) and (7b), for $0 \leq a \leq 1$ and fixed $s, k \geq 0$ there exist parameters $B_a = B_a(s,k)$ such that:*

(I) *In the case $a = 0$ one has*

$$\sup_{0\leq t\leq T} \|\partial\phi(t)\|_{H_{x,k}^s} + \|\phi\|_{\mathrm{WLE}_k^s[0,T]}$$
$$\lesssim_{s,k} \sum_{|I|\leq k}\sum_{|J|\leq s} \|(\tau_x\partial_x)^I \partial_x^J \partial\phi(0)\|_{L_x^2} + \|\phi\|_{H_{k-1}^{s+3}(r<B_0)[0,T]} + \|\Box_g\phi\|_{(\mathrm{WLE}_k^{*,s} + L_t^1 H_{x,k}^s)[0,T]}, \tag{14}$$

where in the case $k = 0$ we define $\|\phi\|_{H_{-1}^{s+3}(r<B_0)[0,T]} = 0$.

(II) *For* $0 < a < 1$ *one has*

$$\sup_{0\le t\le T}\|\phi(t)\|_{E^a_k} + \|\phi\|_{S^a_k[0,T]} \lesssim_{k,a} \sum_{|J|\le k}\|\tau_x^a(\tau_x\partial_x)^J\partial\phi(0)\|_{L^2_x} + \|\tau_+^{a-1}\phi\|_{H^1_{k+1}(r<B_a)[0,T]} + \|\Box_g\phi\|_{N^a_k[0,T]}. \quad (15)$$

(III) *Corresponding to* $a = 1$ *one has the endpoint bound*

$$\sup_{0\le t\le T}\|\phi(t)\|_{E^1_k}+\|\phi\|_{S^{1,\infty}_k[0,T]} \lesssim_k \sum_{|J|\le k}\|\tau_x(\tau_x\partial_x)^J\partial\phi(0)\|_{L^2_x}+\|\phi\|_{\ell^1_t H^1_{k+1}(r<B_1)[0,T]}+\|\Box_g\phi\|_{N^{1,1}_k[0,T]}. \quad (16)$$

Remark 1.17. In the proof of Theorem 1.16 we find that for $a \neq 0, 1$ one has $B_a \to \infty$ as $a \to 0, 1$.

In addition to the estimates above we shall also prove the following analog of Klainerman's estimate:

Theorem 1.18 (global Sobolev estimate). *For* $k \ge 1$ *one has the estimate*

$$\sum_{i+|J|\le k}\|\tau_+^{\frac{3}{2}}\tau_0^{\frac{1}{2}}(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\phi\|_{L^\infty[0,T]} \lesssim_k \sup_{0\le t\le T}\|\phi(t)\|_{E^1_{k+1}} + \|\phi\|_{S^{1,\infty}_{k+2}[0,T]} + \sum_{|J|\le k}\|\tau_x^2(\tau_x\partial)^J\Box_g\phi(0)\|_{L^2_x} + \sum_{i+|J|\le k+1}\|(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\Box_g\phi\|_{N^{1,1}[0,T]}. \quad (17)$$

One can combine the above two results in a straightforward way, which produces the first main conclusion of our paper:

Theorem 1.19 (peeling estimates for the inhomogeneous wave equation). *Given any* $k \ge 1$ *there exists an integer* $N = N(k)$ *depending (linearly) on* k *such that*

$$\sum_{i+|J|\le k}\|\tau_+^{\frac{3}{2}}\tau_0^{\frac{1}{2}}(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\phi\|_{L^\infty[0,T]} \lesssim_k \sum_{|J|\le N}\|\tau_x(\tau_x\partial_x)^J\partial\phi(0)\|_{L^2_x}+\sum_{|J|\le N}\|\tau_x^2(\tau_x\partial)^J\Box_g\phi(0)\|_{L^2_x} + \sum_{i+|J|\le N}\|(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\Box_g\phi\|_{N^{1,1}[0,T]}. \quad (18)$$

Proof that Theorems 1.16 and 1.18 imply estimate (18). We will in fact prove this for the explicit value $N = N(k) = 3k + 13$. Expanding the vector fields $\mathbb{L}$ from (12) we easily see

$$\|\Box_g\phi\|_{N^{1,1}_k[0,T]} \lesssim_k \sum_{i+|J|\le k}\|(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\Box_g\phi\|_{N^{1,1}[0,T]}.$$

Therefore given a value of $k \ge 1$, by combining (17) and (16), it suffices to control $\|\phi\|_{\ell^1_t H^1_{k+3}(r<B_1)[0,T]}$. Note that for any $0 < a < 1$ one has

$$\|\phi\|_{\ell^1_t H^1_{k+3}(r<B_1)[0,T]} \lesssim_a \|\phi\|_{S^a_{k+3}[0,T]}.$$

In particular applying this for $a=\frac{1}{2}$ (say), and then using estimate (15) and $N^{1,1}[0,T]\subseteq N^{\frac{1}{2}}[0,T]$, it suffices to bound the right-hand side of

$$\|\tau_+^{-\frac{1}{2}}\phi\|_{H^1_{k+4}(r<B_{1/2})[0,T]}\lesssim\|\phi\|_{\mathrm{WLE}^1_{k+4}[0,T]}.$$

We now inductively use (14) in the form

$$\|\phi\|_{\mathrm{WLE}^{s_j}_{l_j}[0,T]}\lesssim_{s_j,l_j}\sum_{|I|\le l_j}\sum_{|J|\le s_j}\|(\tau_x\partial_x)^I\partial_x^J\partial\phi(0)\|_{L^2_x}+\|\phi\|_{\mathrm{WLE}^{s_j+3}_{l_j-1}[0,T]}+\|\Box_g\phi\|_{\mathrm{WLE}^{*,s_j}_{l_j}[0,T]},$$

where $l_1=k+4,\ l_2=k+3,\ \dots,\ l_{k+4}=1$ and $s_1=1,\ s_2=4,\ \dots,\ s_{k+4}=3k+13$. This leaves us with finally having to bound the quantity $\|\phi\|_{\mathrm{WLE}^{3k+13}_0[0,T]}$, which can be handled by an additional application of (14) with only initial and source data bounds on the right-hand side (or what amounts to the same thing by directly using our base assumption (7a)). This concludes our demonstration of (18) for the explicit value $N=N(k)=3k+13$ chosen above. □

1E. *Main results, II: Nonlinear estimates.* The estimates of Theorems 1.16 and 1.18 naturally lend themselves bounding solutions to semilinear wave equations of the form $\Box_g\phi=F(t,x,\phi,\partial\phi)$. Rather than develop a comprehensive theory we concentrate on the equations $\Box_g\phi=\mathcal{N}^{\alpha\beta}(t,x,\phi)\partial_\alpha\phi\partial_\beta\phi$, where the quadratic form $\mathcal{N}^{\alpha\beta}$ is sufficiently tame.

Definition 1.20 (generalized null forms). A 2-tensor $\mathcal{N}^{\alpha\beta}$ is called a "generalized null form" with respect to a (weak) optical function u if its Bondi coordinate components satisfy

$$|(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\partial_\phi^k\mathcal{N}^{\alpha\beta}|\lesssim_{i,J}c_k(|\phi|),\qquad |(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\partial_\phi^k\mathcal{N}^{uu}|\lesssim_{i,J}c_k(|\phi|)\tau_0. \tag{19}$$

Remark 1.21. Natural examples of $\mathcal{N}$ satisfying Definition 1.20 include multiples of the inverse metric $g^{\alpha\beta}$ by factors $\mathcal{N}(t,x,\phi)$ which satisfy derivative estimates consistent with (19). This would include the case of wave-maps from $\phi:(\mathcal{M},g)\to(\mathcal{M}',g')$ into Riemannian or Lorentzian targets where ϕ is close to a constant map, in either an intrinsic or extrinsic formulation.

Another example of such $\mathcal{N}$ would be skew symmetric forms $\mathcal{N}^{\alpha\beta}=-\mathcal{N}^{\beta\alpha}$ which obey the first condition in (19).

In order to prove a priori estimates we define the norms

$$\begin{aligned}|||\phi|||_{S_k[0,T]}=\sum_{j=0}^{k+4}\Big(\sup_{0\le t\le T}\|\partial\phi(t)\|_{H^{13+3(k-j)}_{x,j}}+\|\phi\|_{\mathrm{WLE}^{13+3(k-j)}_j[0,T]}\Big)+\|\phi\|_{S^{1/2}_{k+3}[0,T]}\\ +\sup_{0\le t\le T}\|\phi(t)\|_{E^1_{k+2}}+\|\phi\|_{S^{1,\infty}_{k+2}[0,T]}+\sum_{i+|J|\le k}\|\tau_+^{\frac{3}{2}}\tau_0^{\frac{1}{2}}(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\phi\|_{L^\infty[0,T]},\end{aligned} \tag{20}$$

$$\begin{aligned}|||F|||_{N_k[0,T]}=\sum_{j=0}^{k+4}\|F\|_{(\mathrm{WLE}^{*,13+3(k-j)}_j+L^1_tH^{13+3(k-j)}_{x,j})[0,T]}+\|F\|_{N^{1/2}_{k+3}[0,T]}\\ +\sum_{|J|\le k}\|\tau_x^2(\tau_x\partial)^JF(0)\|_{L^2_x}+\sum_{i+|J|\le k+2}\|(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^JF\|_{N^{1,1}[0,T]}.\end{aligned} \tag{21}$$

With this notation the main nonlinear theorem of our paper is the following:

Theorem 1.22. *Let ϕ be a smooth vector-valued function and let $\mathcal{N}(\phi, \partial\phi)$ denote a quadratic form obeying (19), where we are using the notation $\mathcal{N}(\phi, \partial\phi) = \mathcal{N}^{\alpha\beta}(t, x, \phi)\, \partial_\alpha\phi\, \partial_\beta\phi$. Then one has the bounds*

$$\|\!|\phi|\!\|_{S_k[0,T]} \lesssim_k \sum_{|J|\le 3k+13} \|\tau_x(\tau_x\partial_x)^J \partial\phi(0)\|_{L^2_x} + \|\!|\Box_g\phi|\!\|_{N_k[0,T]}, \tag{22}$$

$$\|\!|\mathcal{N}(\phi, \partial\phi)|\!\|_{N_k[0,T]} \lesssim C_k(\|\!|\phi|\!\|_{S_k[0,T]})\|\!|\phi|\!\|_{S_k[0,T]}(\|\!|\phi|\!\|_{S_k[0,T]} + \|\!|\Box_g\phi|\!\|_{N_k[0,T]}), \quad k \ge 18. \tag{23}$$

The functions $C_k(\,\cdot\,)$ are locally bounded functions determined by k as well as the functions c_k on right-hand side of (19).

From this and a standard continuity argument one has the nonlinear analog of Theorem 1.19:

Theorem 1.23 (peeling estimates for solutions to null-form systems). *Suppose quadratic forms $\mathcal{N}$ are given which satisfy (19). Let ϕ be a sufficiently smooth and well-localized solution to the system of semilinear equations*

$$\Box_g\phi = \mathcal{N}^{\alpha\beta}(t, x, \phi)\, \partial_\alpha\phi\, \partial_\beta\phi, \tag{24}$$

which is assumed to hold on a time interval $[0, T]$. Then given any $k \ge 18$ there exists an $\epsilon_0 = \epsilon_0(k) > 0$ such that one has the a priori estimate

$$\sum_{|J|\le 3k+13} \|\tau_x(\tau_x\partial_x)^J \partial\phi(0)\|_{L^2_x} = \epsilon \le \epsilon_0 \quad \Longrightarrow \quad \|\!|\phi|\!\|_{S_k[0,T]} \lesssim_k \epsilon.$$

In particular for sufficiently smooth, small, and well-localized initial data the solution to (24) exists globally and enjoys the peeling estimates

$$\sum_{i+|J|\le k} \|\tau_+^{\frac{3}{2}}\tau_0^{\frac{1}{2}}(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\phi\|_{L^\infty[0,T]} \lesssim_k \epsilon$$

for the same k as above.

1F. ***Some remarks and references.*** The vector field method for the wave equation on asymptotically flat spacetimes has a long and well-developed history. In the case of small perturbations of Minkowski space there are Klainerman's original works [1985; 1986], followed by the proof in [Christodoulou and Klainerman 1993] of the nonlinear stability of Minkowski space itself. In the latter work a vector field method is developed for radiating spacetimes where control is ultimately provided through certain peeling estimates for the curvature tensor of g. In a related vein there is [Bieri 2010] concerning nonlinear stability under the much less restrictive decay assumptions on the initial data. In this case the peeling properties of the metric end up being closer to the thresholds (3).

Going in a different direction there is the proof of stability of Minkowski space and its asymptotics in wave coordinates in [Lindblad and Rodnianski 2010; Lindblad 2017]. Here one proceeds more directly via Minkowski and Schwarzschild vector fields. For scalar wave equations this produces estimates with a small loss due to the divergence of radiating null hypersurfaces compared to their stationary counterparts (at least at the level of energy estimates). On the other hand strong peeling estimates such as (18) cannot hold for the (Minkowski difference of the) metric in wave coordinates due to obstructions at the level of semilinear terms. More specifically, condition (19) seems to fail when writing the Einstein equations in any reasonable way as a system of second-order equations for the metric.

Next, we turn to vector field methods on spacetimes which are locally large perturbations of Minkowski space. There are mainly two innovations here, and we make a heavy use of both in the sequel. The first innovation, due to Klainerman and Sideris [1996], allows one to replace Lorentz boosts with certain weighted identities involving the wave operator $\Box_g$ (specifically see Lemma 4.13 below). The second innovation, due initially to Keel, Smith, and Sogge [Keel et al. 2002] and used by many authors, concerns the use of local energy decay estimates such as (7) in order to control localized errors generated by commutations with vector fields. For further background and developments concerning the combination of local energy decay and vector fields on various large perturbations of Minkowski space, we refer the reader to [Metcalfe and Sogge 2005; 2007; Bony and Häfner 2010; Wang and Yu 2014; Yang 2013].

Concerning the application of vector fields to the class of black hole spacetimes there has recently been a great deal of progress. The works most closely related to the present paper are [Blue and Sterbenz 2006; Dafermos et al. 2016; Luk 2012; 2013], which proceed by way of conformal energy estimates and time-dependent weights. We also mention [Lindblad and Tohaneanu 2018] concerning certain quasilinear wave equations (in this case the vector field approach is similar to [Lindblad and Rodnianski 2010]). We remark that all of these works are written specifically to cover the case of Schwarzschild or Kerr with small angular momentum.

For more general "black hole" backgrounds there is [Moschidis 2016] building on ideas of [Dafermos and Rodnianski 2010] (see also Proposition 6.1 of [Rodnianski and Sterbenz 2010], where weighted estimates based on outgoing null vector fields were introduced). Here one produces a vector field method with time-independent weights. This shares a number of similarities with the present paper, in particular the production of a family of weighted spacetime energy estimates depending on a parameter $0 \leq a \leq 1$ which interpolates between the standard local energy decay estimates and the conformal energy. On the other hand our approach and the one of Dafermos and Rodnianski and Moschidis appear to diverge in many ways, especially regarding the issue of time-dependent weights. Our method also appears to be more directly applicable to studying nonlinear problems. Indeed, with the machinery of estimates (14)–(16) our proof of Theorem 1.22 occupies only a few additional pages.

Finally, one should also mention the works on Price's law for scalar waves [Tataru 2013; Metcalfe et al. 2012] which use vector fields as a launching point for much sharper estimates. An interesting open question in this regard is to find an appropriate collection of norms capable of producing interior decay rates better than $t^{-\frac{3}{2}}$, but which are still compatible with semilinear problems such as in Theorem 1.23.

We close with a few additional comments on the specific methods we employ in this work. These are a natural outgrowth of [Blue and Sterbenz 2006; Lindblad and Sterbenz 2006; Rodnianski and Sterbenz 2010; Oliver 2013; Dafermos and Rodnianski 2010]. In [Lindblad and Sterbenz 2006] we developed a conformal multiplier technique that works well for perturbations of the wave equation, and then used it to produce a collection of conformal energies with weights depending on a parameter. These ideas will again play a central role in the present work. In [Rodnianski and Sterbenz 2010] we introduced a Morawetz-type estimate based on the multiplier $f(r)(\partial_t + \partial_r)$ and used it to control null tangential derivatives for solutions to a certain wave equation with a potential (Proposition 6.1 of that paper). This idea was expanded in [Dafermos and Rodnianski 2010] to prove global decay estimates. We will use similar multipliers in this paper to produce a key portion of our estimates.[1] In [Blue and Sterbenz 2006] we produced a weighted-in-time local energy decay bound, and then used this to control the multiplier error from the conformal vector field. This technique is used again here for a wider range of weights, albeit assuming the local energy decay bound as a black box. Finally, the first author's thesis [Oliver 2013] and [Oliver 2016] developed vector fields on radiating nontrapping spacetimes with even weaker local bounds on $\partial_t g$. We will use much of the setup from [Oliver 2016] in the present work.

1G. ***Outline of the paper.*** In Section 2 we record a number of algebraic identities for energy momentum and deformation tensors. These form the basis for all the multiplier and commutator identities needed in the sequel.

In Section 3 we specialize the formulas of Section 2 to the case of Bondi coordinates satisfying Definition 1.1 and vector fields from Definition 1.15.

Section 4 is the technical heart of the paper. We begin by producing symbol bounds for the Lie derivatives of 2-tensors which satisfy certain natural asymptotic estimates. Building on this we construct a generic multiplier estimate for vector fields satisfying certain structural properties. This estimate covers all of the multiplier bounds needed in the sequel, and may be useful for other applications as well, which is one of our motivations for introducing an axiomatic setup. Following this we move on to estimates for commutators. This is done in a way that allows us to perform some delicate integration by parts later in Section 6, and is also convenient for proving pointwise bounds. We end by proving several generalized Klainerman–Sideris-type identities, and then use these to conclude the discussion of pointwise bounds for commutators.

In Section 5 we apply the abstract multiplier bound from Section 4 to produce the three main estimates of Theorem 1.16 at level $k = 0$. For the bounds (14) and (16) this is done in such a way that the source error term can be integrated by parts; this form is needed later to establish bounds for commutators.

In Section 6 we apply the formulas of the previous two sections to prove the estimates of Theorem 1.16 for an arbitrary number of vector fields. For the bound (15) this is more or less straightforward. However, for the bounds (14) and (16) the argument is more involved because one cannot proceed directly via Hölder's inequality, and several additional integrations seem necessary to close the argument.

In Section 7 we prove some sharp L^∞ decay estimates for functions in terms of the conformal energy norms (10a), (10c), and (11b). This establishes Theorem 1.18.

In Section 8 we prove Theorem 1.22, which concludes the main nonlinear application of the paper.

In Appendix A we provide proofs of a number of the remarks following Definition 1.1. In particular for large class of stationary metrics which are also spherically symmetric to highest order, we construct optical functions satisfying bounds (3). Such metrics include radial long-range perturbations of the Kerr family of metrics.

[1] For a complete the list of multipliers we use here, see Lemma 5.4 and the proofs immediately following.

In Appendix B we show that estimate (7a) implies estimate (7b), at least when the metric satisfies structural assumptions similar to the Kerr family with angular momentum in a certain range.

Finally, in Appendix C we record a number of elementary Hardy- and trace-type estimates. These will be used throughout the body of the paper.

2. Formulas for commutators and multipliers

In this section we recall some basic formulas which underlie the energy method for the wave equation on curved backgrounds. We do this first with respect to the original metric g. We then generalize such formulas cover the case of metric conformal to g. The latter will form the basis for most of the multiplier estimates in the sequel.

2A. *Identities involving g.* We begin with a basic definition:

Definition 2.1 (normalized deformation tensor). Let X be a vector field, and set ${}^{(X)}\pi = \mathcal{L}_X g$. Then we define ${}^{(X)}\hat{\pi} = {}^{(X)}\pi - \frac{1}{2} g \cdot \mathrm{trace}({}^{(X)}\pi)$ to be the "normalized deformation tensor" of X.

The quantity ${}^{(X)}\hat{\pi}$ underlies all of our formulas for multipliers and commutators as the following result shows:

Lemma 2.2 (formulas involving ${}^{(X)}\hat{\pi}$). *Let ϕ be a scalar field and X a vector field. As usual set $T_{\alpha\beta} = \partial_\alpha\phi\partial_\beta\phi - \frac{1}{2} g_{\alpha\beta} g^{\alpha'\beta'}\partial_{\alpha'}\phi\partial_{\beta'}\phi$ to be the energy momentum tensor of ϕ. Then one has the identities*

$$ {}^{(X)}\hat{\pi}^{\alpha\beta} = -\frac{1}{\sqrt{|g|}} X(\sqrt{|g|} g^{\alpha\beta}) - g^{\alpha\beta}\partial_\gamma X^\gamma + g^{\alpha\gamma}\partial_\gamma(X^\beta) + g^{\beta\gamma}\partial_\gamma(X^\alpha), \tag{25a} $$

$$ [\Box_g, X] = \nabla_\alpha {}^{(X)}\hat{\pi}^{\alpha\beta}\nabla_\beta - \tfrac{1}{2}\,\mathrm{trace}({}^{(X)}\hat{\pi})\Box_g, \tag{25b} $$

$$ \nabla^\alpha(T_{\alpha\beta}X^\beta) = \tfrac{1}{2}{}^{(X)}\hat{\pi}^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi + \Box_g\phi X\phi. \tag{25c} $$

Finally, if q is a smooth function then one has the commutator formula

$$ {}^{(qX)}\hat{\pi}^{\alpha\beta} - q\,{}^{(X)}\hat{\pi}^{\alpha\beta} = X^\alpha\nabla^\beta q + X^\beta\nabla^\alpha q - g^{\alpha\beta} X q. \tag{26} $$

We'll prove each of these formulas separately.

Proof of (25a) and (26). We have

$$ {}^{(X)}\pi^{\alpha\beta} = g^{\alpha\alpha'} g^{\beta\beta'}(\mathcal{L}_x g)_{\alpha'\beta'} = -X g^{\alpha\beta} + g^{\alpha\gamma}\partial_\gamma(X^\beta) + g^{\beta\gamma}\partial_\gamma(X^\alpha). $$

On the other hand

$$ \tfrac{1}{2}\,\mathrm{trace}({}^{(X)}\pi) = \nabla_\alpha X^\alpha = \frac{1}{\sqrt{|g|}}\partial_\alpha(\sqrt{|g|}X^\alpha) = \frac{1}{\sqrt{|g|}} X(\sqrt{|g|}) + \partial_\alpha X^\alpha. $$

Subtracting the last two identities gives (25a). A direct application of (25a) shows (26). □

Proof of (25b). We begin with a formula that will also be useful in the sequel. Let $\mathcal{R}^{\alpha\beta}$ be any contravariant 2-tensor; then we claim

$$ [X, \nabla_\alpha\mathcal{R}^{\alpha\beta}\nabla_\beta] = \nabla_\alpha\widetilde{\mathcal{R}}^{\alpha\beta}\nabla_\beta + \widetilde{\mathcal{S}}^\beta\nabla_\beta, \quad \text{where } \widetilde{\mathcal{R}} = \mathcal{L}_X\mathcal{R} \text{ and } \widetilde{\mathcal{S}}^\beta = \mathcal{R}^{\alpha\beta}\nabla_\alpha(\nabla_\gamma X^\gamma). \tag{27} $$

To prove it, first note that a straightforward calculation using the coordinate-based formula for $\nabla_\alpha X^\alpha$ above shows $X(\nabla_\alpha Y^\alpha) - Y(\nabla_\alpha X^\alpha) = \nabla_\alpha [X, Y]^\alpha$ for any pair of vector fields X and Y. Applying this last formula to the vector field $Y^\alpha = \mathcal{R}^{\alpha\beta}\nabla_\beta\phi$ and using the Leibniz rule for Lie derivatives followed by $[\mathcal{L}_X, d] = 0$ for the exterior derivative d gives (27).

Now apply (27) to $\mathcal{R} = g^{-1}$. Using $(\mathcal{L}_X\mathcal{R})^{\alpha\beta} = -{}^{(X)}\pi^{\alpha\beta}$ and $\nabla_\alpha X^\alpha = -\frac{1}{2}\operatorname{trace}({}^{(X)}\hat{\pi})$ gives (25b). □

Proof of (25c). A standard calculation shows $\nabla^\alpha(T_{\alpha\beta}X^\beta) = \frac{1}{2}{}^{(X)}\pi^{\alpha\beta}T_{\alpha\beta} + \Box_g\phi X\phi$ and (25c) follows easily. □

2B. *Conformal changes for scalar fields.* In this section we recall a standard formula from geometry. Let $g_{\alpha\beta}$ be a Lorentzian metric on a (3+1)-dimensional spacetime. We consider a conformally equivalent metric $\tilde{g}_{\alpha\beta}$, where $\Omega^2\tilde{g} = g$ for some weight function $\Omega > 0$. Let $\tilde{\nabla}$ denote the Levi-Civita connection of $\tilde{g}$ and $\Box_{\tilde{g}} = \tilde{\nabla}^\alpha\tilde{\nabla}_\alpha$ the corresponding wave operator. Then we have:

Lemma 2.3 (identity for the conformal wave operator). *Let $\Box_g\phi = F$. Then one has the formula*

$$\Box_{\tilde{g}}\psi + V\psi = \Omega^3 F, \quad \textit{where } \psi = \Omega\phi \textit{ and } V = \Omega^3\Box_g\Omega^{-1}. \tag{28}$$

Proof. Start with

$$\Box_{\tilde{g}} = \Omega^4\frac{1}{\sqrt{|g|}}\partial_\alpha(\Omega^{-2}\sqrt{|g|}g^{\alpha\beta}\partial_\beta) = \Omega^2(\Box_g - 2g^{\alpha\beta}\partial_\alpha\ln(\Omega)\partial_\beta).$$

To eliminate the second term on the right-hand side we rescale ϕ via $\psi = \Omega\phi$, which gives us

$$\Omega^{-2}\Box_{\tilde{g}}\psi = \Box_g\psi - 2g^{\alpha\beta}\partial_\alpha\ln(\Omega)\partial_\beta\psi = \Omega\Box_g\phi - W\phi,$$

where

$$W = -\Box_g(\Omega) + 2\Omega\partial_\alpha\ln(\Omega)\partial^\alpha\ln(\Omega) = \Omega^2\Box_g\Omega^{-1}.$$

A straightforward manipulation of the last two equations gives (28). □

2C. *Conformal multipliers.* We now combine the identities of the last two sections to produce conjugated weighted L^2 identities. Our main result here is:

Lemma 2.4 (the conformal divergence identity). *Let $\Box_g\phi = F$. Let X be a vector field supported in the exterior region $\mathbb{R}^3\setminus\mathcal{K}$. Let $\chi(t,x)$ be a smooth function and $\Omega > 0$ a smooth weight. Then in (t,x)-coordinates one has the divergence identity*

$$\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}} Q(X,\chi,\Omega,\phi)\sqrt{|g|}\,dx\,dt = \int_{\mathbb{R}^3\setminus\mathcal{K}} P(X,\chi,\Omega,\phi)\sqrt{|g|}\,dx\Big|_{t=0}^{t=T}, \tag{29}$$

where

$$P(X,\chi,\Omega,\phi) = \Omega^{-2}g^{0\alpha}\partial_\alpha(\Omega\phi)X(\Omega\phi) - \tfrac{1}{2}\Omega^{-2}X^0(g^{\alpha\beta}\partial_\alpha(\Omega\phi)\partial_\beta(\Omega\phi) - \chi V\phi^2), \tag{30}$$

with $V = \Omega^3\Box_g\Omega^{-1}$, and where

$$Q(X,\chi,\Omega,\phi) = F\cdot\Omega^{-1}X(\Omega\phi) + \Omega^{-2}A^{\alpha\beta}\partial_\alpha(\Omega\phi)\partial_\beta(\Omega\phi) + B^\chi\phi^2 + C^\chi\phi,\ \Omega^{-1}X(\Omega\phi), \tag{31}$$

with

$$A = \tfrac{1}{2}({}^{(X)}\hat{\pi} + 2X\ln(\Omega)g^{-1}), \quad B^\chi = \tfrac{1}{2}\Omega^{-2}(X(\chi V) - \text{trace}(A)\chi V), \quad C^\chi = \Omega^{-2}(\chi - 1)V. \tag{32}$$

On the right-hand sides of the last two equations above all contractions are computed with respect to g.

Proof. First define the tensors

$$\widetilde{T}^\chi_{\alpha\beta} = \partial_\alpha\psi\partial_\beta\psi - \tfrac{1}{2}\tilde{g}_{\alpha\beta}(\tilde{g}^{\gamma\delta}\partial_\gamma\psi\partial_\delta\psi - \chi V\psi^2), \quad {}^{(X)}\widetilde{P}^\chi_\alpha = \widetilde{T}^\chi_{\alpha\beta}X^\beta, \quad \text{where } \psi = \Omega\phi.$$

Then by Stokes' theorem and the support property of X we have

$$\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}} \tilde{\nabla}^{\alpha\,(X)}\widetilde{P}^\chi_\alpha\sqrt{|\tilde{g}|}\,dx\,dt = \int_{\mathbb{R}^3\setminus\mathcal{K}} \tilde{g}^{\alpha 0}\widetilde{T}^\chi_{\alpha\beta}X^\beta\sqrt{|\tilde{g}|}\,dx\Big|_{t=0}^{t=T},$$

and right-hand side of (29) follows by substituting $\Omega^{-2}g = \tilde{g}$ into the volume forms and the right-hand side contractions.

It remains to compute the $\tilde{g}$-contraction $\tilde{\nabla}^{\alpha\,(X)}\widetilde{P}^\chi_\alpha$ and show this produces the terms on left-hand side of (29). To this end suppose $(\Box_{\tilde{g}} + V)\psi = G$. Then one has

$$\tilde{\nabla}^\alpha\widetilde{T}^\chi_{\alpha\beta} = ((\chi-1)V\psi + G)\partial_\beta\psi + \tfrac{1}{2}\partial_\beta(\chi V)\psi^2.$$

Using formula (25c) we have

$$\tilde{\nabla}^{\alpha\,(X)}\widetilde{P}^\chi_\alpha = \tfrac{1}{2}(\widehat{\mathcal{L}_X\tilde{g}})^{\alpha\beta}\partial_\alpha\psi\partial_\beta\psi - \tfrac{1}{4}\,\text{trace}(\widehat{\mathcal{L}_X\tilde{g}})\chi V\psi^2 + ((\chi-1)V\psi + G)X\psi + \tfrac{1}{2}X(\chi V)\psi^2, \tag{33}$$

where all the contractions are computed with respect to $\tilde{g}$. To compute the first two terms on the right-hand side we use

$$\mathcal{L}_X\tilde{g} = \Omega^{-2}(\mathcal{L}_Xg - 2X\ln(\Omega)g), \quad \text{and so} \quad \widehat{\mathcal{L}_X\tilde{g}} = \Omega^{-2}(\widehat{\mathcal{L}_Xg} + 2X\ln(\Omega)g).$$

Substituting the last equation into right-hand side of (33) and using $G = \Omega^3F$ we have (31) and (32). □

3. Algebraic formulas involving Bondi coordinates

In this section we compute the key quantities from Lemmas 2.2 and 2.4 in Bondi coordinates (u, x).

Lemma 3.1 (formulas for deformation tensors). *Let $d = |g|$ be the determinant $g_{\alpha\beta}$ in rectangular Bondi coordinates (u, x^i), and set $\Omega = \tau_x$. Then if $X = X^\alpha\partial^b_\alpha$ is any vector field in Bondi coordinates we have the formula for contravariant tensors*

$$ {}^{(X)}\hat{\pi} + 2X\ln(\Omega)g^{-1} = -d^{-\frac{1}{2}}(\mathcal{L}_Xh + (\partial^b_uX^u + \partial^b_rX^r + \partial^b_i\overline{X}^i)h) + \mathcal{R}, \tag{34}$$

where $\overline{X}^i = X^i - r^{-2}x^ix_jX^j$ denotes the angular portion of X and $X^r = r^{-1}x_iX^i$ the radial portion, and where h is given in (4). The remainder tensor $\mathcal{R}$ is given by the covariant formula

$$\mathcal{R} = -d^{-\frac{1}{2}}\mathcal{L}_X(d^{\frac{1}{2}}g^{-1} - h) - d^{-\frac{1}{2}}(\partial^b_uX^u + \partial^b_rX^r + \partial^b_i\overline{X}^i)(d^{\frac{1}{2}}g^{-1} - h) - 2r^{-1}\tau_x^{-2}X^rg^{-1}. \tag{35}$$

Proof of formula (34). Starting with formula (25a) in Bondi coordinates and then using the identity

$$\partial^b_\gamma X^\gamma = 2X\ln(\tau_x) + \partial^b_u X^u + \partial^b_r X^r + \partial^b_i \overline{X}^i + 2r^{-1}\tau_x^{-2}X^r$$

and setting $\Omega = \tau_x$, we have the formula for raised indices

$${}^{(X)}\hat{\pi} + 2X\ln(\Omega)g^{-1} = -d^{-\frac{1}{2}}\mathcal{L}_X(d^{\frac{1}{2}}g^{-1}) - (\partial^b_u X^u + \partial^b_r X^r + \partial^b_i \overline{X}^i + 2r^{-1}\tau_x^{-2}X^r)g^{-1}.$$

Writing $g^{-1} = d^{-\frac{1}{2}}(d^{\frac{1}{2}}g^{-1} - h) + d^{-\frac{1}{2}}h$ and inserting into this last equation gives (34) and (35). □

Lemma 3.2 (formulas for commutators). *The following commutator formulas hold where $d = |g|$ is computed in Bondi coordinates*:

(I) *For $X \in \{\partial^b_u, \partial^b_i - \omega^i\partial^b_u, \Omega_{ij}\}$ one has the formula*

$$[\Box_g, X] = \nabla_\alpha \mathcal{R}^{\alpha\beta}\nabla_\beta + \tfrac{1}{2}(X\ln(d))\Box_g, \quad \text{where } \mathcal{R} = -d^{-\frac{1}{2}}\mathcal{L}_X(d^{\frac{1}{2}}g^{-1} - h) + \mathcal{R}_1, \tag{36}$$

where $\mathcal{R}_1 = 0$ for $X \in \{\partial^b_u, \Omega_{ij}\}$, and $\mathcal{R}_1^{\alpha\beta} = 2d^{-\frac{1}{2}}\omega^i\tau_x^{-3}\delta^\alpha_u\delta^\beta_u$ when $X = \partial^b_i - \omega^i\partial^b_u$. Again h is as defined in (4).

(II) *For $X = S$ one has*

$$[\Box_g, S] = \nabla_\alpha \mathcal{R}^{\alpha\beta}\nabla_\beta + \tfrac{1}{2}(4 + S\ln(d))\Box_g, \quad \text{where } \mathcal{R} = -d^{-\frac{1}{2}}\mathcal{L}_S(d^{\frac{1}{2}}g^{-1} - h) - 2d^{-\frac{1}{2}}(d^{\frac{1}{2}}g^{-1} - h) + \mathcal{R}_1, \tag{37}$$

and where $\mathcal{R}_1^{\alpha\beta} = d^{-\frac{1}{2}}\omega^i\tau_x^{-2}(\delta^\alpha_i\delta^\beta_u + \delta^\beta_i\delta^\alpha_u)$.

Proof of formulas (36) *and* (37). First note that formula (25a) above can be rewritten as

$${}^{(X)}\hat{\pi} = -d^{-\frac{1}{2}}\mathcal{L}_X(d^{\frac{1}{2}}g^{-1} - h) - d^{-\frac{1}{2}}\mathcal{L}_X h - \partial^b_\alpha X^\alpha g^{-1}, \quad \mathrm{trace}({}^{(X)}\hat{\pi}) = -2\partial^b_\alpha X^\alpha - X\ln(d). \tag{38}$$

Next, for each $X \in \{\partial^b_u, \Omega_{ij}, \partial^b_i - \omega^i\partial^b_u\}$ we have $\partial^b_\alpha X^\alpha = 0$, and for $X \in \{\partial^b_u, \Omega_{ij}\}$ we also have $\mathcal{L}_X h = 0$. On the other hand for $X = \partial^b_i - \omega^i\partial^b_u$ one computes $\mathcal{L}_X h^{\alpha\beta} = 0$ for all but the uu-component, and for this one has $(\mathcal{L}_X h)^{uu} = -2\omega^i\tau_x^{-3}$. Combining this information with (38) above and (25b) gives (36).

Finally, in the case when $X = S$ we compute $\mathcal{L}_S h + 2h = -\omega^i\tau_x^{-2}(\delta^\alpha_i\delta^\beta_u + \delta^\beta_i\delta^\alpha_u)$. This allows us to write for $X = S$

$$d^{-\frac{1}{2}}\mathcal{L}_X h + \partial^b_\alpha X^\alpha g^{-1} = 2d^{-\frac{1}{2}}(d^{\frac{1}{2}}g^{-1} - h) - \mathcal{R}_1 + 2g^{-1}, \quad \text{where } \mathcal{R}_1^{\alpha\beta} = d^{-\frac{1}{2}}\omega^i\tau_x^{-2}(\delta^\alpha_i\delta^\beta_u + \delta^\beta_i\delta^\alpha_u).$$

Using $\mathrm{trace}({}^{(X)}\hat{\pi}) = -8 - S\ln(d)$ and combining everything with (38) above and (25b) gives (37). □

4. Asymptotic estimates involving Bondi coordinates

We now move on to the main technical calculations of the paper. We record these here in a general form that will be used throughout the sequel.

4A. ***Basic estimates for derivatives.***

Lemma 4.1 (estimates for the determinant). *Let $d = |g| = |\det(g)|$ be the absolute determinant of g computed in Bondi coordinates (u, x^i). Then for any $\mu \in \mathbb{R}$ one has the symbol bounds*

$$d^{\mu} - 1 \in \mathcal{Z}^0, \tag{39}$$

where the symbol spaces $\mathcal{Z}^k$ are defined in (2).

Proof of (39). We can write $d^{\mu} - 1 = q_{\mu}(d^{-1} - 1)$, where $q_{\mu}(s)$ is smooth for $s > -1$ and $q(0) = 0$. Thus, by Taylor expansion and the Leibniz and chain rules it suffices to consider the case $\mu = -1$. From (3) we know that

$$d^{-1} - 1 + \tau_x^{-2} = \det(h) - \det(g^{-1}) \in \mathcal{Z}^0,$$

where h is given in (4). Since $\tau_x^{-2} \in \mathcal{Z}^0$, this completes the proof. □

A useful corollary of this last lemma and the assumptions (3) is the following:

Corollary 4.2. *Let $d = |g| = |\det(g)|$ be the absolute determinant of g computed in Bondi coordinates (u, x^i), and g^{-1} the inverse metric of g in Bondi coordinates. Then if h is as in (4) and one sets $\mathcal{R}^{\alpha\beta} = (d^{\frac{1}{2}} g^{-1} - h)^{\alpha\beta}$, there holds the bounds*

$$\mathcal{R}^{ij} \in \mathcal{Z}^0, \quad \mathcal{R}^{ui} \in \mathcal{Z}^{\frac{1}{2}}, \quad \mathcal{R}^{ui} - \omega^i \omega_j \mathcal{R}^{uj} \in \mathcal{Z}^1, \quad \mathcal{R}^{uu} \in \mathcal{Z}^2. \tag{40}$$

Building on the last two results we have the following collection of symbol bounds which will underlie many of the error estimates in the sequel.

Lemma 4.3 (basic Lie derivative estimates). *Let $X = X^{\alpha}\partial_{\alpha}$ be a vector field. Then the following hold:*

(I) *Suppose that in Bondi coordinates X satisfies the symbol-type bounds for $a, b, c \in \mathbb{R}$*

$$|(\tau_- \partial_u^b)^l (\tau_x \partial_x^b)^J X^u| \lesssim_{l,J} \tau_x^a \tau_+^b \tau_-^{c+1} \left(\frac{\tau_x}{\tau_+}\right)^{\min\{|J|,1\}}, \quad |(\tau_- \partial_u^b)^l (\tau_x \partial_x^b)^J X^i| \lesssim_{l,J} \tau_x^{a+1} \tau_+^b \tau_-^c \tag{41}$$

and obeys the conditions

$$\partial_r^b X^u = \partial_u^b (X^i) = \partial_r^b r^{-1}(X^i - r^{-2} x^i x_j X^j) = 0. \tag{42}$$

Let $\mathcal{R}^{\alpha\beta}$ be any contravariant 2-tensor which satisfies (40) with similar estimates for $\mathcal{R}^{iu}$ (if it is nonsymmetric). Then its Lie derivative by X, denoted by $\mathcal{L}_X \mathcal{R} = \mathcal{R}_X$, satisfies

$$\mathcal{R}_X^{ij} \in \tau_x^a \tau_+^b \tau_-^c \cdot \mathcal{Z}^0, \quad \mathcal{R}_X^{ui} \in \tau_x^a \tau_+^b \tau_-^c \cdot \mathcal{Z}^{\frac{1}{2}}, \quad \mathcal{R}_X^{ui} - \omega^i \omega_j \mathcal{R}_X^{uj} \in \tau_x^a \tau_+^b \tau_-^c \cdot \mathcal{Z}^1, \quad \mathcal{R}_X^{uu} \in \tau_x^a \tau_+^b \tau_-^c \cdot \mathcal{Z}^2, \tag{43}$$

with similar bounds for $\mathcal{R}_X^{iu}$.

(II) *Alternatively, if one drops the condition $\partial_r^b X^u = 0$ but keeps the rest of (41) and (42), then the previous conclusion holds with the last bound in (43) replaced by*

$$\mathcal{R}_X^{uu} \in \tau_x^a \tau_+^b \tau_-^c \cdot \mathcal{Z}^{\frac{3}{2}}. \tag{44}$$

(III) *Alternatively, if one drops the condition $\partial^b_u(X^r) = 0$ but retains $\partial^b_u(X^i - r^{-2}x^i x_j X^j) = 0$ and the rest of (41) and (42), then the result of part (I) holds with the first bound in (43) replaced by the pair*

$$\mathcal{R}^{ij}_X \in \tau^a_x \tau^b_+ \tau^c_- \cdot \mathcal{Z}^{-\frac{1}{2}}, \quad \mathcal{R}^{ij}_X - \omega^i \omega^j \omega_k \omega_l \mathcal{R}^{kl}_X \in \tau^a_x \tau^b_+ \tau^c_- \cdot \mathcal{Z}^0. \tag{45}$$

(IV) *Likewise, if $\mathcal{S}^\alpha$ satisfies*

$$\mathcal{S}^i \in \tau_x^{-1} \cdot \mathcal{Z}^{-\frac{1}{2}}, \quad \mathcal{S}^u \in \tau_x^{-1} \cdot \mathcal{Z}^{\frac{1}{2}}, \tag{46}$$

then $\mathcal{L}_X \mathcal{S} = S_X$ satisfies

$$\mathcal{S}^i_X \in \tau^{a-1}_x \tau^b_+ \tau^c_- \cdot \mathcal{Z}^{-\frac{1}{2}}, \quad \mathcal{S}^u_X \in \tau^{a-1}_x \tau^b_+ \tau^c_- \cdot \mathcal{Z}^{\frac{1}{2}}, \tag{47}$$

when X satisfies the symbol bounds (41) (in this case we do not need the extra conditions (42)).

(V) *Finally, let $X \in \mathbb{L}_0 = \{\partial^b_u, \partial^b_i - \omega^i \partial^b_u\}$, and let $\mathcal{R}$ and $\mathcal{S}$ satisfy (40) and (46) respectively. Then $\mathcal{L}_X \mathcal{R}$ and $\mathcal{L}_X \mathcal{S}$ satisfy (43) and (47) with $a = c = -1$ and $b = 1$.*

Remark 4.4. As will become apparent in the proof, if one is only interested in the norms (2) at level $N = 0$ for $\mathcal{R}^{\alpha\beta}_X$ and $\mathcal{S}^\alpha_X$ (i.e., no derivatives), then one can replace the full symbol bounds (41) with first-order conditions

$$\sum_{l+|J|\leq 1} |(\tau_- \partial^b_u)^l (\tau_+ \partial^b_x)^J X^u| \lesssim \tau^a_x \tau^b_+ \tau^{c+1}_-, \qquad \sum_{l+|J|\leq 1} |(\tau_- \partial^b_u)^l (\tau_x \partial^b_x)^J X^i| \lesssim \tau^{a+1}_x \tau^b_+ \tau^c_-.$$

In this case the various implications above are true with the inclusion $\mathcal{R} \in \tau^a_x \tau^b_+ \tau^c_- \cdot \mathcal{Z}^k$ replaced by the bound $|||\tau_x^{-a} \tau_+^{-b} \tau_-^{-c} \mathcal{R}|||_{k,0} < \infty$.

Proof of Lemma 4.3. We prove the various portions separately. First note that the conditions (40) are invariant with respect to dyadic cutoffs in the r-, u-, and $(t+r)$-variables. Therefore by utilizing such cutoffs and the Leibniz rule we may assume $a = b = c = 0$. As a second preliminary note the identity

$$\hat{x}^i - \omega^i = \omega^i r^{-1} (r + \tau_x)^{-1}, \quad \text{where } \hat{x}^i = r^{-1} x^i. \tag{48}$$

This allows us to trade $\hat{x}^i$ for ω^i in the region $r > 1$ as long as errors on the order of r^{-2} are acceptable.

Part 1: (*the $\mathcal{R}$ bounds involving condition* (41)) We begin with the proof of estimates (40) for $\mathcal{L}_X \mathcal{R}$ assuming conditions (41) and (42) or one of the alternatives listed in items (II) and (III) above. The formula for the Lie derivative is $\mathcal{L}_X \mathcal{R}^{\alpha\beta} = X(\mathcal{R}^{\alpha\beta}) - \partial_\gamma(X^\alpha)\mathcal{R}^{\gamma\beta} - \partial_\gamma(X^\beta)\mathcal{R}^{\alpha\gamma}$. We check each component separately:

Case 1a: (*the uu-component assuming $\partial^b_r X^u = 0$*) By assumption we have $\omega^i \partial^b_i X^u = 0$; thus

$$\mathcal{L}_X \mathcal{R}^{uu} = X(\mathcal{R}^{uu}) - 2\partial^b_u(X^u)\mathcal{R}^{uu} - \partial^b_i(X^u)(\mathcal{R}^{ui} - \omega^i \omega_j \mathcal{R}^{uj}) - \partial^b_i(X^u)(\mathcal{R}^{iu} - \omega^i \omega_j \mathcal{R}^{ju}).$$

Then the estimate in (43) for $\mathcal{R}^{uu}_X$ is immediate from the estimates (41) and (40).

Case 1b: (*the uu-component when $\partial^b_r X^u \neq 0$*) By the previous case we have the desired estimate modulo the additional expression $r^2 \tau_x^{-2} \partial^b_r(X^u)(\mathcal{R}^{ur} + \mathcal{R}^{ru})$, which adds a $\mathcal{Z}^{\frac{3}{2}}$-term.

Case 2: (*the ui- and iu-components*) By symmetry it suffices to treat the ui case. We have

$$\mathcal{L}_X\mathcal{R}^{ui} = X(\mathcal{R}^{ui}) - \partial_u^b(X^u)\mathcal{R}^{ui} - \partial_u^b(X^i)\mathcal{R}^{uu} - \partial_j^b(X^u)\mathcal{R}^{ji} - \partial_j^b(X^i)\mathcal{R}^{uj}.$$

Using estimates (41) and (40) we get a $\mathcal{Z}^{\frac{1}{2}}$ symbol bound for this term. In addition one sees that for all parts of the formula above, save for the expression $\mathcal{B}^i = X(\omega^i\omega_j\mathcal{R}^{uj}) - \partial_u^b(X^u)\omega^i\omega_j\mathcal{R}^{uj} - \omega^k\partial_k^b(X^i)\omega_j\mathcal{R}^{uj}$, the bound is on the order of $\mathcal{Z}^1$. To show improved bounds we only need to consider the region $r > 1$. Using (48) we see that $\mathcal{B} \equiv \widetilde{\mathcal{B}} \mod r^{-2}\cdot\mathcal{Z}^{\frac{1}{2}}$, where $\widetilde{\mathcal{B}}^i = X(\hat{x}^i\mathcal{R}^{ur}) - \partial_u^b(X^u)\hat{x}^i\mathcal{R}^{ur} - \partial_r^b(X^i)\mathcal{R}^{ur}$. Again using (48), we see that in order to show $\mathcal{B}^i - \omega^i\omega_j\mathcal{B}^j \in \mathcal{Z}^1$ it suffices to prove $\chi_{r>1}(\widetilde{\mathcal{B}}^i - \hat{x}^i\widetilde{\mathcal{B}}^r) \in \mathcal{Z}^1$. This would follow immediately if $\widetilde{\mathcal{B}}$ is a radially directed vector field. Using $X(\hat{x}^i) = r^{-1}(X^i - \hat{x}^iX^r)$ we compute

$$\widetilde{\mathcal{B}}^i = \hat{x}^i(X(\mathcal{R}^{ur}) - \partial_u^b(X^u)\mathcal{R}^{ur} - \partial_r^b(X^r)\mathcal{R}^{ur}) - r\partial_r^b[r^{-1}(X^i - \hat{x}^iX^r)]\mathcal{R}^{ur},$$

which is manifestly radial thanks to the last condition in (42).

Case 3a: (*the ij-components assuming $\partial_u^b(X^i) = 0$*) Here we have

$$\mathcal{L}_X\mathcal{R}^{ij} = X(\mathcal{R}^{ij}) - \partial_k^b(X^i)\mathcal{R}^{kj} - \partial_k^b(X^j)\mathcal{R}^{ik}.$$

Then the estimate in (43) for $\mathcal{R}_X^{ij}$ is immediate from the estimates (41) and (40).

Case 3b: (*the ij-components assuming $\partial_u^b(X^r) \neq 0$*) In this case we are still assuming $\partial_u^b(X^i - \hat{x}^iX^r) = 0$. Therefore we have

$$\mathcal{L}_X\mathcal{R}_X^{ij} = X(\mathcal{R}^{ij}) - \hat{x}^i\partial_u^b(X^r)\mathcal{R}^{uj} - \hat{x}^i\partial_u^b(X^r)\mathcal{R}^{iu} - \partial_k^b(X^i)\mathcal{R}^{kj} - \partial_k^b(X^j)\mathcal{R}^{ik}.$$

A $\mathcal{Z}^{-\frac{1}{2}}$ symbol bound for this expression in $r > 1$ is again immediate from (41) and (40). On the other hand all but the second and third terms above yield an improved $\mathcal{Z}^0$ bound. Thus, using (48) we have for $r > 1$

$$\mathcal{R}_X^{ij} - \omega^i\omega^j\omega_k\omega_l\mathcal{R}_X^{kl} \equiv -\omega^i\partial_u^b(X^r)(\mathcal{R}^{uj} - \omega^j\omega_k\mathcal{R}^{uk}) - \omega^j\partial_u^b(X^r)(\mathcal{R}^{iu} - \omega^i\omega_k\mathcal{R}^{ku}) \mod r^{-2}\cdot\mathcal{Z}^{-\frac{1}{2}} + \mathcal{Z}^0.$$

By (40) and (41) we have a $\mathcal{Z}^0$ bound for this last term as well.

Part 2: (*the $\mathcal{S}$ bounds involving condition* (41)) Again we can reduce to $a = b = c = 0$. Componentwise we have $\mathcal{L}_X\mathcal{S}^\alpha = X(\mathcal{S}^\alpha) - \partial_\beta(X^\alpha)\mathcal{S}^\beta$.

Case 1: (*the u-component*) Here we have

$$\mathcal{L}_X\mathcal{S}^u = X(\mathcal{S}^u) - \partial_u^b(X^u)\mathcal{S}^u - \partial_i^b(X^u)\mathcal{S}^i,$$

so the second estimate in (46) follows directly by multiplying together the bounds in (41) and (46).

Case 2: (*the i-components*) Here we have

$$\mathcal{L}_X\mathcal{S}^i = X(\mathcal{S}^i) - \partial_u^b(X^i)\mathcal{S}^u - \partial_j^b(X^i)\mathcal{S}^j,$$

and so the first estimate in (46) follows directly from (41) and (46).

Part 3: (*estimates involving* $\mathbb{L}_0$) This is largely a corollary of Parts 1 and 2 above. For any $X \in \mathbb{L}_0$ one has both conditions in (41), with $a = c = -1$ and $b = 1$. In this case one also has to deal with the fact that $\partial_r^b X^u \neq 0$ and $\partial_r^b r^{-1}(X^i - \hat{x}^i \hat{x}_j X^j) \neq 0$, save for when $X = \partial_u^b$. Recall that these two special conditions were only used in Case 1b and Case 2 of Part 1 above. So we review those cases here when $X = \partial_i^b - \omega^i \partial_u^b$.

Case 1: (*the* uu*-component of* $\mathcal{R}_X$) Recall from Case 1b of Part 1 above we only need to handle an expression of the form $\partial_r^b(X^u)(\mathcal{R}^{ur} + \mathcal{R}^{ru})$ in the region $r > 1$, where $X^u = \omega^i$. Using $\partial_r^b(\omega^i) = r^{-1}\tau_x^{-2}\omega^i$ we get an $\tau_x^{-3} \cdot \mathcal{Z}^{\frac{1}{2}} \subseteq \tau_x^{-1}\tau_-^{-1}\tau_+\mathcal{Z}^2$ bound for this expression, which suffices.

Case 2: (*the* ui*-component of* $\mathcal{R}_X$) Recall that the condition $\partial_r^b r^{-1}(X^i - \hat{x}^i \hat{x}_j X^j) = 0$ is only used to establish the improved bound $\mathcal{R}_X^{iu} - \omega^i \omega_j \mathcal{R}_X^{uj} \in \mathcal{Z}^1$. Recall further that this improved bound automatically holds modulo an expression of the form $r\partial_r^b[r^{-1}(X^i - \omega^i \omega_j X^j)]\mathcal{R}^{ur}$. When $X^i = 0, 1$ we see this expression has symbol bounds on the order of $\tau_x^{-1} \cdot \mathcal{Z}^{\frac{1}{2}} \subseteq \tau_x^{-1}\tau_-^{-1}\tau_+\mathcal{Z}^1$, which suffices. □

4B. ***A general exterior multiplier estimate.*** Next, we prove some general multiplier bounds which will be used a number of times in the sequel. To state them we first define the form of an acceptable error.

Definition 4.5 (general form of multiplier estimate errors). For a pair of parameters $0 < a < 1$ and $R > 0$, and a quantity $o_R(1) \to 0$ as $R \to \infty$, we set

$$\begin{aligned}\mathcal{E}(a, R) &= \|\tau_x^a \partial\phi(0)\|^2_{L_x^2(r>\frac{1}{2}R)} + o_R(1) \cdot \big(\sup_{0\le t\le T} \|\phi(t)\|^2_{E^a(r>\frac{1}{2}R)} + \|\phi\|^2_{S^a(r>\frac{1}{2}R)[0,T]}\big) \\ &\quad + \|\phi\|_{S^a(r>\frac{1}{2}R)[0,T]} \cdot \big(R^{-\frac{1}{2}}\|(\tau_-^a\partial_u^b\phi,\, \tau_x^a\partial_x^b\phi,\, \tau_x^{a-1}\phi)\|_{L^2(\frac{1}{2}R<r<R)[0,T]} + \|\Box_g\phi\|_{N^a(r>\frac{1}{2}R)[0,T]}\big). \end{aligned} \tag{49}$$

Corresponding to the cases $a = 0, 1$, for parameter $R > 0$, quantity $o_R(1)$, and vector field X, we set

$$\begin{aligned}\mathcal{E}(0, R, X) &= \|\partial\phi(0)\|^2_{L_x^2(r>\frac{1}{2}R)} + o_R(1) \cdot \big(\sup_{0\le t\le T} \|\partial\phi(t)\|^2_{L_x^2(r>\frac{1}{2}R)} + \|\phi\|^2_{\mathrm{LE}^0(r>\frac{1}{2}R)[0,T]}\big) \\ &\quad + R^{-1}\|(\partial\phi,\, \tau_x^{-1}\phi)\|^2_{L^2(\frac{1}{2}R<r<R)[0,T]} + \left| \int_0^T \int_{\mathbb{R}^3\setminus\mathcal{K}} \Box_g\phi \cdot \tau_x^{-1} X(\tau_x\phi)\, dV_g \right|, \end{aligned} \tag{50}$$

and

$$\begin{aligned}\mathcal{E}(1, R, X) &= \|\tau_x\partial\phi(0)\|^2_{L_x^2(r>\frac{1}{2}R)} + o_R(1) \cdot \big(\sup_{0\le t\le T} \|\phi(t)\|^2_{E^1(r>\frac{1}{2}R)} + \|\phi\|^2_{S^{1,\infty}(r>\frac{1}{2}R)[0,T]}\big) \\ &\quad + \|\phi\|_{S^{1,\infty}(r>\frac{1}{2}R)[0,T]} \cdot R^{-\frac{1}{2}}\|(\tau_-\partial_u^b\phi,\, \tau_x\partial_x^b\phi,\, \phi)\|_{\ell_t^1 L^2(\frac{1}{2}R<r<R)[0,T]} \\ &\quad + \left| \int_0^T \int_{\mathbb{R}^3\setminus\mathcal{K}} \Box_g\phi \cdot \tau_x^{-1} X(\tau_x\phi)\, dV_g \right|. \end{aligned} \tag{51}$$

In the notation above the rate of $o_R(1)$ may change from line to line, but is fixed for any line on which an error of the form $\mathcal{E}$ appears. Also we define $dV_g = \sqrt{|g|}\, dx\, dt$, where $|g| = |\det g|$ is computed in (t, x)-coordinates.

With this notation in mind we have:

Proposition 4.6 (abstract multiplier estimate). *Fix $R>0$ sufficiently large so that $\mathcal{K}\subset\{r<\frac{1}{2}R\}$, and let Y be a vector field such that $Y=Y^u\partial_u^b+Y^r\partial_r^b$, with both $Y^u\geq 0$ and $Y^r\geq 0$ depending only on the (u,r)-variables. Then the following hold:*

(I) *Assume for some $0<a<1$ there holds the symbol-type bounds*

$$\sum_{i+|J|\leq 1}|(\tau_-\partial_u^b)^i(\tau_+\partial_r^b)^J Y^u|\lesssim\tau_+^{2a}\tau_0^{\max\{1,2a\}},\qquad\sum_{i+|J|\leq 1}|(\tau_-\partial_u^b)^i(\tau_x\partial_r^b)^J Y^r|\lesssim\tau_x^{2a}+\tau_+^{2a-1}\tau_x. \tag{52}$$

Then one has the multiplier estimate

$$\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}}\chi_{>R}\big(\mathcal{A}^u(\partial_u^b\phi)^2+\mathcal{A}^{ur}\partial_u^b\phi\cdot\tau_x^{-1}\partial_r^b(\tau_x\phi)+\mathcal{A}^r(\tau_x^{-1}\partial_r^b(\tau_x\phi))^2+\not{\mathcal{A}}|\not\nabla^b\phi|^2\big)\,dx\,dt$$
$$+\|\tau_x^{-1}(\sqrt{Y^u}\partial(\tau_x\phi),\sqrt{Y^r}\partial_x^b(\tau_x\phi))(T)\|^2_{L_x^2(r>R)}\lesssim\mathcal{E}(a,R), \tag{53}$$

where $|\not\nabla^b\phi|^2$ denotes the (Euclidean) angular gradient of ϕ with respect to the spheres $u=\text{const}$ and $r=\text{const}$, and where the components of $\mathcal{A}$ are given by

$$\mathcal{A}^u=-\partial_r^bY^u,\quad\mathcal{A}^r=\tfrac{1}{2}\partial_r^bY^r-\tfrac{1}{2}\partial_u^bY^u-\partial_u^bY^r, \tag{54a}$$
$$\mathcal{A}^{ur}=\partial_r^bY^u,\qquad\not{\mathcal{A}}=r^{-1}Y^r-\tfrac{1}{2}\partial_u^bY^u-\tfrac{1}{2}\partial_r^bY^r. \tag{54b}$$

Here $\chi_{>R}=\chi_{>1}(R^{-1}\cdot)$ is a radial bump function with $\chi_{>R}\equiv 1$ on $r>R$ and $\chi_{>R}\equiv 0$ on $r<\frac{1}{2}R$. The implicit constant in (53) depends only on the bounds from (52), and the metric g.

(II) *In the case $a=1$, still assuming (52), we have (53)–(54) with the right-hand side of (53) replaced by $\mathcal{E}(1,R)=\mathcal{E}(1,R,\chi_{r>R}Y)$, where $\chi_{>R}$ is as above.*

(III) *Alternatively, in the case $a=0$ replace assumption (52) with*

$$\sum_{i+|J|\leq 1}|(\tau_-\partial_u^b)^i(\tau_+\partial_r^b)^J Y^u|\lesssim 1,\qquad\sum_{i+|J|\leq 1}|(\tau_-\partial_u^b)^i(\tau_x\partial_r^b)^J Y^r|\lesssim 1,\qquad\partial_u^bY^r=0. \tag{55}$$

Then (53)–(54) hold with the right-hand side of (53) replaced by $\mathcal{E}(0,R)=\mathcal{E}(0,R,\chi_{r>R}Y)$, where $\chi_{r>R}$ is as above.

In order to prove this proposition we need a few additional supporting lemmas.

Lemma 4.7 (asymptotics of the conformal potential). *Let $\Omega=\tau_x$ and define the quantity $V=\Omega^3\Box_g(\Omega^{-1})$. Then in Bondi coordinates (u,x^i) one has the symbol bounds*

$$V\in\mathcal{Z}^{-\frac{1}{2}}. \tag{56}$$

Proof. First write the wave operator in Bondi coordinates as $\Box_g=d^{-\frac{1}{2}}\Box_h+d^{-\frac{1}{2}}\partial_\alpha^b\mathcal{R}^{\alpha\beta}\partial_\beta^b$, where $d=|g|$ is the Bondi coordinate metric determinant, $\Box_h=\partial_\alpha^bh^{\alpha\beta}\partial_\beta^b$ where h is given in (4), and where $\mathcal{R}=d^{\frac{1}{2}}g-h$ satisfies the estimate (40). A quick calculation shows $\Box_h(\tau_x^{-1})=-3\tau_x^{-5}$, and a little further work reveals

$$V=-d^{-\frac{1}{2}}(r\partial_\alpha^b\mathcal{R}^{\alpha r}+\tau_x^{-2}(1-3r^2)\mathcal{R}^{rr}+3\tau_x^{-2}).$$

By (39) we have $d^{-\frac{1}{2}}-1\in\mathcal{Z}^0$, so estimate (56) follows from (40). □

Lemma 4.8 (formulas for boundary terms). *Let X^r, X^u be nonnegative and set $X = X^u\partial_u^b + X^r\partial_r^b$. Then if $\Omega = \tau_x$, one has the following pointwise estimate involving the quantity $P(X, \chi, \Omega, \phi)$ defined in (30):*

$$X^u|\tau_x^{-1}\partial(\tau_x\phi)|^2 + X^r|\tau_x^{-1}\partial_x^b(\tau_x\phi)|^2$$
$$\lesssim -P(X, \chi, \Omega, \phi) + o_r(1)\cdot((X^u + \tau_0^2X^r)|\tau_x^{-1}\partial(\tau_x\phi)|^2 + \chi(X^u + X^r)\tau_0^{-\frac{1}{2}}\tau_x^{-2}\phi^2). \quad (57)$$

To prove estimate (57) we need the following elementary result:

Lemma 4.9 (approximate null frame). *Let X and Y_A, $A = 1, 2$, be approximately unit-length vectors in the Minkowski space in the sense that $\sup_\alpha |X^\alpha| \approx 1$ and $\sup_\alpha |Y_A^\alpha| \approx 1$. Suppose that there exists $\epsilon > 0$ such that $\langle X, X\rangle = O(\epsilon^2)$, $\langle X, Y_A\rangle = O(\epsilon)$, and in addition $|\langle Y_A, Y_B\rangle - \delta_{AB}| \ll 1$. Then there exists an exact null frame $\{L, \underline{L}, e_A\}$ with $\langle L, L\rangle = \langle L, e_A\rangle = \langle \underline{L}, L\rangle = 0$, $\langle L, \underline{L}\rangle = -1$, and $\langle e_A, e_B\rangle = \delta_{AB}$, and coefficients γ, c_X^A and c_A^B for $A, B = 1, 2$, such that*

$$X = L + c_X^A e_A + \gamma\underline{L}, \quad Y_A = c_A^B e_B, \quad \textit{where } \gamma = O(\epsilon^2) \textit{ and } c_X^A = O(\epsilon) \textit{ and } |c_B^A - \delta_B^A| \ll 1.$$

Proof. Let e_A form an orthonormal basis for the space-like 2-plane spanned by Y_A, with the first e_A in the direction of one of the Y_A. Let c_A^B be the corresponding change of basis. Then $|c_B^A - \delta_B^A| \ll 1$. Let L, $\underline{L}$ generate the two null directions over the span of e_A and Y_A, chosen so that $\langle L, \underline{L}\rangle = -1$ and $X = L + c_X^A e_A + \gamma\underline{L}$ for some set of coefficients c_X^A, γ. From $\langle X, Y_A\rangle = O(\epsilon)$ we have $\langle X, e_A\rangle = O(\epsilon)$ and so $c_X^A = O(\epsilon)$. Then $\langle X, X\rangle = -2\gamma + O(\epsilon^2)$, so $\gamma = O(\epsilon^2)$ follows from $\langle X, X\rangle = O(\epsilon^2)$. □

Proof of (57). Note that it suffices to prove this bound in the region $r \gg 1$. Consider the vector fields $\{\partial_r^b, Y_A\}$, where Y_A is a (local) Euclidean orthonormal basis on the spheres $r = \text{const}$, $u = \text{const}$. Because the metric g is asymptotically Minkowskian, $|\langle Y_A, Y_B\rangle - \delta_{AB}| \ll 1$. On the other hand a quick application of the asymptotic formulas (3) and Cramer's rule shows that $\langle\partial_r^b, \partial_r^b\rangle = o_r(1)\cdot\tau_0^2$ and $\langle\partial_r^b, Y_A\rangle = o_r(1)\cdot\tau_0$. Thus, an application of the previous lemma shows that

$$\partial_r^b = L + o_r(1)\cdot\tau_0\not\partial_x^b + o_r(1)\cdot\tau_0^2\partial,$$

where L is null, $\not\partial_x^b$ denotes derivatives tangent to $u = \text{const}$, $r = \text{const}$ which are also orthogonal to L, and ∂ is arbitrary.

Next, let $T = T[\psi]$ denote the energy momentum tensor of ψ with respect to the metric g. Because $t = \text{const}$ are uniformly spacelike when $r \gg 1$ we have

$$T(L, -\nabla t) \approx |L\psi|^2 + |\not\nabla_x^b\psi|^2, \quad |T(\not\partial_x^b, -\nabla t)| \lesssim |\not\nabla_x^b\psi|\cdot|\partial\psi| + |g^{\alpha\beta}\partial_\alpha\psi\partial_\beta\psi|, \quad |T(\partial, -\nabla t)| \lesssim |\partial\psi|^2.$$

A quick application of (3) and the middle bound above shows that uniformly for $C > 0$

$$\tau_0|T(\not\partial_x^b, -\nabla t)| \lesssim C^{-1}|\partial_x^b\psi|^2 + C\tau_0^2|\partial\psi|^2.$$

Combining the last three displays gives the pointwise estimate

$$|\partial_x^b\psi|^2 \lesssim T(\partial_r^b, -\nabla t) + o_r(1)\cdot\tau_0^2|\partial\psi|^2.$$

In addition to this we also have by the asymptotic flatness of g and standard properties of T

$$|\partial\psi|^2 \lesssim T(\partial_u^b, -\nabla t) + o_r(1)\cdot|\partial\psi|^2.$$

Finally, let $P(X,\chi,\Omega,\phi)$ denote the quantity defined in (30). Then

$$-P(X,\chi,\Omega,\phi) = \Omega^{-2}T[\Omega^{-1}\phi](X,-\nabla t) - \tfrac{1}{2}\Omega^{-2}X^0\chi V\phi^2, \quad \text{where } V = \Omega^3\Box_g(\Omega^{-1}),$$

and where X^0 denotes the time component of X in (t,x)-coordinates. By (1) we have $|X^0| \lesssim X^u + X^r$. Therefore estimate (57) follows from the last three displays above and estimate (56). □

We now return to the proof of the main result of this subsection. Because of the split form of the error terms (49) and (51) there are essentially two cases.

Proof of Proposition 4.6 for $0 < a < 1$. We use the formalism of Section 2, in particular Lemma 2.4. Let $X = X_R = \chi_{>R}Y$, where $\chi_{>R}$ is as in the statement of the proposition. We choose $\Omega = \tau_x$ and set the auxiliary cutoff to $\chi = 0$. Using the divergence identity (29) we need to estimate each spacetime term given by formulas (31) and (32), as well as the boundary term on the right-hand side of (29).

Step 1: (*output of the* $A^{\alpha\beta}$*-contraction*) We'll do this calculation by switching over to polar Bondi coordinates (u,r,x^A), where locally we can choose x^A to be two members of $\hat{x}^i = r^{-1}x^i$. From (34) and (35), and expansion of $\mathcal{L}_X h$, we may write $A^{\alpha\beta} = \chi_{>R}d^{-\frac{1}{2}}A_0^{\alpha\beta} + d^{-\frac{1}{2}}\mathcal{R}^{\alpha\beta}$, where

$$2A_0^{\alpha\beta} = \partial_\gamma^b(Y^\alpha)h^{\gamma\beta} + \partial_\gamma^b(Y^\beta)h^{\alpha\gamma} - Y(h^{\alpha\beta}) - (\partial_u^b Y^u + \partial_r^b Y^r)h^{\alpha\beta}, \tag{58}$$

and $\mathcal{R} = \mathcal{R}_0 + \chi_{>R}\mathcal{R}_1$. Here $2\mathcal{R}_0 = {}^{(X_R)}\hat{\pi} - \chi_{>R}{}^{(Y)}\hat{\pi}$, which according to formula (26) is

$$2\mathcal{R}_0^{\alpha\beta} = \tau_x^{-1}\chi_R(g^{\alpha r}Y^\beta + g^{\beta r}Y^\alpha - Y^r g^{\alpha\beta}), \tag{59}$$

and where $\chi_R = \tau_x\partial_r\chi_{>R}$ is a smooth bump function adapted to $r \approx R$. The second remainder term is

$$2\mathcal{R}_1 = -\mathcal{L}_Y(d^{\frac{1}{2}}g^{-1} - h) - (\partial_u^b Y^u + \partial_r^b Y^r)(d^{\frac{1}{2}}g^{-1} - h) - 2r^{-1}\tau_x^{-2}d^{\frac{1}{2}}X^r g^{-1}. \tag{60}$$

A little further computation shows the coefficients of the quadratic form $A_0^{\alpha\beta}$ from (58) are

$$A_0^{uu} = \mathcal{A}^u + \tau_x^{-1}(r+\tau_x)^{-1}\partial_r^b Y^u, \qquad A_0^{rr} = \mathcal{A}^r + \tau_x^{-1}(r+\tau_x)^{-1}\partial_u^b Y^r, \tag{61a}$$

$$2A_0^{ur} = \mathcal{A}^{ur} + \tau_x^{-3}Y^r, \qquad A^{AB} = r^{-2}\delta^{AB}\mathcal{A}\!\!\!/, \tag{61b}$$

while $A_0^{rA} = A_0^{uA} = 0$. Here the terms $\mathcal{A}$ are given in (54). Recalling the definitions of the norms (10b) and using the conditions (52) we have

$$\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}} \chi_{>R}\big(\mathcal{A}^u(\partial_u^b\phi)^2 + \mathcal{A}^{ur}\partial_u^b\phi\cdot\tau_x^{-1}\partial_r^b(\tau_x\phi) + \mathcal{A}^r(\tau_x^{-1}\partial_r^b(\tau_x\phi))^2 + \mathcal{A}\!\!\!/|\nabla\!\!\!\!/^b\phi|^2\big)\,dx\,dt$$
$$\leq \int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}} \chi_{>R}\tau_x^{-2}d^{-\frac{1}{2}}A_0^{\alpha\beta}\partial_\alpha^b(\tau_x\phi)\partial_\beta^b(\tau_x\phi)\,dV_g + o_R(1)\|\phi\|^2_{S^a(r>\frac{1}{2}R)[0,T]}. \tag{62}$$

To estimate the remainder terms from (59), note that a straightforward calculation involving the conditions (52) and (3) gives in rectangular Bondi coordinates

$$|\mathcal{R}_0^{ij}| \lesssim (\tau_x^{2a-1} + \tau_+^{2a-1})\chi_R, \quad |\mathcal{R}_0^{ui}| \lesssim \tau_x^{-1}\tau_+^{2a}\tau_0\chi_R, \quad |\mathcal{R}_0^{uu}| \lesssim \tau_x^{-1}\tau_+^{2a}\tau_0^{\max\{1,2a\}}\chi_R, \tag{63}$$

where χ_R is supported in $\frac{1}{2}R < r < R$. This yields the estimate

$$\int_0^T\int |\tau_x^{-2}\mathcal{R}_0^{\alpha\beta}\partial_\alpha(\tau_x\phi)\partial_\beta(\tau_x\phi)|\,dx\,dt \lesssim R^{-\frac{1}{2}}\|\phi\|_{S^a(\frac{1}{2}R<r<R)[0,T]}\|(\tau_-^a\partial_u^b\phi,\,\tau_x^a\partial\phi,\,\tau_x^{a-1}\phi)\|_{L^2(\frac{1}{2}R<r<R)[0,T]}. \tag{64}$$

To estimate the remainder terms from (60) we split the range into $0 < a < \frac{1}{2}$ and $\frac{1}{2} \leq a < 1$. When $0 < a < \frac{1}{2}$ the conditions (52) imply

$$\sum_{i+|J|\leq 1} |(\tau_-\partial_u^b)^i(\tau_+\partial_x^b)^J Y^u| \lesssim \tau_x^{2a-1}\cdot\tau_-, \qquad \sum_{i+|J|\leq 1} |(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J Y^r| \lesssim \tau_x^{2a-1}\cdot\tau_x.$$

In addition we must assume $\partial_r^b Y^u \neq 0$ and $\partial_u^b Y^r \neq 0$, although we do have $\partial_u^b(Y^i - \hat{x}^i Y^r) = 0$. Therefore by simultaneously combining cases (II) and (III) of Lemma 4.3 and Remark 4.4, using $|\partial_u^b Y^u + \partial_r^b Y^r| \lesssim \tau_x^{2a-1}$ $\big($again for $0 < a < \frac{1}{2}\big)$ and (40), and directly using (3) for the last term on the right-hand side of (60) we have for $0 < a < \frac{1}{2}$

$$|||w_a^{-1}\mathcal{R}_1^{rr}|||_{-\frac{1}{2},0} < \infty, \qquad |||w_a^{-1}\mathcal{R}_1^{uu}|||_{\max\{1,2a\},0} < \infty, \quad |||w_a^{-1}\mathcal{R}_1^{ru}|||_{\frac{1}{2},0} < \infty, \tag{65a}$$

$$|||w_a^{-1}r\mathcal{R}_1^{uA}|||_{1,0} < \infty, \quad |||w_a^{-1}(r\mathcal{R}_1^{rA}, r^2\mathcal{R}_1^{AB})|||_{0,0} < \infty, \quad \text{where } w_a = (\tau_x^{2a-1} + \tau_+^{2a-1}). \tag{65b}$$

On the other hand in the range $\frac{1}{2} \leq a \leq 1$ the conditions (52) imply

$$\sum_{i+|J|\leq 1} \|(\tau_-\partial_u^b)^i(\tau_+\partial_x^b)^J Y^u| \lesssim \tau_-^{2a-1}\cdot\tau_-, \qquad \sum_{i+|J|\leq 1} |(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J Y^r| \lesssim \tau_+^{2a-1}\cdot\tau_x.$$

Therefore by separately applying cases (II) and (III) of Lemma 4.3 and Remark 4.4 to the vector fields $Y^u\partial_u^b$ and $Y^r\partial_r^b$ respectively, and this time using $|\partial_u^b Y^u| + |\partial_r^b Y^r| \lesssim \tau_+^{2a-1}$, we again have (65). Finally, after several rounds of Hölder's inequality and a straightforward check of definitions (2) and (10b), the error bounds (65) yield the asymptotic estimate

$$\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}} \chi_{>R}|\tau_x^{-2}\mathcal{R}_1^{\alpha\beta}\partial_\alpha(\tau_x\phi)\partial_\beta(\tau_x\phi)|\,dx\,dt \lesssim o_R(1)\|\phi\|^2_{S^a(r>\frac{1}{2}R)[0,T]}. \tag{66}$$

As a last step we combine estimates (62), (64), and (66), while recalling that $d^{\frac{1}{2}}A^{\alpha\beta} = \chi_{>R}A_0^{\alpha\beta} + \mathcal{R}_0 + \chi_{>R}\mathcal{R}_1$. This gives us

$$\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}} \chi_{>R}\big(\mathcal{A}^u(\partial_u^b\phi)^2 + \mathcal{A}^{ur}\partial_u^b\phi\cdot\tau_x^{-1}\partial_r^b(\tau_x\phi) + \mathcal{A}^r(\tau_x^{-1}\partial_r^b(\tau_x\phi))^2 + \mathcal{A}\!\!\!/|\nabla\!\!\!\!/^b\phi|^2\big)\,dx\,dt \leq \int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}} \tau_x^{-2}A^{\alpha\beta}\partial_\alpha(\tau_x\phi)\partial_\beta(\tau_x\phi)\,dV_g + \mathcal{E}(a,R), \tag{67}$$

where the coefficients $\mathcal{A}$ are given by (54).

Step 2: (*estimating the C^χ-term*) Using (56) we have $C^\chi \in \tau_x^{-2}\mathcal{Z}^{-\frac{1}{2}}$, while (52) give the pointwise estimate $|\tau_x^{-1}X_R(\tau_x\phi)| \lesssim \chi_{>R}\tau_+^{2a}(\tau_0^{\max\{1,2a\}}|\partial_u^b\phi| + |\tau_x^{-1}\partial_r^b(\tau_x\phi)|)$. Thus Hölder's inequality and (10b) give

$$\Big|\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}} C^\chi\phi\tau_x^{-1}X_R(\tau_x\phi)\,dV_g\Big| \lesssim o_R(1)\|\phi\|^2_{S^a[0,T](r>\frac{1}{2}R)}. \tag{68}$$

Step 3: (*output of the boundary terms*) Using the bound from (57) we directly have

$$\|\tau_x^{-1}(\sqrt{Y^u}\partial(\tau_x\phi), \sqrt{Y^r}\partial_x^b(\tau_x\phi))(T)\|^2_{L_x^2(r>R)}$$
$$\lesssim \|\tau_x^a\partial\phi(0)\|^2_{L_x^2} + o_R(1)\sup_{0\le t\le T}\|\phi(t)\|^2_{E^a(r>\frac{1}{2}R)} - \int_{\mathbb{R}^3} P(X,\chi,\Omega,\phi)\sqrt{|g|}\,dx\Big|_{t=0}^{t=T}. \tag{69}$$

Step 4: (*output of the source term*) Finally, another application of Hölder's inequality shows

$$\Big|\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}} \Box_g\phi\cdot\tau_x^{-1}X_R(\tau_x\phi)\,dV_g\Big| \lesssim \|\Box_g\phi\|_{N^a[0,T]}\|\phi\|_{S^a(r>\frac{1}{2}R)[0,T]}. \tag{70}$$

Adding together formulas (67)–(70) and using (29), (31), and (32) gives (53). □

For the proof of Proposition 4.6 with $a=1$ it will help to have the following lemma:

Lemma 4.10. *Let Φ be a bounded function supported on $[0,T]$, and let $|||p^2|||_{0,0} < \infty$. Then for $R>0$ one has the estimate*

$$\|\sqrt{\tau_x/\tau_+}p\cdot\Phi\|_{L^2(r>R)} \lesssim o_R(1)\big(\|(1-\chi_{\frac{1}{2}t<r<2t})\Phi\|_{\ell_t^\infty\ell_r^\infty L^2} + \|\chi_{\frac{1}{2}t<r<2t}\Phi\|_{\ell_t^\infty\ell_u^\infty\ell_r^\infty L^2} + \|\tau_x^{\frac{1}{2}}\Phi\|_{L_t^\infty L_x^2}\big). \tag{71}$$

Proof of (71). We split the right-hand side into regions $r<\frac{1}{2}t$, $\frac{1}{2}t<r<2t$, and $r>2t$, all restricted to $r>R$ (we will largely suppress this last condition in the following notation).

In the first region we use

$$\|\sqrt{\tau_x/\tau_+}p\cdot\Phi\|_{L^2(r<\frac{1}{2}t)} \lesssim \|\sqrt{\tau_x/\tau_+}p\|_{\ell_t^2\ell_r^2L^\infty(r<\frac{1}{2}t)}\|\Phi\|_{\ell_t^\infty\ell_r^\infty L^2(r<\frac{1}{2}t)}$$

followed by Young's inequality, which gives $\|\sqrt{\tau_x/\tau_+}p\|_{\ell_t^2\ell_r^2L^\infty(r<\frac{1}{2})} \lesssim \|p\|_{\ell_r^2L^\infty(r>R)} = o_R(1)$.

In the region $\frac{1}{2}t<r<2t$ we use

$$\|\sqrt{\tau_x/\tau_+}p\cdot\Phi\|_{L^2(\frac{1}{2}t<r<2t)} \lesssim \|p\|_{\ell_u^2\ell_r^2L^\infty(\frac{1}{2}t<r<2t)}\|\Phi\|_{\ell_t^\infty\ell_u^\infty\ell_r^\infty L^2(\frac{1}{2}t<r<2t)},$$

followed by $\|p\|_{\ell_u^2\ell_r^2L^\infty(\frac{1}{2}t<r<2t)(r>R)} = o_R(1)$.

Finally, in the region $r>2t$ we use

$$\|\sqrt{\tau_x/\tau_+}p\cdot\Phi\|_{L^2(r>2t)} \lesssim \|\tau_x^{-\frac{1}{2}}p\|_{L_t^2L_x^\infty(r>2t)}\|\tau_x^{\frac{1}{2}}\Phi\|_{L_t^\infty L_x^2},$$

followed by $\|\tau_x^{-\frac{1}{2}}p\|_{L_t^2L_x^\infty(r>2t)} \lesssim \|\chi_{t<\frac{1}{2}r}\tau_x^{-\frac{1}{2}}p\|_{\ell_r^2L_t^2L_x^\infty} \lesssim \|p\|_{\ell_r^2L^\infty(r>R)} = o_R(1)$. □

Proof of Proposition 4.6 for $a=1$. The demonstration is largely similar to the previous proof, with a few key differences. We again choose $\Omega=\tau_x$, but this time set the auxiliary cutoff in Lemma 2.4 to be $\chi=\chi_{<\frac{1}{2}}(r/t)$ which vanishes in $r>\frac{3}{4}t$ with $\chi\equiv 1$ when $r<\frac{1}{2}t$.

Step 1: (*output of the $A^{\alpha\beta}$-contraction*) We again have formulas (58)–(61).

An inspection of the remainder terms on the right-hand side of (61) using condition (52) shows that we also have estimate (62) with the last term on the right-hand side replaced by $o_R(1)\|\phi\|^2_{S^{1,\infty}(r>\frac{1}{2}R)[0,T]}$.

Next, the estimates (63) are again valid except this time we use Hölder's inequality to replace (64) with

$$\int_0^T\int |\tau_x^{-2}\mathcal{R}_0^{\alpha\beta}\partial_\alpha(\tau_x\phi)\partial_\beta(\tau_x\phi)|\,dx\,dt \lesssim R^{\frac{1}{2}}\|\phi\|_{S^{1,\infty}(\frac{1}{2}R<r<R)[0,T]}\|(\tau_-\partial_u^b\phi,\,\tau_x\partial\phi,\,\phi)\|_{\ell_t^1L^2(\frac{1}{2}R<r<R)[0,T]}. \quad (72)$$

To estimate the remainder term involving $\mathcal{R}_1^{\alpha\beta}$ note that (65) is still valid with $a=1$. This allows us to replace (66) in the case $a=1$ with the slight improvement

$$\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}}\chi_{>R}|\tau_x^{-2}\mathcal{R}_1^{\alpha\beta}\partial_\alpha(\tau_x\phi)\partial_\beta(\tau_x\phi)|\,dx\,dt \lesssim \|\sqrt{\tau_x/\tau_+}p\cdot\tau_x^{-\frac{1}{2}}\tau_+(\tau_0\partial\phi,\,\partial_x^b\phi,\,\tau_0^{-\frac{1}{2}}\tau_x^{-1}\partial_r^b(\tau_x\phi),\,\tau_x^{-1}\phi)\|^2_{L^2(r>\frac{1}{2}R)[0,T]}, \quad (73)$$

where $|||p^2|||_{0,0}<\infty$. Then an application of Lemma 4.10 produces

$$\text{LHS (73)} \lesssim o_R(1)\cdot\big(\sup_{0\le t\le T}\|\phi(t)\|^2_{E^1(r>\frac{1}{2}R)}+\|\phi\|^2_{S^{1,\infty}(r>\frac{1}{2}R)[0,T]}\big). \quad (74)$$

Combining (72), (74), and the analog of (62), we have estimate (67) for the case $a=1$.

Step 2: (*estimating the B^χ-term*) Inspection of (32) and estimate (56) shows the main thing is to compute trace(A). By definition trace ${}^{(X_R)}\hat{\pi}=-2\nabla_\alpha X_R^\alpha$, so trace$(A)=4r\tau_x^{-2}X_R^r-\nabla_\alpha X_R^\alpha$. Using the conditions (52) and the support of χ, we conclude $|||\tau_x^2\tau_+^{-1}B^\chi|||_{0,0}<\infty$. Thus

$$\Big|\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}}B^\chi\phi^2\,dV_g\Big| \lesssim \|\sqrt{\tau_x/\tau_+}p\cdot\tau_x^{-\frac{1}{2}}\tau_+(\tau_x^{-1}\phi)\|^2_{L^2(r>\frac{1}{2}R)[0,T]}, \quad (75)$$

where $|||p^2|||_{0,0}<\infty$. From this Lemma 4.10 produces

$$\text{LHS (75)} \lesssim o_R(1)(\|\phi\|^2_{E^1(r>\frac{1}{2}R)[0,T]}+\|\phi\|^2_{S^{1,\infty}(r>\frac{1}{2}R)[0,T]}).$$

Step 3: (*estimating the C^χ-term*) Using (56) and the support property of χ we have $\tau_+^2C^\chi\in\mathcal{Z}^{-\frac{1}{2}}$, while (52) give the pointwise bound $|\tau_x^{-1}X_R(\tau_x\phi)|\lesssim\chi_{>R}\tau_+^2(\tau_0^2|\partial_u^b\phi|+|\tau_x^{-1}\partial_r^b(\tau_x\phi)|)$. Thus

$$\Big|\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}}C^\chi\phi\tau_x^{-1}X_R(\tau_x\phi)\,dV_g\Big| \lesssim \|\sqrt{\tau_x/\tau_+}p\cdot\tau_x^{-\frac{1}{2}}\tau_+(\tau_0\partial\phi,\,\tau_0^{-\frac{1}{2}}\tau_x^{-1}\partial_r^b(\tau_x\phi),\,\tau_x^{-1}\phi)\|^2_{L^2(r>\frac{1}{2}R)[0,T]},$$

where again $p^2\in\mathcal{Z}^0$ so we conclude via Lemma 4.10.

Step 4: (*output of the boundary terms*) Here we simply note that (69) is also valid for $a=1$.

Step 5: (*output of the source term*) This term is included directly in the definition of $\mathcal{E}(1,R,X)$. □

Proof of Proposition 4.6 for $a=0$. This follows the pattern of the previous two proofs. We set $\Omega=\tau_x$ and choose $\chi=0$.

Step 1: (*output of the $A^{\alpha\beta}$-contraction*) We again have formulas (58)–(61). This time one replaces (63) with $|\mathcal{R}_0^{\alpha\beta}| \lesssim \tau_x^{-1}\chi_R$, in which case the right-hand side of (64) becomes

$$R^{-\frac{1}{2}}\|(\partial\phi,\, \tau_x^{-1}\phi)\|^2_{L^2(\frac{1}{2}R<r<R)[0,T]}.$$

The analog of (62) is also valid with the second term on the right-hand side replaced by

$$o_R(1)\|\phi\|^2_{LE^0(r>\frac{1}{2}R)[0,T]}.$$

The main difference is that this time the conditions (55) give $|\partial_u^b X^u + \partial_r^b X^r| \lesssim \tau_x^{-1}\tau_0^{-1}$. Together with the condition $\partial_u^b Y^r = 0$ and an application of Lemma 4.3, this means we need to replace (65) with

$$|||\tau_x\mathcal{R}_1^{ij}|||_{-1,0} < \infty, \quad |||\tau_x\mathcal{R}_1^{ui}|||_{-\frac{1}{2},0} < \infty, \quad |||\tau_x\mathcal{R}_1^{uu}|||_{\frac{1}{2},0} < \infty.$$

This is enough to show the analog of (66) with right-hand side replaced by $o_R(1)\|\phi\|^2_{LE^0(r>\frac{1}{2}R)[0,T]}$.

Step 2: (*estimating the C^X-term*) Using (56) gives $C^X \in \tau_x^{-2}\mathcal{Z}^{-\frac{1}{2}}$, while (55) gives $|\tau_x^{-1}X_R(\tau_x\phi)| \lesssim \chi_{r>R}|(\partial\phi,\, \tau_x^{-1}\phi)|$. Thus (68) is valid with right-hand side replaced by $o_R(1)\|\phi\|^2_{LE^0(r>\frac{1}{2}R)[0,T]}$.

Step 4: (*output of the boundary terms*) Here we note that (69) is also valid for $a = 0$.

Step 5: (*output of the source term*) This term is included directly in the definition of $\mathcal{E}(0, R, X)$. □

4C. *Abstract bounds for commutators.* We now turn to some further consequences of Lemma 4.3.

Lemma 4.11 (abstract bounds for commutators). *Let $\mathcal{R}$ and $\mathcal{S}$ be a contravariant 2-tensor and vector field respectively. From them define the operator $\mathcal{Q} = \nabla_\alpha\mathcal{R}^{\alpha\beta}\nabla_\beta + \mathcal{S}^\alpha\nabla_\alpha$. Then the following results hold*:

(I) *Suppose that $\mathcal{R}$ and $\mathcal{S}$ satisfy (40) and (46) respectively. Then if X is any vector field which satisfies (41) and (42) with $a = b = c = 0$, one has $[X, \mathcal{Q}] = \nabla_\alpha\widetilde{\mathcal{R}}^{\alpha\beta}\nabla_\beta + \widetilde{\mathcal{S}}^\alpha\nabla_\alpha$, where $\widetilde{\mathcal{R}}$ and $\widetilde{\mathcal{S}}$ satisfy (40) and (46) respectively as well.*

(II) *Alternatively suppose $X \in \mathbb{L}_0 = \{\partial_u^b,\, \partial_i^b - \omega^i\partial_u^b\}$ with the same conditions on $\mathcal{R}$ and $\mathcal{S}$ as above. Then the previous result holds with bound (43) for $\widetilde{\mathcal{R}}$ and bound (47) for $\widetilde{\mathcal{S}}$ with $a = c = -1$ and $b = 1$.*

(III) *For any $\mathcal{Q}$ as defined above with $\mathcal{R}$ and $\mathcal{S}$ satisfying (40) and (46) we have the pointwise bound*

$$|\mathcal{Q}\phi| \lesssim q\cdot\tau_x^{-2}\sum_{1\le l+|J|\le 2}|(\tau_x\tau_0\partial_u^b)^l(\tau_x\partial_x^b)^J\phi|, \quad \text{where } q \in \mathcal{Z}^{-\frac{1}{2}}. \tag{76}$$

Alternatively, suppose $\mathcal{R}$ and $\mathcal{S}$ satisfy (43) and (47) with $a = c = -1$ and $b = 1$. Then

$$|\mathcal{Q}\phi| \lesssim q\cdot\tau_x^{-1}\sum_{|I|\le 1}(\tau_0^{\frac{1}{2}}|\partial_u^b\Gamma^I\phi| + |\partial_x^b\Gamma^I\phi|), \quad \text{where } q \in \mathcal{Z}^{-1}, \tag{77}$$

and where $\Gamma \in \mathbb{L}_0 = \{\partial_u^b,\, \partial_i^b - \omega^i\partial_u^b\}$.

(IV) *Finally, let $\mathcal{R}$ satisfy any combination of conditions (43), (44), and (45), and let $\mathcal{S}$ satisfy (47). Let w be a smooth weight function with*

$$|(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J w| \lesssim_{I,J} \tau_x^{a'}\tau_+^{b'}\tau_-^{c'}. \tag{78}$$

Then with $\mathcal{Q}$ as defined above we have $w\mathcal{Q} = \nabla_\alpha \widetilde{\mathcal{R}}^{\alpha\beta}\nabla_\beta + \widetilde{\mathcal{S}}^\alpha \nabla_\alpha$, where $\widetilde{\mathcal{R}}^{\alpha\beta}$ satisfies the appropriate combination of conditions (43), (44), and (45), and $\widetilde{\mathcal{S}}^\alpha$ satisfies (47), in each case with coefficients $a+a'$, $b+b'$, and $c+c'$.

Proof. We'll show each part separately.

Part 1: (*the commutator property for X satisfying (41) and (42)*) By formula (27) and parts (I) and (IV) of Lemma 4.3, it suffices to show that $\mathcal{R}^{\alpha\beta}\nabla_\alpha(\nabla_\gamma X^\gamma)$ satisfies (46). We'll do this in a bit more generality here for use in the sequel.

Let w be any weight function which satisfies (78), and let $\mathcal{R}$ be any contravariant 2-tensor which satisfies the weaker conditions (44) and (45), and the remaining conditions in (43). Then we claim (47) holds for $\mathcal{S}^\alpha = \mathcal{R}^{\alpha\beta}\nabla_\alpha(w)$ with weights $a+a'$, $b+b'$, $c+c'$. To see this note

$$\mathcal{S}^i = \mathcal{R}^{ui}\partial_u^b w + \mathcal{R}^{ji}\partial_j^b(w), \qquad \mathcal{S}^u = \mathcal{R}^{uu}\partial_u^b w + \mathcal{R}^{ju}\partial_j^b(w),$$

so the desired estimates follow easily by multiplying together bounds (78) and the appropriate combination of (43), (44), and (45).

To show part (I) of the lemma note that if X is as stated, then (39) shows $w = \nabla_\gamma X^\gamma$ satisfies (78) with $a' = b' = c' = 0$. The desired result now follows from the discussion of the previous paragraph.

Part 2: (*the commutator property for $X \in \mathbb{L}_0$*) This follows at once from part (V) of Lemma 4.3 and the main calculation of Part 1 above. Note that if $X \in \mathbb{L}_0$ then $w = \nabla_\gamma X^\gamma$ satisfies (78) with $a' = c' = -1$ and $b' = 1$.

Part 3: (*the pointwise estimates (76) and (77)*) The bound for the $\mathcal{S}$ potion of $\mathcal{Q}\phi$ follows at once from (47). For the $\mathcal{R}$-contraction we write in Bondi coordinates

$$\nabla_\alpha \mathcal{R}^{\alpha\beta}\nabla_\beta\phi = \mathcal{R}^{\alpha\beta}\partial_\alpha^b(\ln\sqrt{|g|})\partial_\beta^b\phi + (\partial_\alpha^b \mathcal{R}^{\alpha\beta})\partial_\beta^b\phi + \mathcal{R}^{\alpha\beta}\partial_\alpha^b\partial_\beta^b\phi = \widetilde{S}^\alpha\partial_\alpha^b u + \mathcal{R}^{\alpha\beta}\partial_\alpha^b\partial_\beta^b\phi.$$

We only need to show that $\widetilde{S}$ satisfies (47) with $a=b=c=0$ in the case of estimate (76), and $a=c=-1$, $b=1$ in the case of estimate (76); then study $\mathcal{R}^{\alpha\beta}\partial_\alpha^b\partial_\beta^b\phi$.

For the first term of $\widetilde{S}$ we use the fact that $w = \ln\sqrt{|g|}$ satisfies (78) with $a' = b' = c' = 0$, which follows from (39). Then by the main calculation of Part 1 above we have (47) for $\mathcal{R}^{\alpha\beta}\partial_\alpha^b(\ln\sqrt{|g|})$.

For the expression $\partial_\alpha^b\mathcal{R}^{\alpha\beta}$ the appropriate version of (47) follows at once from (43).

For the final term note that if $\mathcal{R}$ satisfies (43) then one has the pointwise estimate

$$|\mathcal{R}^{\alpha\beta}\partial_\alpha^b\partial_\beta^b\phi| \lesssim q\cdot\tau_x^a\tau_+^b\tau_-^c(\tau_0^{\frac{5}{2}}|(\partial_u^b)^2\phi| + \tau_0|\partial_u^b\partial_x^b\phi| + \tau_0^{\frac{1}{2}}|(\partial_x^b)^2\phi|), \quad \text{where } q \in \mathcal{Z}^{-\frac{1}{2}}.$$

This is bounded by the right-hand side of (76) when $a=b=c=0$, and the right-hand side of (77) when $a=c=-1$ and $b=1$.

Part 4: (*proof of the algebra property (IV)*) It is immediate that bound (47) for $\mathcal{S}$ is stable under multiplication by w satisfying (78) with the appropriate change of weights. For the quadratic term of $\mathcal{Q}$ we write $w\nabla_\alpha\mathcal{R}^{\alpha\beta}\nabla_\beta = \nabla_\alpha w\mathcal{R}^{\alpha\beta}\nabla_\beta - \mathcal{R}^{\alpha\beta}\nabla_\alpha(w)\nabla_\beta$. For the first term we use bounds (43) which are also stable under multiplication by w. For the second term we use the main calculation of Part 1 above. □

Parts (I) and (II) of the last lemma imply the following:

Corollary 4.12 (estimates for multicommutators). *Let g be a metric which satisfies (3), and as usual set $\mathbb{L}_0 = \{\partial^b_u, \partial^b_i - \omega^i \partial^b_u\}$ and $\mathbb{L} = \{S, \Omega_{ij}\} \cup \mathbb{L}_0$. Then the following hold:*

(I) *If I is any multiindex then for products of vector fields in $\mathbb{L}$ one has the identity*

$$[\Box_g, \Gamma^I] = \sum_{I' \subsetneq I} ([\nabla_\alpha \mathcal{R}^{\alpha\beta}_{I'} \nabla_\beta + \mathcal{S}^\alpha_{I'} \nabla_\alpha]\Gamma^{I'} + w_{I'}\Gamma^{I'}\Box_g), \tag{79}$$

where the sum is taken over the collection of all multiindices I' strictly contained in I; in particular each I' satisfies $|I'| \le |I| - 1$. Here $\mathcal{R}_{I'}$ and $\mathcal{S}_{I'}$ satisfy (40) and (46) respectively, while there exist constants $w^0_{I'} \in \mathbb{R}$ such that

$$w_{I'} - w^0_{I'} \in \mathcal{Z}^0. \tag{80}$$

(II) *If the product in the previous part is restricted to vector fields in $\mathbb{L}_0$, then one has identity (79) with estimates (43) and (47) for $\mathcal{R}_{I'}$ and $\mathcal{S}_{I'}$ with $a = c = -1$ and $b = 1$. In addition (80) in this case is replaced by $w_{I'} \in \tau_x^{-1} \cdot \mathcal{Z}^{-1}$.*

(III) *Let Γ^I be a product of vector fields in $\mathbb{L}$ and $\widetilde{\Gamma}^J$ a product of vector fields in $\mathbb{L}_0$. Then one has the identity*

$$\widetilde{\Gamma}^J[\Box_g, \Gamma^I] = \sum_{I' \subsetneq I,\, J' \subseteq J} ([\nabla_\alpha \mathcal{R}^{\alpha\beta}_{I',J'} \nabla_\beta + \mathcal{S}^\alpha_{I',J'} \nabla_\alpha]\widetilde{\Gamma}^{J'}\Gamma^{I'} + w_{I',J'}\widetilde{\Gamma}^{J'}\Gamma^{I'}\Box_g), \tag{81}$$

where $\mathcal{R}_{I',J'}$, $\mathcal{S}_{I',J'}$, and $w_{I',J'}$ satisfy respectively (40), (46), and (80).

Proof. We'll show the different parts separately.

Part 1: (*proof of* (79) *for* $\mathbb{L}$ *and* $\mathbb{L}_0$) Here we will focus only on the case of products of vector fields in $\mathbb{L}$, as the case of $\mathbb{L}_0$ is completely analogous. Using the algebra property of Part (IV) of Lemma 4.11 and an induction, it suffices to show

$$[\Box_g, \Gamma^I] = \sum_{I' \subsetneq I} [\nabla_\alpha \mathcal{R}^{\alpha\beta}_{I'} \nabla_\beta + \mathcal{S}^\alpha_{I'} \nabla_\alpha + w_{I'}\Box_g]\Gamma^{I'}, \quad \text{where } \Gamma^0 = \mathrm{Id}. \tag{82}$$

We shall prove this last bound itself by induction on the length $|I|$ of the product Γ^I.

Case 1: ($|I| = 1$) When Γ^I consists of a single vector field in $\mathbb{L}$, formula (82) follows from a combination of formulas (36) and (37), followed by estimates (39), (40), and part (I) of Lemma 4.3.

Case 2: ($|I| \ge 2$) Assume formula (82) holds for all multiindices $|I| < k$ and choose some $|I| = k$, and write $\Gamma^I = \Gamma^{I_0}\Gamma^{I_1}$ for some $|I_0| = 1$. By the Leibniz rule we have $[\Box_g, \Gamma^I] = [\Box_g, \Gamma^{I_0}]\Gamma^{I_1} + \Gamma^{I_0}[\Box_g, \Gamma^{I_1}]$. By the same calculations as in the previous step we have

$$[\Box_g, \Gamma^{I_0}]\Gamma^{I_1} = \nabla_\alpha \mathcal{R}^{\alpha\beta} \nabla_\beta \Gamma^{I_1} + w\Box_g \Gamma^{I_1},$$

where $\mathcal{R}$, w are of the desired form. On the other hand by induction we have

$$\Gamma^{I_0}[\Box_g, \Gamma^{I_1}] = \sum_{I' \subsetneq I_1} (\nabla_\alpha \mathcal{R}^{\alpha\beta}_{I'} \nabla_\beta + \mathcal{S}^\alpha_{I'} \nabla_\alpha + w_{I'}\Box_g)\Gamma^{I_0}\Gamma^{I'} + \sum_{I' \subsetneq I_1} [\Gamma^{I_0}, (\nabla_\alpha \mathcal{R}^{\alpha\beta}_{I'} \nabla_\beta + \mathcal{S}^\alpha_{I'} \nabla_\alpha + w_{I'}\Box_g)]\Gamma^{I'}.$$

By formula (27) and Part (I) of Lemma 4.11 the commutator $[\Gamma^{I_0}, (\nabla_\alpha \mathcal{R}^{\alpha\beta}_{I'} \nabla_\beta + \mathcal{S}^\alpha_{I'} \nabla_\alpha)]$ again yields an operator of the form $\mathcal{Q}_{I'} = \nabla_\alpha \widetilde{\mathcal{R}}^{\alpha\beta}_{I'} \nabla_\beta + \widetilde{\mathcal{S}}^\alpha_{I'} \nabla_\alpha$. Using (39), the same calculations of the previous step, and part (IV) of Lemma 4.11, we see the commutator $[\Gamma^{I_0}, w_{I'} \Box_g]$ yields another such operator. Combining all this yields (82).

Part 2: (*proof of* (81)) Applying $\widetilde{\Gamma}^J$ to formula (79) and then computing $[\widetilde{\Gamma}^J, (\nabla_\alpha \mathcal{R}^{\alpha\beta}_{I'} \nabla_\beta + \mathcal{S}^\alpha_{I'} \nabla_\alpha)]$ through a repeated use of part (II) of Lemma 4.11 yields the desired result. □

4D. *Klainerman–Sideris inequalities.* We now prove an analog of the Klainerman–Sideris identity [1996].

Lemma 4.13 (Klainerman–Sideris-type identity). *One has the pointwise estimates*

$$\sum_{1\leq l+|J|\leq 2} |(\tau_x \tau_0 \partial^b_u)^l (\tau_x \partial^b_x)^J \phi| \lesssim \sum_{\substack{l+|J|=1 \\ |I|\leq 1}} |(\tau_x \tau_0 \partial^b_u)^l (\tau_x \partial^b_x)^J \Gamma^I \phi| + \tau_x^2 \tau_0 |\Box_g \phi|, \tag{83}$$

$$\sum_{1\leq l+|J|\leq k} |(\tau_- \partial^b_u)^l (\tau_x \partial^b_x)^J \phi| \lesssim \sum_{\substack{l+|J|=1 \\ |I|\leq k-1}} |(\tau_- \partial^b_u)^l (\tau_x \partial^b_x)^J \Gamma^I \phi| + \sum_{l+|J|\leq k-2} \tau_x^2 \tau_0 |(\tau_- \partial^b_u)^l (\tau_x \partial^b_x)^J \Box_g \phi|, \tag{84}$$

where in the second bound the implicit constant depends on $k \geq 2$. Here all $\Gamma \in \mathbb{L}$.

Proof of estimates (83) *and* (84). Both estimates follow from essentially the same computation.

Step 1: (*a preliminary reduction*) First, note that in either case it suffices to restrict to $r > R \gg 1$, as estimates (83) and (84) with some implicit constant $C = C(R, k)$ are automatic in $r \leq R$ by choosing all $\Gamma \in \mathbb{L}_0 \cup \{S\}$.

Next, let $\Box_\eta$ be the Minkowski wave operator in Bondi coordinates which satisfies

$$\Box_\eta = -2\partial^b_u \partial_r + (\partial^b_r)^2 - 2r^{-1}\partial^b_u + 2r^{-1}\partial^b_r + r^{-2}\sum_{i<j}(\Omega_{ij})^2,$$

$$\Box_\eta = d^{\frac{1}{2}}\Box_g - \partial^b_\alpha \mathcal{R}^{\alpha\beta}\partial^b_\beta + O(r^{-2})\partial^b_x \partial^b_u + O(r^{-3})\partial^b_u,$$

where $d = |g|$ is the metric determinant in Bondi coordinates, and $\mathcal{R} = d^{\frac{1}{2}} g - h$, where h is given in (4) and $\mathcal{R}$ satisfies (40). Thanks to (76) one has

$$\tau_x^2 \tau_0 |\Box_\eta \phi| \lesssim \tau_x^2 \tau_0 |\Box_g \phi| + o_R(1) \cdot \sum_{1\leq l+|J|\leq 2} |(\tau_x \tau_0 \partial^b_u)^l (\tau_x \partial^b_x)^J \phi| \quad \text{in } r > R.$$

Therefore it suffices to replace $\Box_g$ by $\Box_\eta$ in (83), and also in (84) when proving it for $k = 2$.

Step 2: (*proof of* (83) *and* (84) *for $k = 2$ and $\Box_g$ replaced by $\Box_\eta$*) Start with the two identities

$$r^2 \tau_+^{-1} \partial^b_r S = r^2 u \tau_+^{-1} \partial^b_u \partial^b_r + r^3 \tau_+^{-1} (\partial^b_r)^2 + r^2 \tau_+^{-1} \partial^b_r,$$

$$\tfrac{1}{2} r^2 u \tau_+^{-1} \Box_\eta = -r^2 u \tau_+^{-1} \partial^b_u \partial^b_r + \tfrac{1}{2} r^2 u \tau_+^{-1} (\partial^b_r)^2 - r u \tau_+^{-1} \partial^b_u + r u \tau_+^{-1} \partial^b_r + \tfrac{1}{2} u \tau_+^{-1} \sum_{i<j} (\Omega_{ij})^2.$$

Adding the two operators on the left-hand sides above and applying to ϕ yields

$$|(\tau_x \partial^b_r)^2 \phi| \lesssim \sum_{l+|J|=1,\ |I|\leq 1} |(\tau_x \tau_0 \partial^b_u)^l (\tau_x \partial^b_x)^J \Gamma^I \phi| + \tau_x^2 \tau_0 |\Box_\eta \phi|.$$

Note that by Remark 1.4 we can assume $u + 2r \approx \tau_+$ in $r > R$.

Next, the vector fields S and ∂_u^b alone give the pair of inequalities

$$|(\tau_x\tau_0\partial_u^b)^2\phi| + |(\tau_x\tau_0\partial_u^b)(\tau_x\partial_r^b)\phi| \lesssim |(\tau_x\partial_r^b)^2\phi| + \sum_{l+|J|=1,\,|I|\leq 1} |(\tau_x\tau_0\partial_u^b)^l(\tau_x\partial_x^b)^J\Gamma^I\phi|,$$

$$|(\tau_-\partial_u^b)^2\phi| + |(\tau_-\partial_u^b)(\tau_x\partial_r^b)\phi| \lesssim |(\tau_x\partial_r^b)^2\phi| + \sum_{l+|J|=1,\,|I|\leq 1} |(\tau_-\partial_u^b)^l(\tau_x\partial_x^b)^J\Gamma^I\phi|.$$

Finally, note that all other combinations of derivatives on the left-hand side of (83) (respectively (84) when $k=2$) are automatically controlled by the first sums on right-hand side of (83) (respectively (84) when $k=2$) thanks to the rotation vector fields.

Step 3: (*estimate* (84) *when* $k>2$) It suffices to show (84) assuming it's true for $k-1$. Applying (84) with $k-1$ to $X\phi$, where $X\in\{\tau_-\partial_u^b,\,\tau_x\partial_i^b\}$, and then feeding the results back into (84) with $k=2$ applied to $\Gamma^I\phi$ where $\Gamma\in\mathbb{L}$ and $|I|\leq k-2$, we need to show the commutator estimates

$$\sum_{\substack{|I'|,|I''|=1\\ |I|\leq k-2}} |X^{I'}[\Gamma^I, X^{I''}]\phi| \lesssim \sum_{1\leq|I|\leq k-1} |X^I\phi|,$$

$$\sum_{|I|\leq k-2} \tau_x^2\tau_0|[\Box_g,\Gamma^I]\phi| \lesssim \sum_{1\leq|I|\leq k-1} |X^I\phi|,$$

$$\sum_{\substack{|I|\leq k-3\\ |I'|=1}} \tau_x^2\tau_0|X^I[\Box_g, X^{I'}]\phi| \lesssim \sum_{1\leq|I|\leq k-1} |X^I\phi|,$$

where each X^I, $X^{I'}$, and $X^{I''}$ is a product of members of $\{\tau_-\partial_u^b,\,\tau_x\partial_i^b\}$. The validity of these last three bounds is easily checked by using (79) and (76) to evaluate $[\Box_g,\Gamma^I]\phi$, and by referring to the following lemma. □

Lemma 4.14 (products of "standard" vector fields). *Let X^I and Y^J be products of vector fields whose Bondi coordinate coefficients satisfy*

$$|(\tau_-\partial_u^b)^l(\tau_x\partial_x^b)^J X^u| \lesssim_{l,J} \tau_-, \qquad |(\tau_-\partial_u^b)^l(\tau_x\partial_x^b)^J X^i| \lesssim_{l,J} \tau_x,$$

with the convention a product of length zero is a scalar satisfying bounds of the form (78) *with $a=b=c=0$. Then the following hold:*

(I) *The vector field $[X^I, Y^J]$ is a sum of products of similar vector fields each with word length $|I|+|J|-1$.*

(II) *For any nonzero multiindex I we have*

$$|X^I\phi| \lesssim \sum_{1\leq l+|J|\leq|I|} |(\tau_-\partial_u^b)^l(\tau_x\partial_x^b)^J\phi|.$$

(III) *If X^I is any product of such vector fields then*

$$\tau_x^2\tau_0 X^I\tau_x^{-2}\tau_0^{-1} - X^I = \sum_{|J|\leq|I|-1} Y^J$$

for some other collection of products of similar vector fields Y^I.

(IV) *One has* $\tau_x^2\tau_0\Box_g = \sum_{1\leq|I|\leq 2} X^I$ *for some collection of such vector fields* X^I.

Proof. The proof of the first three parts boils down to more or less elementary calculations. The last part follows from a direct calculation involving the conditions (3) and (39) is also left to the reader. □

4E. *A pointwise bound for multicommutators.* To conclude this section, we record a combined consequence of Lemma 4.11, Corollary 4.12, and Lemma 4.13. This will be our main tool for controlling commutators in the sequel.

Lemma 4.15 (pointwise bound for commutators). *For pair of multiindices* I, J *with* $|I|\geq 1$ *one has*

$$|\widetilde{\Gamma}^J[\Box_g,\Gamma^I]\phi| \lesssim \sum_{\substack{I'\subsetneq I\\ |J'|\leq|J|}}\Big(\sum_{\substack{l+|K|=1\\ |I''|\leq 1}} q\cdot\tau_x^{-1}|(\tau_0\partial_u^b)^l(\partial_x^b)^K\widetilde{\Gamma}^{J'}\Gamma^{I'+I''}\phi| + |\widetilde{\Gamma}^{J'}\Gamma^{I'}\Box_g\phi|\Big),\quad \text{where } q\in\mathcal{Z}^{-\frac{1}{2}}, \tag{85}$$

where Γ^I (*etc.*) *denotes a product of vector fields in* $\mathbb{L}=\{S,\Omega_{ij},\partial_u^b,\partial_i^b-\omega^i\partial_u^b\}$, *and* $\widetilde{\Gamma}^J$ (*etc.*) *denotes a product of vector fields in* $\mathbb{L}_0=\{\partial_u^b,\partial_i^b-\omega^i\partial_u^b\}$.

Proof. Using (81) followed by (76) and then (83) we have

$$\text{LHS (85)} \lesssim \text{RHS (85)} + \sum_{I'\subsetneq I,\ J'\subseteq J}|\Box_g(\widetilde{\Gamma}^{J'}\Gamma^{I'}\phi)|.$$

Thus, modulo induction on $|I|$ we have reduced matters to estimating the commutator $|[\Box_g,\widetilde{\Gamma}^{J'}](\Gamma^{I'}\phi)|$. Applying part (II) of Corollary 4.12, again followed by (76) and then (83), and inducting on $|I|$ we have

$$|[\Box_g,\widetilde{\Gamma}^{J'}](\Gamma^{I'}\phi)| \lesssim \text{RHS (85)} + \sum_{J''\subsetneq J'}|\Box_g(\widetilde{\Gamma}^{J''}\Gamma^{I'}\phi)|,$$

so the proof concludes with an additional round of induction, this time with respect to $|J|$. □

5. Proof of the weighted L^2 estimates for $k=0$

Theorem 5.1 (generalized local energy decay estimates). *Assume estimates* (7a) *and* (7b) *for* $s=0$; *then the following are true*:

(I) *For* R *sufficiently large there exists* $C_R>0$ *and vector fields* X_j *such that one has the uniform bound*

$$\|\phi\|_{WLE^0[0,T]} \lesssim \sup_{0\leq t\leq T}\|\partial\phi(t)\|_{L_x^2} + C_R\|\phi\|_{WLE^0_{\text{class}}[0,T]} + \sup_j\Big|\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}}\Box_g\phi\cdot X_j\phi\,dV_g\Big|^{\frac{1}{2}}, \tag{86}$$

where $X_j=\chi_{>R}q_j\partial_u^b$, *with* $q_j=q_j(u)$ *obeying the uniform bounds* $|(\tau_-\partial_u^b)^k q_j|\lesssim_k 1$. *In addition* $\chi_{>R}=\chi_{>1}(R^{-1}\cdot)$ *is a radial bump function with* $\chi_{>R}\equiv 1$ *on* $r>R$ *and* $\chi_{>R}\equiv 0$ *on* $r<\frac{1}{2}R$ *for some* R *sufficiently large so that* $\mathcal{K}\subseteq\{r<\frac{1}{2}R\}$.

(II) *For each* $0<a<1$ *there exists an* R_a *sufficiently large so that*

$$\sup_{0\leq t\leq T}\|\phi(t)\|_{E^a} + \|\phi\|_{S^a[0,T]} \lesssim_a \|\tau_x^a\partial\phi(0)\|_{L_x^2} + \|\tau_+^{a-1}\phi\|_{H_1^1(r<R_a)[0,T]} + \|\Box_g\phi\|_{N^a[0,T]}, \tag{87}$$

where the implicit constant depends continuously on $a\in(0,1)$.

(III) *For R sufficiently large there exists $C_R > 0$ and vector fields X_j such that one has uniformly*

$$\sup_{0\le t\le T} \|\phi(t)\|_{E^1} + \|\phi\|_{S^{1,\infty}[0,T]} \lesssim \|\tau_x \partial\phi(0)\|_{L^2} + C_R\|\phi\|_{\ell^1_t H^1_1(r<R)[0,T]} + \|\Box_g\phi\|_{\ell^\infty_t N^1[0,T]} + \sup_j \left| \int_0^T \int_{\mathbb{R}^3\setminus\mathcal{K}} \Box_g\phi \cdot \tau_x^{-1} X_j(\tau_x\phi)\, dV_g \right|^{\frac{1}{2}}, \quad (88)$$

where $X_j = \chi_{>R_1} q_j K_0$, with K_0 given by the formula $K_0 = (1+u^2)\partial_u^b + 2(u+r)r\partial_r^b$, where $q_j = q_j(u)$ has the uniform bounds $|(\tau_-\partial_u^b)^k q_j| \lesssim_k 1$, and where $\chi_{>R}$ is the same as in (86) above.

5A. ***Splitting into interior and exterior estimates.*** To control the solution in the interior we use:

Proposition 5.2 (weighted LE bounds in time-like regions). *Let $R_0 \ge 1$ be as in Definition 1.7. Then for any $a \ge 0$ and $1 \le p \le \infty$ one has the uniform bound*

$$\sup_{0\le t\le T} \|\tau_+^a(\partial\phi, \tau_x^{-1}\phi)(t)\|_{L^2_x(r<\frac{1}{2})} + \|\tau_+^a(\partial\phi, \tau_x^{-1}\phi)\|_{\ell^p_t \mathrm{LE}(r<\frac{1}{2})[0,T]}$$
$$\lesssim \|\tau_x^a \tau_+^{-\frac{1}{2}}(\partial\phi, \tau_+^{-1}\phi)\|_{\ell^p_t L^2(r<\frac{3}{4}t)[0,T]} + \|\tau_+^{a-1}\phi\|_{\ell^p_t H^1_1(r<R_0)[0,T]} + \|\tau_+^a \tau_0 \Box_g\phi\|_{\ell^p_t \mathrm{WLE}^{*,0}[0,T]}. \quad (89)$$

Here the ℓ^p_t sum is taken over a collection of dyadic regions $\langle t\rangle \approx 2^j \ge 1$.

In the exterior we use multipliers to show that:

Proposition 5.3 (weighted exterior LE bounds). *One has the following estimates uniformly for R sufficiently large that $\mathcal{K} \subseteq \{r \le \frac{1}{2}R\}$:*

(I) *The following null energy bounds hold*:

$$\|(\partial_x^b\phi, \tau_x^{-1}\phi)\|_{\mathrm{NLE}(r>R)[0,T]} \lesssim R^{\frac{1}{2}}\|\phi\|_{\mathrm{WLE}^0_{\mathrm{class}}[0,T]} + \sqrt{\mathcal{E}(0,R)}, \quad (90)$$

where $\mathcal{E}(0,R) = \sup_j \mathcal{E}(0,R,X_j)$ is given by formula (50) with X_j as in Theorem 5.1.

(II) *For fixed $0 < a < 1$ there holds*

$$\sup_{0\le t\le T} \|\phi(t)\|_{E^a(r>\max\{R,\frac{1}{2}t\})} + \|\phi\|_{S^a(r>\max\{R,\frac{1}{2}t\})[0,T]} + \|\tau_x^a\tau_+^{-\frac{1}{2}}(\partial\phi, \tau_x^{-1}\phi)\|_{L^2(R<r<\frac{3}{4}t)[0,T]} \lesssim_a \sqrt{\mathcal{E}(a,R)}, \quad (91)$$

where $\mathcal{E}(a,R)$ is given by formula (49).

(III) *Corresponding to $a = 1$ there holds the estimate*

$$\sup_{0\le t\le T} \|\phi(t)\|_{E^1(r>\max\{R,\frac{1}{2}t\})} + \|\phi\|_{S^{1,\infty}(r>\max\{R,\frac{1}{2}t\})[0,T]} + \|\tau_x\tau_+^{-\frac{1}{2}}(\partial\phi, \tau_+^{-1}\phi)\|_{\ell^\infty_t L^2(R<r<\frac{3}{4}t)[0,T]}$$
$$\lesssim \sup_{0\le t\le T} R^{\frac{1}{2}}\|\tau_-^{\frac{1}{2}}(\partial\phi, \tau_x^{-1}\phi)(t)\|_{L^2_x(\frac{1}{2}R<r<R)} + \sqrt{\mathcal{E}(1,R)}, \quad (92)$$

where $\mathcal{E}(1,R) = \sup_j \mathcal{E}(1,R,X_j)$ is given in terms of formula (51) with X_j as in Theorem 5.1.

Proof that Propositions 5.2 and 5.3 imply Theorem 5.1. We do this separately for each estimate.

Case 1: ($a=0$) Here we need to show (86) follows directly from (90) and the assumed bounds (7). From inspection of $\sqrt{\mathcal{E}(0,R)}$ and taking R sufficiently large, we only need to bound $\|(\partial_x^b\phi, \tau_x^{-1}\phi)\|_{\mathrm{NLE}(R_0<r<R)[0,T]}$ in terms of C_R times $\|\phi\|_{\mathrm{WLE}^0_{\mathrm{class}}[0,T]}$. This follows by taking $C_R \approx R^{\frac{1}{2}}$.

Case 2: $(0 < a < 1)$ Adding together a suitable linear combination of estimates (89) with $p = 2$ and (91), and using the inclusions $\ell_t^2\mathrm{LE} \subseteq \mathrm{LE}$ and $\mathrm{LE}^* \subseteq \ell_t^2\mathrm{LE}^*$ (from Minkowski's inequality), we have uniformly

$$\sup_{0\le t\le T} \|\phi(t)\|_{E^a} + \|\phi\|_{S^a[0,T]} \lesssim_a \sup_{0\le t\le 2R} \|\phi(t)\|_{E^a(r<R)} + \|\phi\|_{S^a(r<R)[0,2R]} + \|\tau_x^a\tau_+^{-\frac12}(\partial\phi,\, \tau_x^{-1}\phi)\|_{L^2(r<\min\{R,\frac34 t\})[0,T]} + \|\tau_+^{a-1}\phi\|_{H_1^1(r<R_0)[0,T]} + \|\Box_g\phi\|_{N^a[0,T]} + \sqrt{\mathcal{E}(a,\,R)},$$

Next, uniformly for $T_0 \ge 2R \ge 1$ there hold the pair of bounds

$$\|\tau_x^a\tau_+^{-\frac12}(\partial\phi,\, \tau_x^{-1}\phi)\|_{L^2(r<\min\{R,\frac34 t\})[0,T]} \lesssim \ln(R)\|\phi\|_{S^a(r<R)[0,T_0]} + (R/T_0)^{\frac12}\|\phi\|_{S^a[T_0,T]},$$

and

$$\mathcal{E}(a,\,R) \lesssim \|\tau_x^a\partial\phi(0)\|_{L_x^2}^2 + (o_R(1) + (R/T_0)^a)\cdot\Big(\sup_{0\le t\le T}\|\phi(t)\|_{E^a}^2 + \|\phi\|_{S^a[0,T]}^2\Big) + \|\phi\|_{S^a[0,T]}\cdot(\|\phi\|_{S^a(r<R)[0,T_0]} + R^{\frac32}\|\tau_+^{a-1}\phi\|_{H_1^0(r<R)[0,T]} + \|\Box_g\phi\|_{N^a[0,T]}).$$

In addition for $T_0 \ge 2R$ there is the simple bound

$$\sup_{0\le t\le 2R}\|\phi(t)\|_{E^a(r<R)} + \ln(R)\|\phi\|_{S^a(r<R)[0,T_0]} \lesssim_{R,T_0} \|\tau_+^{a-1}\phi\|_{H_1^1(r<R)[0,T_0]}.$$

Therefore combining the last four inequalities with $R = R_a \ge R_0$ sufficiently large depending on a, and for some $T_0 = T_a \ge R_a$ which depends on both the size of R_a and a, we have estimate (87).

Case 3: $(a = 1)$ Here we apply (89) with $a = 1$ and $p = \infty$, and add to this a suitable linear combination of (92). Note that an application of the weighted trace estimate (188c) followed by (184) with $a = 1$ gives

$$\sup_{0\le t\le T} R^{\frac12}\|\tau_-^{\frac12}(\partial\phi,\, \tau_x^{-1}\phi)(t)\|_{L_x^2(\frac12 R<r<R)} \lesssim \|\tau_x\partial\phi(0)\|_{L^2} + R^{\frac32}\|\phi\|_{\ell_t^1 H_1^1(r<R)[0,T]}.$$

The rest of the proof follows a similar pattern to Case 2 above. □

5B. *Proof of the interior estimate.* Before moving on to the proof of the exterior estimates, which are more involved, we first demonstrate Proposition 5.2.

Proof of estimate (89). Without loss of generality we work with the time interval $[1, T]$. We apply estimate (8) to $2^{ak}\chi_0(2^{-k}t)\chi_{<1}(r/t)\phi$ for $k \ge 0$, where $\chi_0(s)$ is a smooth bump function adapted to $1 \le s \le 2$, and $\chi_{<1}(s)$ is a smooth function $= 1$ for $s \le \frac12$ and $= 0$ for $s > \frac34$. Using the Hardy estimate (184) this yields

$$\sup_{2^k\le t\le 2^{k+1}}\|\tau_+^a(\partial\phi,\, \tau_x^{-1}\phi)(t)\|_{L_x^2(r<\frac12)} + \|\tau_+^a(\partial\phi,\, \tau_x^{-1}\phi)\|_{\mathrm{LE}(r<\frac12)[2^k,2^{k+1}]} \lesssim \|\tau_+^a(\partial_u^b\phi,\, \tau_+^{-1}\phi)\|_{H^1(r<R_0)[2^{k-1},2^{k+2}]} + \|\tau_+^{a-1}(\partial\phi,\, \tau_x^{-1}\phi)\|_{\mathrm{WLE}^{*,0}(r<\frac34)[2^{k-1},2^{k+2}]} + \|\tau_+^a\tau_0\Box_g\phi\|_{\mathrm{WLE}^{*,0}[2^{k-1},2^{k+2}]}.$$

For a fixed value of k, dyadic summation in r gives the uniform estimate

$$\|\tau_+^{a-1}(\partial\phi,\tau_x^{-1}\phi)\|_{\mathrm{WLE}^{*,0}(r<\frac34)[2^{k-1},2^{k+2}]} \lesssim \|\tau_+^{a-\frac12}(\partial\phi,\tau_x^{-1}\phi)\|_{L^2(r<\frac34)[2^{k-1},2^{k+2}]} + \|\tau_+^{a-1}\phi\|_{H^2(r<R_0)[2^{k-1},2^{k+2}]}.$$

By splitting into regions $r < \gamma t$ and $r > \gamma t$ we have the following uniform estimate for $0 < \gamma < \frac{1}{2}$:

$$\|\tau_+^{a-\frac{1}{2}}(\partial\phi,\tau_x^{-1}\phi)\|_{L^2(r<\frac{3}{4})[2^{k-1},2^{k+2}]} \lesssim \gamma^{\frac{1}{2}}\|\tau_+^a(\partial\phi,\tau_x^{-1}\phi)\|_{\mathrm{LE}(r<\frac{1}{2})[2^{k-1},2^{k+2}]} + \gamma^{-a-1}\|\tau_x^a\tau_+^{-\frac{1}{2}}(\partial\phi,\tau_+^{-1}\phi)\|_{L^2(r<\frac{3}{4})[2^{k-1},2^{k+2}]}.$$

On the one hand with the help of the scaling vector field we have

$$\|\tau_+^a(\partial_u^b\phi,\tau_+^{-1}\phi)\|_{H^1(r<R_0)[2^{k-1},2^{k+2}]} \lesssim \|\tau_+^{a-1}\phi\|_{H_1^1(r<R_0)[2^{k-1},2^{k+2}]}.$$

Thus (89) follows by summing the last four displays in ℓ_t^p and taking $\gamma \ll 1$. □

5C. ***Proof of the exterior estimates.*** At the level of multipliers there are essentially four cases here: $a = 0$, $0 < a \le \frac{1}{2}$, $\frac{1}{2} \le a < 1$, and $a = 1$. Collectively these are stated in the following lemma.

Lemma 5.4 (core multiplier bounds). *One has the following collection of estimates uniformly for R large enough that $\mathcal{K} \subseteq \{r \le \frac{1}{2}R\}$:*

(I) *Corresponding to the case $a = 0$ one has*

$$\|\tau_-^{-\frac{1}{2}}\tau_x^{-1}\partial_x^b(\tau_x\phi)\|^2_{\ell_u^\infty L^2(r>R)[0,T]} \lesssim \mathcal{E}(0,R). \tag{93}$$

(II) *In the range $0 < a \le \frac{1}{2}$ one has*

$$\begin{aligned}\sup_{0\le t\le T}\|\tau_+^a(\tau_0^{\frac{1}{2}}\partial_u^b\phi,\tau_x^{-1}\partial_x^b(\tau_x\phi))(t)\|^2_{L_x^2(r>R)} &+ \|\tau_x^{a-\frac{3}{2}}\partial_x^b(\tau_x\phi)\|^2_{L^2(r>R)[0,T]} + \|\tau_+^{a-\frac{1}{2}}\tau_0^2\partial_u^b\phi\|^2_{L^2(r>R)[0,T]} \\ &+ \|\tau_-^{-\frac{1}{2}}\tau_x^{a-1}\partial_r^b(\tau_x\phi)\|^2_{\ell_u^\infty L^2(r>R)[0,T]} + \|\tau_+^{a-\frac{1}{2}}\tau_0^{\frac{1}{2}}\partial_u^b\phi\|^2_{\ell_{\underline{u}}^\infty L^2(r>R)[0,T]} \lesssim a^{-1}\mathcal{E}(a,R).\end{aligned} \tag{94}$$

Here we have set $\underline{u} = u + 2r$.

(III) *When $\frac{1}{2} \le a < 1$ one has*

$$\begin{aligned}\sup_{0\le t\le T}\|\tau_+^a(\tau_0^a\partial_u^b\phi,\tau_x^{-1}\partial_x^b(\tau_x\phi))(t)\|^2_{L_x^2(r>R)} &+ \|\tau_x^{a-\frac{3}{2}}\partial_x^b(\tau_x\phi)\|^2_{L^2(r>R)[0,T]} + \|\tau_+^{a-\frac{1}{2}}\tau_0^2\partial_u^b\phi\|^2_{L^2(r>R)[0,T]} \\ &+ \|\tau_-^{-\frac{1}{2}}\tau_x^{a-1}\partial_r^b(\tau_x\phi)\|^2_{\ell_u^\infty L^2(r>R)[0,T]} + \|\tau_+^{-\frac{1}{2}}\tau_-^a\partial_u^b\phi\|^2_{\ell_{\underline{u}}^\infty L^2(r>R)[0,T]} \lesssim (1-a)^{-1}\mathcal{E}(a,R).\end{aligned} \tag{95}$$

Here again we have set $\underline{u} = u + 2r$.

(IV) *Finally, corresponding to the case $a = 1$ we have*

$$\sup_{0\le t\le T}\|\tau_+(\tau_0\partial_u^b\phi,\tau_x^{-1}\partial_x^b(\tau_x\phi))(t)\|^2_{L_x^2(r>R)} + \|\tau_-^{-\frac{1}{2}}\tau_+\tau_x^{-1}\partial_r^b(\tau_x\phi)\|^2_{\ell_u^\infty L^2(r>R)[0,T]} \lesssim \mathcal{E}(1,R). \tag{96}$$

We'll prove each of these estimates separately. In each case we invoke Proposition 4.6.

Proof of estimate (93). Fix a $j \in \mathbb{N}$ and define the multiplier

$$Y^u = \chi_{<j}(u), \qquad Y^r = 0,$$

where $\partial_s \chi_{<j}(s) = -2^{-j}\chi_j(s)$, with $\chi_j(s) \approx 1$ when $\langle s\rangle \approx 2^j$ and vanishing away from this region. It can also be arranged that $\chi_{<j}(s) = 0$ for $s \gtrsim 2^j$ by addition of a suitable constant. It is immediate these coefficients satisfy (55), so we may use case (III) of Proposition 4.6. A short calculation then shows

$$\mathcal{A}^u = \mathcal{A}^{ur} = 0, \quad \mathcal{A}^r = \not{\mathcal{A}} = 2^{-j-1}\chi_j(u)$$

for the coefficients in (54). Repeatedly applying estimate (53) for different $j \in \mathbb{N}$, then taking $\sup_j$ of the result concludes the proof. □

Proof of estimate (94). First freeze $j, k \in \mathbb{N}$ and define the multiplier

$$Y^u = a(2^{(2a-1)k}\tau_-\chi_{<k}(\underline{u}) + \tau_+^{2a}\tau_0^4), \quad Y^r = Cr^{2a}(1+\chi_{<j}(u)), \quad \underline{u} = u + 2r, \quad \tau_+ = 1 + \underline{u},$$

where $\chi_{<k}$, $\chi_{<j}$ is as above and $C > 0$ will be chosen shortly. It is easy to check that these coefficients satisfy the conditions on (52) (in $r > R$). Computing the formulas from (54) and choosing $C \gg 1$ we have the following uniform estimates for $0 < a \le \frac{1}{2}$ in the region $r > 1$:

$$\begin{aligned} \mathcal{A}^u &\gtrsim a(\tau_+^{2a-1}\tau_0\chi_k(\underline{u}) + \tau_+^{2a-1}\tau_0^4), \quad & \mathcal{A}^r &\gtrsim \tau_-^{-1}\tau_x^{2a}\chi_j(u) + a\tau_x^{2a-1}, \\ |\mathcal{A}^{ur}| &\ll \sqrt{\mathcal{A}^u\mathcal{A}^r}, & \not{\mathcal{A}} &\gtrsim \tau_x^{2a-1}. \end{aligned}$$

Repeatedly applying estimate (53) for different $j, k \in \mathbb{N}$ and on time intervals $[0, t] \subseteq [0, T]$ and then taking $\sup_{j,k,t}$ concludes the proof. □

Proof of estimate (95). In this case we set

$$Y^u = (1-a)(\tau_-^{2a}\chi_{<k}(\underline{u}) + \tau_+^{2a}\tau_0^4), \quad Y^r = C^2r^{2a}(1+\chi_{<j}(u)) + C(1-a)\tau_+^{2a-1}r.$$

These satisfy the conditions in (52), and the formulas from (54) yield when $C \gg 1$ the following estimates uniformly in $\frac{1}{2} \le a < 1$:

$$\begin{aligned} \mathcal{A}^u &\gtrsim (1-a)(\tau_+^{-1}\tau_-^{2a}\chi_k(\underline{u}) + \tau_+^{2a-1}\tau_0^4), \quad & \mathcal{A}^r &\gtrsim \tau_-^{-1}\tau_x^{2a}\chi_j(u) + (1-a)\tau_x^{2a-1}, \\ |\mathcal{A}^{ur}| &\ll \sqrt{\mathcal{A}^u\mathcal{A}^r}, & \not{\mathcal{A}} &\gtrsim (1-a)\tau_x^{2a-1}. \end{aligned}$$

These suffice to give (95) through an application of (53). □

Proof of estimate (96). For this estimate we use the multiplier

$$Y^u = (1+\chi_{<j}(u))\tau_-^2, \quad Y^r = (1+\chi_{<j}(u))2(u+r)r.$$

This satisfies the conditions in (52) with $a = 1$. The formulas from (54) yield

$$\begin{aligned} \mathcal{A}^u &= 0, \quad & \mathcal{A}^r &\gtrsim \tau_-^{-1}\tau_+^2\chi_j(u), \\ \mathcal{A}^{ur} &= 0, & \not{\mathcal{A}} &\ge 0. \end{aligned}$$

This suffices to give (96) through an application of case (II) of Proposition 4.6. □

It remains to close the gap between Proposition 5.3 and Lemma 5.4.

Proof of (90). Applying the Hardy estimate (187) with $a = \frac{1}{2}$ to the function $2^{-\frac{1}{2}j}\chi_j(u)\phi$, followed by the Hardy estimate (184) with $a = 0$ for the term on the right-hand side of (187) at $t = 0$, one has

$$\|\tau_-^{-\frac{1}{2}}\tau_x^{-1}\phi\|_{\ell_u^\infty L^2(r>R)[0,T]} \lesssim \|\tau_-^{-\frac{1}{2}}\tau_x^{-1}\partial_x^b(\tau_x\phi)\|_{\ell_u^\infty L^2(r>R)[0,T]} + R^{\frac{1}{2}}\|\phi\|_{\mathrm{WLE}^0_{\mathrm{class}}[0,T]} + \|\partial\phi(0)\|_{L_x^2}, \quad (97)$$

where the extra factor of $R^{\frac{1}{2}}$ results from replacing $\tau_-^{-\frac{1}{2}}$ with $\tau_x^{-\frac{1}{2}}$ in the spacetime region $\frac{1}{2}R < r < R$. Combining this with estimate (93) and using the definition (50) we have (90). □

Proof of (91). Estimate (187) of Appendix C and the definition of $\mathcal{E}(a, R)$ from (49) imply

$$\|\tau_x^{a-1}\phi(T)\|^2_{L_x^2(r>R)} + \|\tau_x^{a-\frac{3}{2}}\phi\|^2_{L^2(r>R)[0,T]} \lesssim_a \|\tau_x^{a-\frac{3}{2}}\partial_r^b(\tau_x\phi)\|^2_{L^2(r>R)[0,T]} + \mathcal{E}(a, R).$$

Adding this to (94) and (95) we have (91). □

Proof of (92). Adding estimate (186) to (96) gives

$$\sup_{0\le t\le T}\|\phi(t)\|_{E^1(r>\max\{R,\frac{1}{2}t\})} + \|\tau_x\tau_+^{-\frac{1}{2}}(\partial\phi, \tau_+^{-1}\phi)\|_{\ell_t^\infty L^2(R<r<\frac{3}{4}t)[0,T]} + \|\tau_+\tau_x^{-1}\partial_r^b(\tau_x\phi)\|_{\mathrm{NLE}(r>R)[0,T]}$$
$$\lesssim \sup_{0\le t\le T} R^{\frac{1}{2}}\|\tau_-^{\frac{1}{2}}(\partial\phi, \tau_x^{-1}\phi)(t)\|_{L_x^2(\frac{1}{2}R<r<R)} + \sqrt{\mathcal{E}(1, R)}.$$

On the other hand a straightforward integration of fixed-time norms in the region $r > \frac{1}{2}t$ also gives

$$\|\phi\|_{\ell_t^\infty S^1(r>\max\{R,\frac{1}{2}t\})[0,T]} \lesssim \sup_{0\le t\le T}\|\phi(t)\|_{E^1(r>\max\{R,\frac{1}{2}t\})}.$$

Combining the previous two equations with the definition of $S^{1,\infty}$ finishes the proof. □

6. Estimates for commutators

In this section we complete the proofs of estimates (14), (15), and (16).

6A. ***Splitting into interior and exterior estimates.*** In the interior we use:

Proposition 6.1 (general interior estimates for commutators). *Fix a multiindex $|I| = k \geqslant 1$ and let Γ^I denote a product in $\mathbb{L} = \{S, \Omega_{ij}, \partial_u^b, \partial_i^b - \omega^i\partial_u^b\}$. Then for R sufficiently large and any $R' > R$ one has the following estimates*:

$$\|[\Box_g, \Gamma^I]\phi\|_{\mathrm{WLE}^{*,s}(r<R)[0,T]} \lesssim_R \|\phi\|_{H^{s+3}_{k-1}(r<R)[0,T]}, \quad (98)$$

$$\|[\Box_g, \Gamma^I]\phi\|_{N^a(r<R)[0,T]} \lesssim_{R,R'} \|\tau_+^{a-1}\phi\|_{H^1_{k+1}(r<R')[0,T]} + \|\phi\|_{S^a_{k-1}[0,T]} + \|\Box_g\phi\|_{N^a_k[0,T]}, \quad (99)$$

$$\|[\Box_g, \Gamma^I]\phi\|_{\ell_t^\infty N^1(r<R)[0,T]} \lesssim_{R,R'} \|\phi\|_{\ell_t^\infty H^1_{k+1}(r<R')[0,T]} + \|\phi\|_{S^{1,\infty}_{k-1}[0,T]} + \|\Box_g\phi\|_{\ell_t^\infty N^1_k[0,T]}. \quad (100)$$

The bound (99) *is uniform in* $0 \le a \le 1$.

For the exterior we use:

Proposition 6.2 (general exterior bounds for commutators). *Fix multiindices $|I| = k \geqslant 1$ and $|J| = s \geqslant 0$, and let Γ^I denote a product of vector fields in $\mathbb{L} = \{S, \Omega_{ij}, \partial_u^b, \partial_i^b - \omega^i\partial_u^b\}$, while $\widetilde{\Gamma}^J$ denotes a product of vector fields in $\mathbb{L}_0 = \{\partial_u^b, \partial_i^b - \omega^i\partial_u^b\}$. Let R_0 be as in the definition of the norms (6). Then for $R \geq 2R_0 > 0$ sufficiently large one has the following*:

(I) *Corresponding to $a = 0$ one has*

$$\|[\Box_g, \Gamma^I]\phi\|_{(\mathrm{WLE}^{*,s}+L_t^1H_x^s)(r>R)[0,T]} \lesssim o_R(1)\cdot\|\phi\|_{\mathrm{WLE}_k^s(r>R)[0,T]} + \|\Box_g\phi\|_{(\mathrm{WLE}_{k-1}^{*,s}+L_t^1H_{x,k-1}^s)(r>R)[0,T]}. \quad (101)$$

In addition let $X = \chi_{>R}q\partial_u^b$, where $q = q(u)$ has the uniform bounds $|(\tau_-\partial_u^b)^l q| \lesssim_l 1$, where $\chi_{>R}$ is supported in $r > \frac{1}{2}R$ with the usual derivative bounds. Then one has the integral estimate

$$\left|\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}}[\Box_g, \widetilde{\Gamma}^J]\phi\cdot X\widetilde{\Gamma}^J\phi\, dV_g\right| \lesssim o_R(1)\cdot\big(\|\phi\|^2_{\mathrm{WLE}^s[0,T]} + \sup_{0\leq t\leq T}\|\partial\phi(t)\|^2_{H_x^s}\big) + \big(\sup_{0\leq t\leq T}\|\partial\phi(t)\|_{H_x^s} + \|\phi\|_{\mathrm{WLE}^s[0,T]}\big)\cdot\|\Box_g\phi\|_{(\mathrm{LE}^{*,s}+L_t^1H_x^s)[0,T]}. \quad (102)$$

(II) *For $0 < a < 1$ one has uniformly the collection of estimates*

$$\|[\Box_g, \Gamma^I]\phi\|_{N^a(r>R)[0,T]} \lesssim o_R(1)\cdot\|\phi\|_{S_k^a[0,T]} + \|\Box_g\phi\|_{N_{k-1}^a[0,T]}. \quad (103)$$

(III) *Finally, corresponding to the case $a = 1$ we have*

$$\|[\Box_g, \Gamma^I]\phi\|_{\ell_t^\infty N^1(r>R)[0,T]} \lesssim o_R(1)\cdot\|\phi\|_{S_k^{1,\infty}[0,T]} + \|\Box_g\phi\|_{\ell_t^\infty N_{k-1}^1[0,T]}. \quad (104)$$

In addition let $X = \chi_{>R}qK_0$, where $K_0 = (1+u^2)\partial_u^b + 2(u+r)r\partial_r^b$, and where q and $\chi_{>R}$ are as previously stated. Then one has

$$\left|\int_0^T\int_{\mathbb{R}^3\setminus\mathcal{K}}[\Box_g, \Gamma^I]\phi\cdot\tau_x^{-1}X(\tau_x\Gamma^I\phi)\, dV_g\right| \lesssim o_R(1)\cdot\big(\|\phi\|^2_{S_k^{1,\infty}[0,T]} + \sup_{0\leq t\leq T}\|\phi(t)\|^2_{E_k^1}\big) + \|\phi\|_{S_k^{1,\infty}[0,T]}\cdot\|\Box_g\phi\|_{N_{k-1}^{1,1}[0,T]}. \quad (105)$$

We will also need the following initial data bound.

Lemma 6.3 (initial data bound). *Assume that the level sets $t =$ const are uniformly spacelike. Then one has the uniform estimate for $0 \leqslant a \leqslant 1$*

$$\|\tau_x^a\partial\phi(0)\|_{H_{x,k}^s} \lesssim \sum_{|I|\leq k}\sum_{|J|\leq s}\|\tau_x^a(\tau_x\partial_x)^I\partial_x^J\partial\phi(0)\|_{L_x^2} + \|\tau_x^a\Box_g\phi\|_{L_t^1H_{x,k}^s[0,1]}. \quad (106)$$

Before giving the proofs of these individual components we use them to establish Theorem 1.16.

Proof of estimate (14). We need to treat separately the cases $k = 0$ and $k > 0$.

Case 1: ($k = 0$) In light of assumption (7a) and the data bound (106) we need to show

$$\sum_{|J|\leq s}\|(\partial^b)^J\phi\|_{\mathrm{WLE}^0[0,T]} \lesssim \|\partial\phi(0)\|_{H_x^s} + \|\Box_g\phi\|_{(\mathrm{WLE}^{*,s}+L_t^1H_x^s)[0,T]}.$$

It suffices to show this bound for $(\partial^b)^J\phi$ replaced by $\widetilde{\Gamma}^J\phi$, where the product is taken over vector fields in $\mathbb{L}_0 = \{\partial_u^b, \partial_i^b - \omega^i\partial_u^b\}$. Applying estimate (86) to $\widetilde{\Gamma}^J\phi$, then using assumption (7a) and (102), we find that for $|J| \leq s$

$$\|(\partial^b)^J\phi\|^2_{\mathrm{WLE}^0[0,T]} \lesssim C_R(\|\partial\phi(0)\|^2_{H^s_x} + \|\Box_g\phi\|^2_{(\mathrm{WLE}^{*,s}+L^1_t H^s_x)[0,T]}) + o_R(1)\cdot\|\phi\|^2_{\mathrm{WLE}^s[0,T]}.$$

Summing this in $|J| \leq s$ and taking R sufficiently large finishes the proof.

Case 2: $(k > 0)$ This is a straightforward application of estimate (14) with $k = 0$ applied to $\Gamma^I\phi$, where Γ^I denotes a product in $\mathbb{L} = \{S, \Omega_{ij}, \partial_u^b, \partial_i^b - \omega^i\partial_u^b\}$. Using estimates (98) and (101) for sufficiently large R to control the commutator, followed by the data bound (106), completes the proof. Note that the value R becomes the definition of $B_0(s,k)$ on (14). □

Proof of estimate (15). This follows by combining estimate (87) with (99), (103), and (106), and then using an induction on k. Note that the value R becomes the definition of $B_a(s,k)$ in (15). □

Proof of estimate (16). This follows from estimates (88), (100), (104), (105), and (106) in a similar pattern as the previous proof. □

6B. ***Proof of the initial data bound and the interior estimates.***

Proof of (106). By Remark 1.4 we can assume $u \equiv t - \tau_x$ when $t \in [0,1]$. Therefore without loss of generality we may replace Γ^I with products of vector fields in $\mathbb{L}_{\mathrm{Minkowski}} = \{t\partial_t + r\partial_r, x^i\partial_j - x^j\partial_i, \partial_\alpha\}$. Notice that for products in $\mathbb{L}_{\mathrm{Minkowski}}$ we have the following equivalence at $t = 0$:

$$\|\tau_x^a\phi^{(k)}(0)\|_{H^s_x} \approx \sum_{|I|+j\leq k} \|\tau_x^a(\tau_x\partial_x)^I\partial_t^j\phi(0)\|_{H^s_x} \tag{107}$$

for a similarity constant that depends only on k. In addition to this, from the uniform spacelike condition of $t = 0$ we have the identity

$$\partial_t^2 = (g^{00})^{-1}(\Box_g - P(t,x;D)), \tag{108}$$

where $P(t,x;D)$ is second-order with uniformly $(r\partial)^J$ homogeneous coefficients, is at most first-order in ∂_t, and contains no zero-order term.

First we repeatedly use the $\lesssim$ direction of (107) for $\phi(0)$, followed by the substitution (108) for terms containing more than one copy of ∂_t. We then use the $\gtrsim$ direction of (107) for any terms which are produced which contain $\Box_g\phi(0)$. From this sequence of steps we have the bound

$$\|\tau_x^a\partial\phi(0)\|_{H^s_{x,k}} \lesssim \sum_{|I|\leq k} \|\tau_x^a(\tau_x\partial_x)^I\partial\phi(0)\|_{H^s_x} + \|\tau_x^a(\Box_g\phi)^{(k-1)}(0)\|_{H^s_x}.$$

After a local $L^1(dt)$ trace estimate, the second term on the right-hand side above matches the right-hand side of (106).

For the first term on the right-hand side on the previous inequality we again use identity (108) to successively get rid of all additional ∂_t-derivatives from the H^s_x norm, followed by another application of the $\gtrsim$ direction of (107) for terms produced containing $\Box_g\phi(0)$. This gives

$$\sum_{|I|\le k}\|\tau_x^a(\tau_x\partial_x)^I\partial\phi(0)\|_{H_x^s}\lesssim\sum_{|I|\le k}\sum_{|J|\le s}\|\tau_x^a(\tau_x\partial_x)^I\partial_x^J\partial\phi(0)\|_{L_x^2}+\|\tau_x^a(\Box_g\phi)^{(k)}(0)\|_{H_x^{s-1}}.$$

The proof now concludes with a final local $L^1(dt)$ trace estimate the second term on the right-hand side above. ☐

We now move on to the proof of Proposition 6.1. Note that estimate (98) follows more or less immediately from (79). Therefore we focus attention on the last two estimates.

Proof of estimates (99) *and* (100). We will only focus here in showing (99). The proof of (100) follows from identical calculations by replacing all L_t^2 norms with $\ell_t^\infty L_t^2$ norms.

Step 1: (*inductive setup*) It suffices to show that for multiindex $|I|=k$ and R sufficiently large, for $R'>R$ we have

$$\begin{aligned}\|\tau_+^a[\Box_g,\Gamma^I]\phi\|_{H^1(r<R)[0,T]}\lesssim_{R,R'}&\|\tau_+^{a-1}\phi^{(k+1)}\|_{H^1(r<R')[0,T]}+\|\tau_+^a\phi^{(k-1)}\|_{L^2(r<R')[0,T]}\\&+\|\tau_+^a(\Box_g\phi)^{(k)}\|_{H^1(r<R_0)[0,T]}+\|\tau_+^a(\Box_g\phi)^{(k)}\|_{L^2(r<R')[0,T]}.\end{aligned}\tag{109}$$

This boils down to an induction. Indeed, for fixed nontrivial I and integer $s\ge1$ it suffices to show that under the same assumptions one has

$$\begin{aligned}\|\tau_+^a[\Box_g,\Gamma^I]\phi\|_{H^s(r<R)[0,T]}\lesssim_{R,R'}&\|\tau_+^{a-1}\phi^{(|I|+s)}\|_{H^1(r<R')[0,T]}+\|\tau_+^a\phi^{(|I|-1)}\|_{L^2(r<R')[0,T]}\\&+\|\tau_+^a(\Box_g\phi)^{(|I|-1)}\|_{H^{s+1}(r<R_0)}+\|\tau_+^a(\Box_g\phi)^{(|I|-1)}\|_{H^s(r<R')}\\&+\sum_{I'\subsetneq I}\|\tau_+^a[\Box_g,\Gamma^{I'}]\phi\|_{H^{s+1}(r<R')[0,T]}.\end{aligned}\tag{110}$$

By repeatedly applying this last estimate for $|I|+s=k+1$, where $s=1,\dots,k$, and a sequence $R<R'_s<R'_{s+1}<R'$ we have (109).

Step 2: (*elliptic estimate in* $\frac12R_0<r<R$) To prove (110) start with (79), which implies that

$$\|\tau_+^a[\Box_g,\Gamma^I]\phi\|_{H^s(r<R)[0,T]}\lesssim\sum_{I'\subsetneq I}\|\tau_+^a\Gamma^{I'}\phi\|_{H^{s+2}(r<R)[0,T]}.\tag{111}$$

Without loss of generality we may assume R_0 is chosen large enough that in the region $r>\frac12R_0$ the operator $P(x,D)=\Box_g-Q_0(x,D)$ is uniformly elliptic, where Q_0 contains all terms with a $g^{0\alpha}$-factor. Standard elliptic estimates then give

$$\|\tau_+^a\Gamma^{I'}\phi\|_{H^{s+2}(\frac12R_0<r<R)[0,T]}\lesssim_{R,R'}\|\tau_+^aP(x,D)\Gamma^{I'}\phi\|_{H^s(r<R')[0,T]}+\|\tau_+^a\Gamma^{I'}\phi\|_{L^2(r<R')[0,T]}.$$

On the other hand due to the fact that any term in $Q_0(x,D)$ contains at least one time derivative, and using the metric conditions (3), we have for $I'\subsetneq I$

$$\|\tau_+^aQ_0(x,D)\Gamma^{I'}\phi\|_{H^s(r<R')[0,T]}\lesssim_{R'}\|\tau_+^{a-1}\phi^{(|I|+s)}\|_{H^1(r<R')[0,T]}.$$

Thus, combining the last two inequalities we have

$$\sum_{I'\subsetneq I}\|\tau_+^a\Gamma^{I'}\phi\|_{H^{s+2}(\frac12 R_0<r<R)[0,T]}$$
$$\lesssim_{R,R'}\|\tau_+^{a-1}\phi^{(|I|+s)}\|_{H^1(r<R')[0,T]}+\|\tau_+^a\phi^{(|I|-1)}\|_{L^2(r<R')[0,T]}$$
$$+\|\tau_+^a(\Box_g\phi)^{(|I|-1)}\|_{H^s(r<R')}+\sum_{I'\subsetneq I}\|\tau_+^a[\Box_g,\Gamma^{I'}]\phi\|_{H^s(r<R')[0,T]}. \quad (112)$$

Step 3: (LE *estimate in* $r<\frac12 R_0$) It remains to bound the portion of the right-hand side of (111) which is contained in $r<\frac12 R_0$. Note that we only need to focus on the region $t>1$ as the right-hand side of (111) restricted to the time slab [0, 1] is automatically bounded by $\|\tau_+^{a-1}\phi^{(|I|+s)}\|_{H^1(r<R')[0,T]}$.

We begin by applying the stationary LE bound (7b) at regularity $s+1$ to $\chi_{t>1}\chi_{r<\frac12 R_0}\tau_+^a\Gamma^{I'}\phi$, where $\chi_{r<\frac12 R_0}=1$ on $r<\frac12 R_0$ and $\chi_{r<\frac12 R_0}=0$ on $r>R_0$, and with similar properties for $\chi_{t>1}$. This results in the estimate

$$\sum_{I'\subsetneq I}\|\tau_+^a\Gamma^{I'}\phi\|_{H^{s+2}(r<\frac12 R_0)[1,T]}\lesssim_{R,R'}T_1+T_2+T_3+T_4,$$

where

$$T_1=\sum_{I'\subsetneq I}\|\Gamma^{I'}\phi\|_{H^{s+2}(r<R_0)[0,2]},\qquad T_2=\sum_{I'\subsetneq I}\|(\partial_t\tau_+^a\Gamma^{I'}\phi,\ \tau_+^a\Gamma^{I'}\phi)\|_{H^{s+1}\times L^2(r<R_0)[0,T]},$$
$$T_3=\sum_{I'\subsetneq I}\|[\Box_g,\ \chi_{r<\frac12 R_0}\tau_+^a]\Gamma^{I'}\phi\|_{H^{s+1}[0,T]},\qquad T_4=\sum_{I'\subsetneq I}\|\tau_+^a\Box_g\Gamma^{I'}\phi\|_{H^{s+1}(r<R_0)[0,T]}.$$

The term T_1 results from differentiation of $\chi_{t>1}$. The terms T_1 and T_4 are already compatible with the right-hand side of (110). It is also easy to see that

$$T_2\lesssim\|\tau_+^{a-1}\phi^{(|I|+s)}\|_{H^1(r<R_0)[0,T]}+\|\tau_+^a\phi^{(|I|-1)}\|_{L^2(r<R_0)[0,T]}.$$

It remains to bound T_3. Expanding the commutator gives

$$T_3\lesssim\|\tau_+^{a-1}\phi^{(|I|+s)}\|_{H^1(r<R_0)[0,T]}+\sum_{I'\subsetneq I}\|\tau_+^a\Gamma^{I'}\phi\|_{H^{s+2}(\frac12 R_0<r<R_0)[0,T]}.$$

The first term above is of the correct form. The second term is handled by estimate (112). □

6C. ***Proof of the exterior estimates.*** We first list some general calculations which take care of a large portion of the desired estimates.

Lemma 6.4. *For integers* $s,k\geq 0$ *define*

$$\Phi^{(s,k)}=\sum_{|J|\leq s}\tau_x^{-\frac12}|(\tau_0\partial(\partial^b)^J\phi^{(k)},\ \partial_x^b(\partial^b)^J\phi^{(k)},\ \tau_x^{-1}(\partial^b)^J\phi^{(k)})|,\qquad \not{\Phi}^{(k)}=\tau_x^{-\frac12}\tau_0^{-\frac12},|\tau_x^{-1}\partial_r^b(\tau_x\phi^{(k)})|.$$

We also use the shorthand $\Phi^{(0,k)}=\Phi^{(k)}$. *With this notation we have*

$$\|\Phi^{(s,k)}\|_{\ell_r^\infty L^2(r>R)[0,T]}+\|\tau_0^{-\frac12}\Phi^{(s,k)}\|_{\ell_u^\infty\ell_r^\infty L^2(\frac12 t<r<2t)\cap(r>R)[0,T]}\lesssim\|\phi\|_{\mathrm{LE}_k^s(r>R)[0,T]}, \quad (113)$$
$$\|\tau_+^a\Phi^{(k)}\|_{\ell_r^\infty L^2[0,T]}+\|\tau_+^a\not{\Phi}^{(k)}\|_{\ell_u^\infty L^2(\frac12 t<r<2t)[0,T]}\lesssim\|\phi\|_{S_k^a[0,T]}, \quad (114)$$

$$\|\tau_+\Phi^{(k)}\|_{\ell_t^\infty\ell_r^\infty L^2[0,T]} + \|\tau_+\not{\Phi}^{(k)}\|_{\ell_u^\infty L^2(\frac{1}{2}t<r<2t)[0,T]} \lesssim \|\phi\|_{S_k^{1,\infty}[0,T]}, \tag{115}$$

$$\|\tau_x^{\frac{1}{2}}\tau_+\Phi^{(k)}\|_{L_t^\infty L_x^2[0,T]} \lesssim \sup_{0\le t\le T}\|\phi(t)\|_{E_k^1}. \tag{116}$$

In addition if Γ^I *is a product of vector fields in* $\mathbb{L} = \{S, \Omega_{ij}, \partial_u^b, \partial_i^b - \omega^i\partial_u^b\}$, *and* $\widetilde{\Gamma}^J$ *a product of vector fields in* $\mathbb{L}_0 = \{\partial_u^b, \partial_i^b - \omega^i\partial_u^b\}$, *where* $|I| = k$ *and* $|J| = s$, *then we have the pointwise estimates*

$$|\widetilde{\Gamma}^J[\Box_g, \Gamma^I]\phi| \lesssim q\cdot\tau_0^{-\frac{1}{2}}\tau_x^{-\frac{1}{2}}\Phi^{(s,k)} + \sum_{|J|\le s}|\partial^J(\Box_g\phi)^{(k-1)}|, \quad \text{where } q\in\mathcal{Z}^0. \tag{117}$$

Proof of Lemma 6.4. The proof of (113)–(116) is a straightforward application of the definition of the various spaces involved. On the other hand (117) is immediate from (85). □

Note that a direct combination of (117) with either (113), (114), or (115) shows (101), (103), or (104) respectively. Thus, the remainder of the subsection is devoted to showing the integral estimates (102) and (105). In both cases the key step is to integrate by parts the bilinear operator resulting from the commutator with $\Box_g$. The relevant result here is:

Lemma 6.5. *Let* R *be sufficiently large so that* $\mathcal{K} \subseteq \{|x| < \frac{1}{2}R\}$, *and let* $\chi_{>R}$ *be a cutoff supported in* $r > \frac{1}{2}R$, *constant for* $r > R$, *with the usual derivative bounds. Then one has the following integral estimates*:

(I) *Let* $q = q(u)$ *be a smooth function such that* $|(\tau_-\partial_u^b)^l q| \lesssim_l 1$. *Furthermore let* $|J| = s \ge 1$ *and* $|J'| \le s-1$ *and let* $\widetilde{\Gamma}^J$ *and* $\widetilde{\Gamma}^{J'}$ *denote products of vector fields in* $\mathbb{L}_0 = \{\partial_u^b, \partial_i^b - \omega^i\partial_u^b\}$. *Then if* $\mathcal{R}$ *obeys the conditions (43) with* $a = c = -1$ *and* $b = 1$ *one has*

$$\Big|\int_0^T\!\!\int_{\mathbb{R}^3}\chi_{>R}\nabla_\alpha\mathcal{R}^{\alpha\beta}\nabla_\beta\widetilde{\Gamma}^{J'}\phi\cdot q\partial_u^b\widetilde{\Gamma}^J\phi\, dV_g\Big| \lesssim o_R(1)\cdot\big(\|\phi\|^2_{\mathrm{LE}^s(r>\frac{1}{2}R)[0,T]} + \sup_{0\le t\le T}\|\partial\phi(t)\|^2_{H_x^s}\big). \tag{118}$$

(II) *Alternatively, suppose that* $q = q(u)$ *satisfies* $|(\tau_-\partial_u^b)^l q| \lesssim_l \tau_-$. *Let* Γ^I *and* $\Gamma^{I'}$ *be products of vector fields in* $\mathbb{L} = \{S, \Omega_{ij}, \partial_u^b, \partial_i^b - \omega^i\partial_u^b\}$ *for multiindices* $|I| = k \ge 1$ *and* $|I'| \le k-1$. *Then if* $\mathcal{R}$ *obeys the conditions (40) one has*

$$\Big|\int_0^T\!\!\int_{\mathbb{R}^3}\chi_{>R}\nabla_\alpha\mathcal{R}^{\alpha\beta}\nabla_\beta\Gamma^{I'}\phi\cdot qS\Gamma^I\phi\, dV_g\Big| \lesssim o_R(1)\cdot\big(\|\phi\|^2_{S_k^{1,\infty}[0,T]} + \sup_{0\le t\le T}\|\phi(t)\|^2_{E_k^1}\big). \tag{119}$$

(III) *Finally, with the same setup of estimate (119) let* $\mathcal{S}$ *obey the conditions (46). Then one has*

$$\Big|\int_0^T\!\!\int_{\mathbb{R}^3}\chi_{>R}\mathcal{S}^\alpha\partial_\alpha\Gamma^{I'}\phi\cdot qS\Gamma^I\phi\, dV_g\Big| \lesssim o_R(1)\cdot\big(\|\phi\|^2_{S_k^{1,\infty}[0,T]} + \sup_{0\le t\le T}\|\phi(t)\|^2_{E_k^1}\big). \tag{120}$$

Proof of estimate (118). For functions F and G, a vector field X, and quadratic operator $\nabla_\alpha\mathcal{R}^{\alpha\beta}\nabla_\beta$, we have the pointwise identity

$$\begin{aligned}\nabla_\alpha\mathcal{R}^{\alpha\beta}\nabla_\beta F\cdot XG &= \nabla_\alpha[\mathcal{R}^{\alpha\beta}\nabla_\beta F\cdot XG - X^\alpha\mathcal{R}^{\beta\gamma}\partial_\beta F\cdot\partial_\gamma G]\\ &\qquad + (\nabla_\gamma X^\gamma)\mathcal{R}^{\alpha\beta}\partial_\alpha F\cdot\partial_\beta G + \mathcal{R}_X^{\alpha\beta}\partial_\alpha F\cdot\partial_\beta G + \mathcal{R}^{\alpha\beta}\partial_\alpha XF\cdot\partial_\beta G,\\ &= \nabla_\alpha T_1^\alpha + T_2 + T_3 + T_4,\end{aligned} \tag{121}$$

where $\mathcal{R}_X = \mathcal{L}_X\mathcal{R}$. Setting $F = \widetilde{\Gamma}^{J'}\phi$, $G = \widetilde{\Gamma}^J\phi$, and $X = q\partial_u^b$ in the above formula we estimate the integral of each term separately.

<u>Case 1</u>: (*the T_1-term*) Using the divergence theorem gives

$$\Big|\int_0^T\int_{\mathbb{R}^3} \chi_{>R}\nabla_\alpha T_1^\alpha\, dV_g\Big| \lesssim \sup_{0\leq t\leq T}\int_{|x|>\frac{1}{2}R}|\nabla_\alpha t|\cdot|T_1^\alpha|\,dx + R^{-1}\Big|\int_0^T\int_{\mathbb{R}^3}\chi_R' T_1^r\, dV_g\Big|, \tag{122}$$

where $|\chi_R'|\lesssim 1$ and is supported where $\frac{1}{2}R < r < R$. Based on the fact that all components of X are uniformly bounded, and all components of $\mathcal{R}$ are $o_R(1)$ (in either Bondi or (t,x)-coordinates), we directly have the pointwise estimate

$$|\nabla_\alpha t|\cdot|T_1^\alpha| \lesssim \sup_\alpha |T_1^\alpha| \lesssim o_R(1)\sum_{|J''|\leq s}|\partial\partial^{J''}\phi|^2,$$

which in turn produces a bound for the right-hand side of (122) in terms of

$$o_R(1)\cdot(\|\phi\|^2_{LE^s(r>\frac{1}{2}R)[0,T]} + \sup_{0\leq t\leq T}\|\partial\phi(t)\|^2_{H_x^s}).$$

<u>Case 2</u>: (*the T_2-term*) Notice that estimate (39) gives $|\nabla_\gamma X^\gamma|\lesssim 1$ for $X = q\partial_u^b$. On the other hand for $\mathcal{R}$ satisfying conditions (43) with $a = c = -1$ and $b = 1$ we have the pointwise estimate

$$|\mathcal{R}^{\alpha\beta}\partial_\alpha\widetilde{\Gamma}^{J'}\phi\partial_\beta\widetilde{\Gamma}^J\phi| \lesssim p\cdot\sum_{|J''|\leq s}\tau_x^{-1}(|\partial\partial^{J''}\phi|^2 + \tau_0^{-1}|\partial_x^b\partial^{J''}\phi|^2), \quad \text{where } p\in\mathcal{Z}^0. \tag{123}$$

This suffices to give

$$\Big|\int_0^T\int_{\mathbb{R}^3}\chi_{>R}T_2\, dV_g\Big| \lesssim o_R(1)\cdot\|\phi\|^2_{LE^s(r>\frac{1}{2}R)[0,T]}.$$

<u>Case 3</u>: (*the T_3-term*) This is similar to the previous step. Notice that $X = q\partial_u^b$ obeys the symbol bounds (41) with $a = b = 0$ and $c = -1$, and satisfies all conditions in (42). Therefore, thanks to case (I) of Lemma 4.3 we have that $\mathcal{R}_X$ satisfies the bounds in (43) with $a = c = -1$ and $b = 1$. This is enough to show (123) holds for $\mathcal{R}$ replaced by $\mathcal{R}_X$.

<u>Case 4</u>: (*the T_4-term*) Modulo another bound similar to (123) it suffices to show

$$|(\partial_\alpha q)\mathcal{R}^{\alpha\beta}\partial_u^b\widetilde{\Gamma}^{J'}\phi\partial_\beta\widetilde{\Gamma}^J\phi| \lesssim \text{RHS}\,(123).$$

This follows from direct inspection of various terms involved. □

Proof of estimate (119). We again use the identity (121) and estimate each term separately. As a preliminary note that with the assumptions of (119) and notation of Lemma 6.4 one has the pointwise estimate

$$\tau_-|\mathcal{R}^{\alpha\beta}\partial_\alpha\Gamma^{I'}\phi\cdot\partial_\beta\Gamma^I\phi| \lesssim p\cdot\tau_x\tau_+|\Phi^{(k)}|^2, \quad \text{where } p\in\mathcal{Z}^0. \tag{124}$$

A similar bound holds if we replace $\Gamma^{I'}\phi$ by $S\Gamma^{I'}\phi$.

Likewise, when $X = qS$ with $q = q(u)$ and bounds $|(\tau_-\partial_u^b)^l q| \lesssim_l \tau_-$ we have condition (41) with $a = b = 0$ and $c = 1$, and also the conditions in (42) save for the second identity. Therefore by (III) of Lemma 4.3 the tensor $\mathcal{R}_X = \mathcal{L}_X\mathcal{R}$ satisfies (in Bondi coordinates)

$$\mathcal{R}_X^{ij} \in \tau_+ \cdot \mathcal{Z}^0, \quad \mathcal{R}_X^{ui} \in \tau_+ \cdot \mathcal{Z}^1, \quad \mathcal{R}_X^{uu} \in \tau_+ \cdot \mathcal{Z}^2.$$

In particular we have the pointwise estimate

$$|\mathcal{R}_X^{\alpha\beta}\partial_\beta\Gamma^{I'}\phi \cdot \partial_\beta\Gamma^I\phi| \lesssim p \cdot \tau_x\tau_+|\Phi^{(k)}|^2, \quad \text{where } p \in \mathcal{Z}^0. \tag{125}$$

<u>Case 1</u>: (*the T_1-term*) Here we again use (122). For the first term on the right-hand side of (122) estimate (116) shows it suffices to prove

$$\sup_\alpha |T_1^\alpha| \lesssim o_R(1) \cdot \tau_x\tau_+^2|\Phi^{(k)}|^2. \tag{126}$$

For the first term in T_1^α we use

$$\sup_\alpha \tau_-|\mathcal{R}^{\alpha\beta}\partial_\beta\Gamma^{I'}\phi \cdot S\Gamma^I\phi| \lesssim o_R(1)\tau_x\tau_+^2|\Phi^{(k)}|^2,$$

which follows from $|\mathcal{R}^{\alpha\beta}| \lesssim 1$ and expanding S into ∂_u^b- and ∂_x^b-derivatives. For the second term in T_1^α we have (126) thanks to (124) and $|X^\alpha| \lesssim \tau_+\tau_-$.

For the second term on the right-hand side of (122) the pointwise bound (126) is not sufficient to recover the ℓ_t^∞ structure needed on the right-hand side of (119). However, similar calculations to those above show $T_1^r = T_{11}^r + T_{12}^r$, where

$$R^{-1}\chi_R' T_{11}^r = \chi_R'\mathcal{S}^\alpha\partial_\alpha\Gamma^{I'}\phi \cdot qS\Gamma^I\phi, \quad R^{-1}|\chi_R' T_{12}^r| \lesssim o_R(1) \cdot \chi_R(r)\tau_x\tau_+|\Phi^{(k)}|^2.$$

Here χ_R is a cutoff on a dyadic region $\approx R$, and $\mathcal{S}^\alpha = R^{-1}\chi_R(r)\mathcal{R}^{r\alpha}$ is a smooth vector field satisfying the assumptions of estimate (120) (which will be proved independently). Therefore we only need bound the second term in the display above. Using (116) in the region $t \le R$, and (115) in the region $t > R$, we have

$$R^{-1}\int_0^T\int_{\mathbb{R}^3} |\chi_R' T_{12}^r|\,dx\,dt \lesssim o_R(1) \cdot \big(\|\phi\|^2_{S_k^{1,\infty}[0,T]} + \sup_{0\le t\le T}\|\phi(t)\|^2_{E_k^1}\big).$$

<u>Case 2</u>: (*the T_2-term*) For the remaining three terms in (121) we will set things up so as to appeal to Lemma 4.10. For $i = 2, 3, 4$ we will show

$$\int_0^T\int_{\mathbb{R}^3} \chi_{>R}|T_i|\,dV_g \lesssim \|\sqrt{\tau_x/\tau_+}p \cdot \tau_+\Phi^{(k)}\|^2_{L^2[0,T](r>R)}, \quad \text{where } p^2 \in \mathcal{Z}^0, \tag{127}$$

which by a combination of estimate (71) and estimates (115) and (116) produces

$$\int_0^T\int_{\mathbb{R}^3} \chi_{>R}|T_i|\,dV_g \lesssim o_R(1) \cdot \big(\|\phi\|^2_{S_k^{1,\infty}[0,T]} + \sup_{0\le t\le T}\|\phi(t)\|^2_{E_k^1}\big).$$

For the specific case of the T_2-term note that (39) shows $|\nabla_\gamma X^\gamma| \lesssim \tau_-$. Then (124) immediately gives (127).

<u>Case 3</u>: (*the T_3-term*) In this case (127) follows at once from (125).

Case 4: (*the T_4-term*) Modulo an application of (124) with $\Gamma^{I'}\phi$ replaced by $S\Gamma^{I'}\phi$, to produce (127) for this case it suffices to prove the pointwise estimate

$$|(\partial_\alpha q)\mathcal{R}^{\alpha\beta}S\Gamma^{I'}\phi\cdot\partial_\beta\Gamma^I\phi| \lesssim p\cdot\tau_x\tau_+|\Phi^{(k)}|^2, \quad \text{where } p\in\mathcal{Z}^0,$$

which follows from expanding S to get

$$|\mathcal{R}^{u\beta}S\Gamma^{I'}\phi\cdot\partial_\beta\Gamma^I\phi| \lesssim p\cdot\tau_+(\tau_0^2|\partial\phi^{(k)}|^2+|\partial_x^b\phi^{(k)}|^2), \quad \text{where } p\in\mathcal{Z}^0.$$

This completes the proof of (119). □

Proof of (120). For functions F and G, and vector fields $\mathcal{S}$ and X, we have the pointwise identity

$$\mathcal{S}^\alpha\partial_\alpha F\cdot XG = \nabla_\alpha(X^\alpha\mathcal{S}^\beta\partial_\beta F\cdot G) - (\nabla_\alpha X^\alpha)\mathcal{S}^\beta\partial_\beta F\cdot G - \mathcal{S}_X^\beta\partial_\beta F\cdot G - \mathcal{S}^\beta\partial_\beta XF\cdot G. \tag{128}$$

Here $\mathcal{S}_X = \mathcal{L}_X\mathcal{S} = [X,\mathcal{S}]$. We again need to estimate each term separately.

As a general first step note if $\mathcal{S}$ satisfies (46), $|I'|\le k-1$, and $|I|\le k$ then

$$\tau_-|\mathcal{S}^\alpha\partial_\alpha\Gamma^{I'}\phi\cdot\Gamma^I\phi| \lesssim p\cdot\tau_x\tau_+|\Phi^{(k)}|^2, \quad \text{where } p\in\mathcal{Z}^0. \tag{129}$$

A similar bound holds if we replace $\Gamma^{I'}\phi$ by $S\Gamma^{I'}\phi$.

Likewise, when $X = qS$, where $q = q(u)$ with bounds $|(\tau_-\partial_u^b)^l q| \lesssim_l \tau_-$, we have from part (IV) of Lemma 4.3 the estimates

$$\mathcal{S}_X^i \in \tau_x^{-1}\tau_+\cdot\mathcal{Z}^{\frac{1}{2}}, \quad \mathcal{S}_X^u \in \tau_x^{-1}\tau_+\cdot\mathcal{Z}^{\frac{3}{2}}.$$

This gives the pointwise estimate

$$|\mathcal{S}_X^\alpha\partial_\alpha\Gamma^{I'}\phi\cdot\Gamma^I\phi| \lesssim p\cdot\tau_x\tau_+|\Phi^{(k)}|^2, \quad \text{where } p\in\mathcal{Z}^0. \tag{130}$$

Finally,

$$|(\partial_\alpha q)\mathcal{S}^\alpha S\Gamma^{I'}\phi\cdot\Gamma^I\phi| \lesssim p\cdot\tau_x\tau_+|\Phi^{(k)}|^2, \quad \text{where } p\in\mathcal{Z}^0, \tag{131}$$

which follows from $(\partial_\alpha q)\mathcal{S}^\alpha \in \tau_x^{-2}\tau_+\cdot\mathcal{Z}^{\frac{1}{2}}$ and $\tau_x^{-2}\tau_+|S\Gamma^{I'}\phi\cdot\Gamma^I\phi| \lesssim \tau_x^{-2}\tau_+|\phi^{(k)}|^2$. With (128)–(131) in hand, the remainder of the proof of (120) is essentially identical to the proof of (119) above. □

Proof of estimate (102). Using part (II) of Corollary 4.12 we may write

$$\text{LHS (102)} \lesssim \left|\int_0^T\int_{\mathbb{R}^3}\chi_{>R}T_1\,dV_g\right| + \int_0^T\int_{\mathbb{R}^3}\chi_{>R}|T_2|\,dx\,dt$$
$$+\Big(\sup_{0\le t\le T}\|\partial\phi(t)\|_{H_x^s}+\|\phi\|_{LE^s(r>R)[0,T]}\Big)\cdot\|\Box_g\phi\|_{(LE^{*,s}+L_t^1H_x^s)(r>R)[0,T]},$$

where

$$T_1 = \sum_{J'\subsetneq J}\nabla_\alpha\mathcal{R}_{J'}^{\alpha\beta}\nabla_\beta\widetilde{\Gamma}^{J'}\phi\cdot q\partial_u^b\widetilde{\Gamma}^J\phi, \quad T_2 = \sum_{J'\subsetneq J}\mathcal{S}_{J'}^\alpha\partial_\alpha\widetilde{\Gamma}^{J'}\phi\cdot q\partial_u^b\widetilde{\Gamma}^J\phi,$$

and where $\mathcal{R}_{J'}$ and $\mathcal{S}_{J'}$ satisfy estimates (43) and (47) respectively with $a=c=-1$ and $b=1$. We are assuming the weight $q=q(u)$ satisfies $|(\tau_-\partial_u^b)^l q| \lesssim_l 1$. The term T_1 is therefore handled by estimate

(118). On the other hand the conditions on $\mathcal{S}_{I'}$ and inspection give the pointwise bound

$$|T_2| \lesssim p \cdot \sum_{|J''|\leq s} \tau_x^{-1}(|\partial\partial^{J''}\phi|^2 + \tau_0^{-1}|\partial_x^b(\partial^b)^{J''}\phi|^2), \quad \text{where } p \in \mathcal{Z}^0.$$

This suffices to produce $\int_0^T\int_{\mathbb{R}^3} \chi_{>R}|T_2| \lesssim o_R(1)\cdot\|\phi\|^2_{\mathrm{LE}^s(r>R)[0,T]}$. □

Proof of estimate (105). Note that the definition of X and the notation of Lemma 6.4 give the pointwise bound

$$|\tau_x^{-1}X(\tau_x\Gamma^I\phi)| \lesssim \tau_x^{\frac{1}{2}}\tau_+^2\tau_0^{\frac{1}{2}}\cdot(\Phi^{(k)} + \not{\Phi}^{(k)}).$$

Next, we take the decomposition $\tau_x^{-1}X\tau_x = uS + X'$, where $X' = q(\partial_u^b + (u+2r)r\tau_x^{-1}\partial_r^b\tau_x + ur^2\tau_x^{-2})$. This leads to the following improvement of the previous inequality

$$|X'(\Gamma^I\phi)| \lesssim \tau_x^{\frac{3}{2}}\tau_+\tau_0^{\frac{1}{2}}\cdot(\Phi^{(k)} + \not{\Phi}^{(k)}).$$

Therefore combining (79), (117), the previous two inequalities, and (115) gives

$$\text{LHS (102)} \lesssim \sum_{i=1,2}\left|\int_0^T\int_{\mathbb{R}^3}\chi_{>R}T_i\,dV_g\right| + \|\sqrt{\tau_x/\tau_+}p\cdot\tau_+(\Phi^{(k)}+\not{\Phi}^{(k)})\|^2_{L^2[0,T](r>R)} + \|\phi\|_{S_k^{1,\infty}[0,T]}\cdot\|\Box_g\phi\|_{N_{k-1}^{1,1}[0,T]},$$

where $p^2 \in \mathcal{Z}^0$, where

$$T_1 = \sum_{I'\subsetneq I}\nabla_\alpha\mathcal{R}_{I'}^{\alpha\beta}\nabla_\beta\Gamma^{I'}\phi\cdot\tilde{q}S\Gamma^I\phi, \quad T_2 = \sum_{I'\subsetneq I}\mathcal{S}_{I'}^\alpha\partial_\alpha\Gamma^{I'}\phi\cdot\tilde{q}S\Gamma^I\phi, \tag{132}$$

and where $\tilde{q} = uq(u)$ satisfies the assumptions of Lemma 6.5. The proof of (105) is concluded by an application of estimate (119) to handle the contribution of T_1, estimate (120) to handle the contribution of T_2, and Lemma 4.10 followed by (115) and (116), which together show

$$\|\sqrt{\tau_x/\tau_+}p\cdot\tau_+(\Phi^{(k)}+\not{\Phi}^{(k)})\|^2_{L^2[0,T](r>R)} \lesssim o_R(1)\cdot\big(\|\phi\|^2_{S_k^{1,\infty}[0,T]} + \sup_{0\leq t\leq T}\|\phi(t)\|^2_{E_k^1}\big).$$ □

7. L^∞ estimates

The purpose of this section is to prove Theorem 1.18. In fact we will prove the slightly stronger bound:

Proposition 7.1 (fixed-time global Sobolev inequality). *Let $R_1 \geq 1$ be large enough so that $\mathcal{K} \subseteq \{r < R_1\}$. Then there exists $R \geq R_1$ sufficiently large such that given any $k \geq 1$ one has the following fixed-time estimate uniform in $t \geq 0$:*

$$\sum_{i+|J|\leq k}\|\tau_+^{\frac{3}{2}}\tau_0^{\frac{1}{2}}(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\phi(t)\|_{L_x^\infty} \lesssim \|\tau_+^{\frac{3}{2}}\phi^{(k)}(t)\|_{H_x^2(r<R)} + \|\phi(t)\|_{E_{k+1}^1} + \sum_{i+|J|\leq k}\|\tau_+^{\frac{3}{2}}\tau_x^{\frac{1}{2}}\tau_0(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\Box_g\phi(t)\|_{\ell_r^1L_x^2(r<t)\cap L_x^2}. \tag{133}$$

In the estimate above the implicit constant depends on R.

We first give a quick demonstration of how Proposition 7.1 produces Theorem 1.18.

Proof that (133) implies (17). By an application of $T \leq 1$ and (188a) (when $T > 1$), and using the identity $\tau_- \partial_u^b = S + q\partial$ for some smooth q with $|\partial^J q| \lesssim \tau_x$, we have uniformly for $0 \leq t \leq T$ the bound

$$\|\tau_+^{\frac{3}{2}} \phi^{(k)}(t)\|_{H_x^2(r<R)} \lesssim_R \sup_{0\leq t\leq T} \|\phi(t)\|_{E_{k+1}^1} + \|\phi\|_{S_{k+2}^{1,\infty}[0,T]}.$$

Likewise, by an appropriate combination of (188a)–(188c) we have

$$\sup_{0\leq t\leq T} \sum_{i+|J|\leq k} \|\tau_+^{\frac{3}{2}} \tau_x^{\frac{1}{2}} \tau_0 (\tau_- \partial_u^b)^i (\tau_x \partial_x^b)^J \Box_g \phi(t)\|_{\ell_r^1 L_x^2(r<t)\cap L_x^2}$$
$$\lesssim \sum_{|J|\leq k} \|\tau_x^2 (\tau_x \partial)^J \Box_g \phi(0)\|_{L_x^2} + \sum_{i+|J|\leq k+1} \|(\tau_- \partial_u^b)^i (\tau_x \partial_x^b)^J \Box_g \phi\|_{N^{1,1}[0,T]}. \quad \Box$$

7A. *Reduction of Proposition 7.1.* The proof of (133) will rest on previous material and the following three lemmas. In each of these t is a fixed parameter and $\phi = \phi(t)$ only depends on x. We also assume $R > 0$ is chosen as in Proposition 7.1.

Lemma 7.2 (basic L^∞ estimates). *Let ϕ be a test function supported in $r < \frac{3}{4}\langle t\rangle$. Then one has*

$$\|(\tau_x \partial_x \phi, \phi^{(1)})\|_{L_x^\infty} \lesssim \|\tau_x^{\frac{1}{2}} (\partial_x^2 \phi^{(1)}, \tau_x^{-1} \partial_x \phi^{(1)}, \tau_x^{-2} \phi^{(1)})\|_{\ell_r^\infty L_x^2}. \tag{134}$$

On the other hand, without any support conditions imposed on ϕ we have

$$\|\tau_x^{\frac{3}{2}} \tau_0^{\frac{1}{2}} \phi\|_{L_x^\infty} \lesssim \sum_{l+|J|\leqslant 2} \|(\tau_x \tau_0 \partial_u^b)^l (\tau_x \partial_x^b)^J \phi\|_{L_x^2}. \tag{135}$$

Lemma 7.3 (global elliptic estimate). *Let ϕ be a function supported in $r < \frac{3}{4}\langle t\rangle$. Then for R sufficiently large one has the fixed-time estimate*

$$\|\tau_x^{\frac{1}{2}} (\partial_x^2 \phi, \tau_x^{-1} \partial_x \phi, \tau_x^{-2} \phi)\|_{\ell_r^\infty L_x^2} \lesssim \|\phi\|_{H_x^2(r<R)} + \|\tau_x^{\frac{1}{2}} (\partial \partial_u^b \phi, \tau_x^{-1} \partial_u^b \phi, \tau_+^{-1} \partial \phi)\|_{\ell_r^1 L_x^2} + \|\tau_x^{\frac{1}{2}} \Box_g \phi\|_{\ell_r^1 L_x^2}, \tag{136}$$

where the implicit constant depends on R.

Lemma 7.4 (fixed-time commutator estimate). *Let ϕ be a function which is supported in the region $r < \frac{3}{4}\langle t\rangle$, and let Γ^I denote a product of vector fields in $\mathbb{L} = \{S, \Omega_{ij}, \partial_u^b, \partial_i^b - \omega^i \partial_u^b\}$ with $|I| = k \geqslant 1$. Then for R as in Lemma 7.3 one has the fixed-time estimate*

$$\|\tau_x^{\frac{1}{2}} [\Box_g, \Gamma^I] \phi\|_{\ell_r^1 L_x^2} \lesssim \|\phi^{(k-1)}\|_{H_x^2(r<R)} + \langle t\rangle^{-\frac{3}{2}} \|\phi\|_{E_k^1} + \|\tau_x^{\frac{1}{2}} (\Box_g \phi)^{(k-1)}\|_{\ell_r^1 L_x^2}, \tag{137}$$

where the implicit constant again depends on R.

We postpone the proofs of these in order to first establish Proposition 7.1.

Proof of (133). We estimate the timelike and null/spacelike regions separately.

Step 1: (*proof of* (133) *in* $r > \frac{1}{2}\langle t\rangle$) Applying estimate (135) to $\chi_{r>\frac{1}{2}\langle t\rangle}(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\phi$ followed by (84) we have

$$\sum_{I+|J|\leqslant k}\|\tau_+^{\frac{3}{2}}\tau_0^{\frac{1}{2}}(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\phi\|_{L_x^\infty(r>\frac{1}{2}\langle t\rangle)}\lesssim\sum_{I+|J|\leqslant k+2}\|(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\phi\|_{L_x^2},$$
$$\lesssim\|\phi\|_{E^1_{k+1}}+\sum_{I+|J|\leqslant k}\|\tau_+^{\frac{3}{2}}\tau_x^{\frac{1}{2}}\tau_0(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\Box_g\phi\|_{L_x^2}.$$

Step 2: (*reduction of* (133) *in* $r < \frac{1}{2}\langle t\rangle$ *to truncated functions*) To prove (133) in the region $r < \frac{1}{2}\langle t\rangle$ we claim it suffices to show the fixed-time estimate

$$\sum_{I+|J|\leqslant k}\|(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\psi\|_{L_x^\infty}\lesssim\|\psi^{(k)}\|_{H_x^2(r<R)}+\langle t\rangle^{-\frac{3}{2}}\|\psi\|_{E^1_{k+1}}+\sum_{I+|J|\leqslant k}\|\tau_x^{\frac{1}{2}}(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\Box_g\psi\|_{\ell_r^1L_x^2},\qquad(138)$$

for functions ψ supported in the region $r < \frac{3}{4}\langle t\rangle$. Indeed, applying (138) to $\psi=\chi_{r<\frac{1}{2}\langle t\rangle}\phi$ and multiplying the result by $\langle t\rangle^{\frac{3}{2}}$ we have shown (133) in $r < \frac{1}{2}\langle t\rangle$ after using $\|\chi_{r<\frac{1}{2}t}\phi\|_{E^1_{k+1}}\lesssim\|\phi\|_{E^1_{k+1}}$ as well as

$$\sum_{I+|J|\leqslant k}\|\tau_+^{\frac{3}{2}}\tau_x^{\frac{1}{2}}(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J[\Box_g,\chi_{r<\frac{1}{2}\langle t\rangle}]\phi\|_{\ell_r^1L_x^2}\lesssim\sum_{I+|J|\leqslant k+1}\|(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\phi\|_{L_x^2},$$
$$\lesssim\|\phi\|_{E^1_k}+\sum_{I+|J|\leqslant k-1}\|\tau_+^{\frac{3}{2}}\tau_x^{\frac{1}{2}}\tau_0(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\Box_g\phi\|_{L_x^2}.$$

On the last line we have again used (84).

Step 3: (*reduction of* (138) *to the case* $k=1$) Using (84) we have

$$\text{LHS (138)}\lesssim\sum_{I+|J|\leqslant 1}\|(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\psi^{(k-1)}\|_{L_x^\infty}+\sum_{I+|J|\leqslant k-2}\|\tau_x^2(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\Box_g\psi\|_{L_x^\infty}.$$

On the other hand for functions ψ supported in $r < \frac{3}{4}\langle t\rangle$ estimate (135) gives

$$\sum_{I+|J|\leqslant k-2}\|\tau_x^2(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\Box_g\psi\|_{L_x^\infty}\lesssim\sum_{I+|J|\leqslant k}\|\tau_x^{\frac{1}{2}}(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\Box_g\psi\|_{L_x^2}.$$

Therefore, with the help of (137) we have reduced (138) to showing

$$\sum_{I+|J|\leqslant 1}\|(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\psi\|_{L_x^\infty}\lesssim\|\psi^{(1)}\|_{H_x^2(r<R)}+\langle t\rangle^{-\frac{3}{2}}\|\psi\|_{E^1_2}+\|\tau_x^{\frac{1}{2}}\Box_g\psi^{(1)}\|_{\ell_r^1L_x^2}\qquad(139)$$

for functions ψ supported in $r < \frac{3}{4}\langle t\rangle$.

Step 4: (*proof of* (139)) As a first step we have for ψ supported in $r < \frac{3}{4}\langle t\rangle$

$$\sum_{I+|J|\leqslant 1}\|(\tau_-\partial_u^b)^I(\tau_x\partial_x^b)^J\psi\|_{L_x^\infty}\lesssim\|(\tau_x\partial_x\psi,\psi^{(1)})\|_{L_x^\infty},$$

which follows from Remark 1.4 and by writing $S=(u-ru_r)\partial_u^b+r\partial_r$ and $\partial_i^b=\partial_i-u_i\partial_u^b$.

Next, concatenating (134) and (136) we have

$$\|(\tau_x\partial_x\psi,\psi^{(1)})\|_{L^\infty_x}\lesssim\|\psi^{(1)}\|_{H^2_x(r<R)}+\langle t\rangle^{-1}\|\tau_x^{\frac12}(\partial u\partial_u^b\psi^{(1)},\tau_x^{-1}u\partial_u^b\psi^{(1)},\partial\psi^{(1)})\|_{\ell^1_rL^2_x}+\|\tau_x^{\frac12}\Box_g\psi^{(1)}\|_{\ell^1_rL^2_x}.$$

Here we have used that $|\langle t\rangle^{-1}u|\approx 1$ on the support of ψ, as well as $|[\partial,u]|\lesssim 1$.

Finally, we use the expansion $u\partial_u^b=S+q\partial$, where $|\partial^Jq|\lesssim\tau_x^{1-|J|}$ on $r<\frac34\langle t\rangle$, as well as the estimate $\tau_x^{-\frac12}\times$ (83) for terms involving ∂^2, which altogether gives

$$\|\tau_x^{\frac12}(\partial u\partial_u^b\psi^{(1)},\tau_x^{-1}u\partial_u^b\psi^{(1)},\partial\psi^{(1)})\|_{\ell^1_rL^2_x}\lesssim\langle t\rangle^{-\frac12}\|\psi\|_{E^1_2}+\langle t\rangle\|\tau_x^{\frac12}\Box_g\psi^{(1)}\|_{L^2_x}. \tag{140}$$

This completes the proof of (139), and hence our demonstration of (133). □

7B. *Proof of the supporting lemmas.* We now prove Lemmas 7.2, 7.3, and 7.4.

Proof of estimate (134). First note that a rescaled version of the usual $H^2\to L^\infty$ Sobolev estimates gives

$$\|\phi\|_{L^\infty_x}\lesssim\|\tau_x^{\frac12}(\partial_x^2\phi,\tau_x^{-1}\partial_x\phi,\tau_x^{-2}\phi)\|_{\ell^\infty_rL^2_x},$$

which applied to $\phi^{(1)}$ proves half of (134).

It remains to prove (134) for $\tau_x\partial_x\phi$, and this is really only an issue where $r>1$. In this case the result follows from Remark 1.4 and applying the following global Sobolev estimate to $\partial_x\phi$ for $R\geq 1$:

$$\|\phi\|_{L^\infty_x(\frac12R<r<2R)}\lesssim\sum_{|J|\leq 1}R^{-\frac12}\|(\partial_x\Omega^J\phi,R^{-1}\Omega^J\phi)\|_{L^2_x(\frac14R<r<4R)},\quad\text{where }\Omega\in\{x^i\partial_j-x^j\partial_i\}.$$

Note that this bound is scale-invariant so it suffices to prove it for $R=1$. After using a set of angular cutoffs and a local chart on $\mathbb{S}^2$, it becomes the mixed Sobolev embedding

$$\|\phi\|_{L^\infty(\mathbb{R}^3)}\lesssim\sum_{|I|\leq1,|J|\leq1}\|\partial_x^I\partial_{x'}^J\phi\|_{L^2(\mathbb{R}^3)}, \tag{141}$$

where the coordinates are written as $x=(x^1,x^2,x^3)=(x^1,x')\in\mathbb{R}^3=\mathbb{R}\times\mathbb{R}^2$. Estimate (141) follows in the usual way by combining the Fourier inversion formula, the Cauchy–Schwarz inequality, and Plancherel's theorem, and using the fact that the multiplier $\langle\xi\rangle^{-1}\langle\xi'\rangle^{-1}$ is in $L^2(\mathbb{R}^3)$. □

Proof of estimate (135). It suffices to consider the region $\frac12\langle t\rangle<r<\frac32\langle t\rangle$ and $r\gg1$, as the complementary bound follows by rescaling the $H^2\to L^\infty$ Sobolev embedding. Using dyadic cutoffs we may further assume ϕ is supported where $\tau_-\approx 2^k$ and $\tau_x\approx 2^j$, and by using angular sector cutoffs in the x-variable we may further restrict this support to a $\frac{\pi}{4}$ wedge about the x^1-axis.

Next, we introduce the following variables on $t=\text{const}$, $r\gg1$, and the $\frac{\pi}{4}$ wedge about the x^1-axis:

$$y^1=2^{-k}u,\quad y^2=2^{-j}x^2,\quad y^3=2^{-j}x^3.$$

Changing variables we have the formulas on $t=\text{const}$

$$\partial_{y^1}=2^k\Big(\partial_u^b+\frac{1}{u_{x^1}}\partial_{x^1}^b\Big),\quad\partial_{y'}=2^j\Big(\partial_{x'}^b-\frac{u_{x'}}{u_{x^1}}\partial_{x^1}^b\Big),\quad 2^k2^{2j}\,dy=|u_{x^1}|\,dx,$$

where the derivatives u_{x^i} are also with respect to $t = \text{const}$, and $y' = (y^2, y^3)$. In particular by condition (1) there exist coefficients c^i_α which are uniformly bounded where $\tau_- \approx 2^k$, $\tau_x \approx 2^j$, and within the $\frac{\pi}{4}$ wedge about the x^1-axis, such that $\partial_{y^i} = \sum_\alpha c^i_\alpha e_\alpha$, where $e_\alpha \in \{\tau_- \partial^b_u, \tau_x \partial^b_x\}$.

By the change of measures formula in the previous display we have

$$2^{\frac{1}{2}k+j} \sum_{|I|\leqslant 2} \|e^I \phi\|_{L^2(dy)} \approx \sum_{a+|\beta|\leqslant 2} \|(\tau_x \tau_0 \partial^b_u)^l (\tau_x \partial^b_x)^J \phi\|_{L^2(dx)}.$$

To finish the proof it suffices to establish $H^2 \to L^\infty$ Sobolev estimates in the y-coordinates in terms of the e_α vector fields. Since the coefficients c^i_α are possibly very rough with respect to the y-variable we do this by concatenating H^1 Sobolev embeddings in the following way:

$$\|\phi\|_{L^\infty(dy)} \lesssim \sum_{|I|\le 1} \|\partial^I_y \phi\|_{L^6(dy)} \lesssim \sum_{|I|\le 1} \|e^I \phi\|_{L^6(dy)} \lesssim \sum_{|J|\le 1} \sum_{|I|\le 1} \|\partial^J_y e^I \phi\|_{L^2(dy)} \lesssim \sum_{|I|\le 2} \|e^I \phi\|_{L^2(dy)}.$$

This completes the proof of (135). □

Proof of estimate (136). Following Remark 1.4 we see that the metric in (t, x^i)-coordinates satisfies

$$\|(\langle t\rangle \partial_t)^l (\tau_x \partial_x)^J (g^{\alpha\beta} - \eta^{\alpha\beta})\|_{\ell^1_r L^\infty(r<\frac{3}{4}\langle t\rangle)} \lesssim_{l,J} 1,$$

where $\eta = \mathrm{diag}(-1, 1, 1, 1)$ is the standard Minkowski metric. This gives the pointwise estimate

$$|\Delta\phi| \lesssim q \cdot |(\partial^2_x \phi, \tau_x^{-1} \partial_x \phi)| + |\partial \partial^b_u \phi| + \tau_x^{-1} |\partial^b_u \phi| + \tau_+^{-1} |\partial \phi| + |\Box_g \phi|, \quad \text{where } q \in \ell^1_r L^\infty,$$

and where Δ is the standard 3-dimensional Laplacian. Therefore, by choosing R sufficiently large we see that to prove (136) it suffices to show

$$\|\tau_x^{\frac{1}{2}} (\partial^2_x \phi, \tau_x^{-1} \partial_x \phi, \tau_x^{-2} \phi)\|_{\ell^\infty_r L^2_x} \lesssim \|\phi\|_{H^2_x(r<R_1)} + \|\tau_x^{\frac{1}{2}} \Delta\phi\|_{\ell^1_r L^2_x}, \tag{142}$$

where Δ is the standard 3-dimensional Laplacian and $R_1 \geq 1$ is chosen so that $\mathcal{K} \subseteq \{r < R_1\}$.

Next, using the endpoint Hardy estimate (185) and truncating ϕ smoothly so it is supported away from $\mathcal{K}$, we can reduce (142) to the following global estimate on $\mathbb{R}^3$:

$$\|\tau_x^{-\frac{1}{2}} \partial \Delta^{-1} F\|_{\ell^1_r L^2_x(\mathbb{R}^3)} + \|\tau_x^{\frac{1}{2}} \partial^2 \Delta^{-1} F\|_{\ell^1_r L^2_x(\mathbb{R}^3)} \lesssim \|\tau_x^{\frac{1}{2}} F\|_{\ell^1_r L^2_x(\mathbb{R}^3)}.$$

This last inequality follows by taking the decomposition $\partial^J \Delta^{-1} F = \sum_{k,j} \chi_j \partial^J \Delta^{-1} \chi_k F$, where χ_j is a partition of unity adapted to dyadic regions $r \approx 2^k$, $2^j \geq 1$, and using Young's inequality to sum over

$$\|\tau_x^{-\frac{1}{2}} \chi_j \partial \Delta^{-1} \chi_k F\|_{L^2_x(\mathbb{R}^3)} + \|\tau_x^{\frac{1}{2}} \chi_j \partial^2 \Delta^{-1} \chi_k F\|_{L^2_x(\mathbb{R}^3)} \lesssim 2^j 2^k 2^{-2\max\{j,k\}} \|\tau_x^{\frac{1}{2}} \chi_k F\|_{L^2_x(\mathbb{R}^3)}$$

for $j, k \geq 0$. This final estimate follows from standard L^2 fractional/singular integral bounds. □

Finally we prove the commutator estimate:

Proof of estimate (137). Using (79) in $r < \frac{3}{4}\langle t\rangle$, we have for ϕ supported in that region

$$\|\tau_x^{\frac{1}{2}} [\Box_g, \Gamma^I] \phi\|_{\ell^1_r L^2_x} \lesssim \|\tau_x^{\frac{1}{2}} (\partial^2 \phi^{(k-1)}, \tau_x^{-1} \partial \phi^{(k-1)})\|_{\ell^\infty_r L^2_x} + \|\tau_x^{\frac{1}{2}} (\Box_g \phi)^{(k-1)}\|_{\ell^1_r L^2_x}.$$

It remains to estimate the first term on right-hand side above. First, note that by the same reasoning used to establish (140) we have

$$\|\tau_x^{\frac{1}{2}}(\partial\partial_u^b\phi^{(k-1)}, \tau_x^{-1}\partial_u^b\phi^{(k-1)}, \tau_+\partial\phi^{(k-1)})\|_{\ell_r^1 L_x^2} \lesssim \langle t\rangle^{-\frac{3}{2}}\|\phi\|_{E_k^1} + \|\tau_x^{\frac{1}{2}}\Box_g\phi^{(k-1)}\|_{L_x^2}$$

for functions ϕ supported in $r < \frac{3}{4}\langle t\rangle$. In addition we have by applying (136) to $\phi^{(k-1)}$ and then using the previous bound to handle the resulting middle term on the right-hand side of (136) the bound

$$\|\tau_x^{\frac{1}{2}}(\partial_x^2\phi^{(k-1)}, r^{-1}\partial_x\phi^{(k-1)})\|_{\ell_r^\infty L_x^2} \lesssim \|\phi^{(k-1)}\|_{H_x^2(r<R)} + \langle t\rangle^{-\frac{3}{2}}\|\phi\|_{E_k^1} + \|\tau_x^{\frac{1}{2}}\Box_g\phi^{(k-1)}\|_{\ell_r^1 L_x^2}.$$

Combining the last three inequalities and using (1) to handle $[\partial_u^b, \partial]$ we see that estimate (137) follows via induction on k. □

8. Estimates for nonlinear problems

This section is devoted to the proof of Theorem 1.22.

8A. *Proof of the $N_k \to S_k$ mapping property.* This section is devoted to the first half of Theorem 1.22.

Proof of (22). We treat each component of the norm separately.

Case 1: ($L^\infty H^s$*- and* WLE*-components*) These are handled via an induction on the index j. For $j = 0$ use (14). For $j \geq 1$ we again use (14) and estimate the middle term on the right-hand side via

$$\|\phi\|_{H_{j-1}^{16+3(k-j)}(r<B_0)[0,T]} \lesssim \|\phi\|_{\mathrm{WLE}_{j-1}^{13+3(k-j+1)}[0,T]}.$$

Case 2: ($S^{\frac{1}{2}}$*-components*) This term is handled by (15) with $a = \frac{1}{2}$. The middle term on right-hand side of (15) is bounded by the output of the previous step at level $j = k+4$ as follows:

$$\|\tau_+^{-\frac{1}{2}}\phi\|_{H_{k+4}^1(r<B_{1/2})[0,T]} \lesssim \|\phi\|_{\mathrm{WLE}_{k+4}^1[0,T]}.$$

Case 3: ($S^{1,\infty}$*- and* E^1*-components*) This follows from (16). The middle term on right-hand side of (16) is bounded by the output of the previous step as follows:

$$\|\phi\|_{\ell_t^1 H_{k+3}^1(r<B_{1/2})[0,T]} \lesssim \|\phi\|_{S_{k+3}^{1/2}[0,T]}.$$

Note also that expanding vector fields via the basis ∂^b shows

$$\|F\|_{N_{k+2}^{1,1}[0,T]} \lesssim \sum_{i+|J|\leq k+2} \|(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J F\|_{N^{1,1}[0,T]}.$$

Case 4: (L^∞*-components*) The bound for this term follows at once from (17). □

8B. *Proof of the null form estimate.* Here we prove the estimate (23). This may be broken up into a number of pieces according the constituent parts of (20) and (21).

Proposition 8.1 (constituent null form estimates). *Let $\mathcal{N} = \mathcal{N}^{\alpha\beta}(t,x,\phi)\partial_\alpha\phi\partial_\beta\phi$ denote a quadratic form satisfying the conditions of Definition 1.20. Then there exist locally bounded functions C_k, depending on the c_k from (19), such that for $k \geq 18$ one has*

$$\sum_{j=0}^{k+4} \|\mathcal{N}\|_{(\mathrm{WLE}^{*,13+3(k-j)}_j + L^1_t H^{13+3(k-j)}_{x,j})[0,T]} \lesssim C_k(|||\phi|||_{S_k[0,T]})|||\phi|||^2_{S_k[0,T]}, \tag{143}$$

$$\|\mathcal{N}\|_{N^{1/2}_{k+3}[0,T]} \lesssim C_k(|||\phi|||_{S_k[0,T]})|||\phi|||^2_{S_k[0,T]}, \tag{144}$$

$$\sum_{i+|J|\leq k+2} \|(\tau_-\partial^b_u)^i(\tau_x\partial^b_x)^J\mathcal{N}\|_{N^{1,1}[0,T]} \lesssim C_k(|||\phi|||_{S_k[0,T]})|||\phi|||_{S_k[0,T]}(|||\phi|||_{S_k[0,T]} + |||\Box_g\phi|||_{N_k[0,T]}), \tag{145}$$

$$\sum_{|J|\leq k} \|\tau_x^2(\tau_x\partial)^J\mathcal{N}(0)\|_{L^2_x} \lesssim C_k(|||\phi|||_{S_k[0,T]})|||\phi|||_{S_k[0,T]}(|||\phi|||_{S_k[0,T]} + |||\Box_g\phi|||_{N_k[0,T]}). \tag{146}$$

Remark 8.2. In the sequel we will prove Proposition 8.1 assuming $\mathcal{N} = \mathcal{N}^{\alpha\beta}(t,x)\partial_\alpha\phi\partial_\beta\phi$. In the more general case when $\mathcal{N}^{\alpha\beta}$ depends on ϕ as well, repeated application of the Leibniz rule leads to higher-order products of the form $\prod_{i=1}^m \phi^{(k_i)}\partial\phi^{(l_1)}\partial\phi^{(l_2)}$. Such cubic and higher-order expressions are much easier to handle via the uniform norms employed in this paper (e.g., they do not require a null structure), and treating them explicitly only serves to clutter notation.

We will prove (143)–(146) with the help of the following lemma:

Lemma 8.3 (Leibniz rules). *One has*:

(I) *Let $\mathcal{N}$ be a quadratic satisfying the conditions of Definition 1.20 but not depending on ϕ, and use the notation $\mathcal{N}(\phi,\psi) = \mathcal{N}^{\alpha\beta}(t,x)\partial_\alpha\phi\partial_\beta\psi$. Then if X^I denotes a product in $\{\partial^b_u, \partial^b_i - \omega^i\partial^b_u, S, \Omega_{ij}, \tau_-\partial^b_u, \tau_x\partial^b_x\}$, we have the identity*

$$X^I\mathcal{N}(\phi,\psi) = \sum_{I'+I''\subseteq I} \mathcal{N}_{I',I''}(X^{I'}\phi, X^{I''}\psi), \tag{147}$$

where each $\mathcal{N}_{I',I''}$ satisfies the conditions of Definition 1.20 as well.

(II) *Let f, g be smooth functions compactly supported in both time and in the exterior region $\{r > R_0\}$, where R_0 is given in Definition 1.7. Let w_a be a weight satisfying $|\partial^J w_a| \lesssim_J \tau_0^a$. Then one has the balanced product estimates*

$$\|w_0 fg\|_{\mathrm{WLE}^{*,s}} \lesssim \|\tau_x^{\frac12}\tau_-^{\frac12} f\|_{\ell^1_r\ell^1_u L^\infty}\|\tau_-^{-\frac12} g\|_{\ell^\infty_r\ell^\infty_u H^s} + \|\tau_x^{-\frac12} f\|_{\ell^\infty_r H^s}\|\tau_x g\|_{\ell^1_r L^\infty}, \tag{148}$$

$$\|w_a fg\|_{\mathrm{WLE}^{*,s}} \lesssim \|\tau_x\tau_0^a f\|_{\ell^1_r L^\infty}\|\tau_x^{-\frac12} g\|_{\ell^\infty_r H^s} + \|\tau_x^{-\frac12} f\|_{\ell^\infty_r H^s}\|\tau_x\tau_0^a g\|_{\ell^1_r L^\infty}. \tag{149}$$

In addition there are also hold the unbalanced versions

$$\|w_0 fg\|_{\mathrm{WLE}^{*,s}} \lesssim \|\tau_x^{\frac12}\tau_-^{\frac12}\partial^s f\|_{\ell^1_r\ell^1_u L^\infty}\|\tau_-^{-\frac12} g\|_{\ell^\infty_r\ell^\infty_u H^s}, \tag{150}$$

$$\|w_a fg\|_{\mathrm{WLE}^{*,s}} \lesssim \|\tau_x\tau_0^a\partial^s f\|_{\ell^1_r L^\infty}\|\tau_x^{-\frac12} g\|_{\ell^\infty_r H^s}, \tag{151}$$

where we are using the shorthand $|\partial^s f| = \sum_{|J|\leq s}|\partial^J f|$.

Proof of (147). By induction it suffices to prove this identity for a single vector field X. Using the notation $\mathcal{N}_X(\phi,\psi)=X\mathcal{N}(\phi,\psi)-\mathcal{N}(X\phi,\psi)-\mathcal{N}(\phi,X\psi)$, we have $\mathcal{N}_X^{\alpha\beta}=X(\mathcal{N}^{\alpha\beta})-\partial_\gamma^b(X^\alpha)\mathcal{N}^{\gamma\beta}-\partial_\gamma^b(X^\beta)\mathcal{N}^{\alpha\gamma}$. The first bound in (19) for $\mathcal{N}_X$ follows at once from this identity and the assumption of (19) for $\mathcal{N}$. To prove the second bound in (19) for $\mathcal{N}_X$, it suffices to study the contractions $\partial_\gamma^b(X^u)\mathcal{N}^{u\gamma}$ and $\partial_\gamma^b(X^u)\mathcal{N}^{\gamma u}$. For $\gamma=u$ the bound is again immediate from assuming (19). On the other hand for $\gamma=i$ we need $|(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\partial_i^b(X^u)|\lesssim_{i,J}\tau_0$ for each $X\in\{\partial_u^b,\partial_i^b-\omega^i\partial_u^b,S,\Omega_{ij},\tau_-\partial_u^b,\tau_x\partial_x^b\}$. This follows from inspection. □

Proof of estimates (148)–(149). First let both f,g be not only (spacetime) compactly supported in the exterior region $\mathbb{R}^4\setminus\mathbb{R}\times\mathcal{K}$, but also supported in dyadic regions $\tau_x\approx 2^j$ and $\tau_-\approx 2^k$. Then one has the Moser estimate

$$\|fg\|_{H^s}\lesssim\|f\|_{L^\infty}\|g\|_{H^s}+\|f\|_{H^s}\|g\|_{L^\infty}.$$

Multiplying this through by the appropriate combination of 2^j and 2^k and then summing in $\ell^1_r(\ell^2_u)$, we have both (148) and (149). □

Proof of estimates (150)–(151). This follows from distributing the derivatives and Hölder's inequality. □

Proof of estimate (143). Let $\chi_{\mathcal{C}}$ be a smooth cutoff which is $\equiv 1$ on the cylinder $\mathcal{C}=[1,T-1]\times\{r>R_0+1\}$ and which vanishes for $r<R_0$ and $t\in\big[0,\tfrac12\big]\cup\big[T-\tfrac12,T\big]$. Our plan is to show

$$\sum_{j=0}^{k+4}\|(1-\chi_{\mathcal{C}})\mathcal{N}\|_{L^1_tH^{13+3(k-j)}_{x,j}[0,T]}\lesssim|||\phi|||^2_{S_k[0,T]},\tag{152}$$

$$\sum_{j=0}^{k+4}\|\chi_{\mathcal{C}}\mathcal{N}\|_{\mathrm{WLE}_j^{*,13+3(k-j)}}\lesssim|||\phi|||^2_{S_k[0,T]}.\tag{153}$$

These bounds are further broken down into a number of cases.

Case 1: ($L^1_tH^s_{x,j}$ *bounds in* $r<R_0+1$) Expanding all derivatives into the product and discarding $\mathcal{N}^{\alpha\beta}$ it suffices to prove for $j\le k+4$ that

$$\|\partial^I\partial\phi^{(j_1)}\partial^{I'}\partial\phi^{(j_2)}\|_{L^1_tL^2_x(r<R_0+1)[0,T]}\lesssim|||\phi|||^2_{S_k[0,T]},\quad\text{when } j_1+j_2=j,\ \ |I|+|I'|\le 13+3(k-j).\tag{154}$$

There are now two subcases:

Case 1a: (*evenly split derivatives*) If both $|I|+j_1\le 10+3k-2j$ and $|I'|+j_2\le 10+3k-2j$, then each factor can absorb at least two more derivatives and still go in $L^2(r<R_0+1)[0,T]$. Thus, after an $L^2_x\to L^\infty_x$ Sobolev embedding on one factor we may use a product of

$$\|\partial^I\partial\phi^{(j_1)}\|_{L^2_tL^\infty_x(r<R_0+1)[0,T]}\lesssim|||\phi|||_{S_k[0,T]},\qquad \|\partial^{I'}\partial\phi^{(j_2)}\|_{L^2(r<R_0+1)[0,T]}\lesssim|||\phi|||_{S_k[0,T]}.$$

Case 1b: (*uneven split*) The alternative to the previous case is that one factor, say the first, is such that $|I|+j_1\ge 11+3k-2j$. But this forces $|I'|+j_2\le 2$. In particular for $k\ge 3$ we may use a product of

$$\|\partial^I\partial\phi^{(j_1)}\|_{L^\infty_tL^2_x(r<R_0+1)}\lesssim|||\phi|||_{S_k[0,T]},$$

$$\|\partial^{I'}\partial\phi^{(j_2)}\|_{L^1_tL^\infty_x(r<R_0+1)}\lesssim\|\tau_+^{\frac32}\partial^{I'}\partial\phi^{(j_2)}\|_{L^\infty(r<R_0+1)}\lesssim|||\phi|||_{S_k[0,T]}.$$

This completes the proof of (154).

Case 2: ($L^1_t H^s_{x,j}$ *bounds in* $[0, 1] \cup [T-1, T]$) In this case we show the analog of (154) where the integral is restricted to $([0, 1] \cup [T-1, T]) \times (\mathbb{R}^3 \setminus \mathcal{K})$. This follows by taking a product of two $L^\infty_t H^s_{x,j}$ bounds after an $L^2_x \to L^\infty_x$ embedding for the factor with the least number of derivatives.

Case 3: ($\mathrm{WLE}^{*,s}$ *bounds in* $\mathcal{C}$) Using (147) it suffices to show for $j \le k+4$ and $\alpha, \beta = u, 1, 2, 3$

$$\|\tilde{\chi}^2_{\mathcal{C}} \mathcal{N}^{\alpha\beta}_{j_1,j_2} \partial^b_\alpha \phi^{(j_1)} \partial^b_\beta \phi^{(j_2)}\|_{\mathrm{WLE}^{*,13+3(k-j)}} \lesssim \|\!|\phi|\!\|^2_{S_k[0,T]}, \quad \text{where } j_1 + j_2 = j, \tag{155}$$

where $\tilde{\chi}_{\mathcal{C}}$ is also supported in the exterior region $(0, T) \times \{r > R_0\}$. This estimate is further broken down based on the values of j_1, j_2 and α, β.

Case 3a: ($\max\{j_1, j_2\} \le k-1$) In this case we plan to use (148) and (149) to cleanly distribute the ∂^I-derivatives. To facilitate this freeze the values of j_1, j_2 and define

$$f_\alpha = \tilde{\chi}_{\mathcal{C}} \partial^b_\alpha \phi^{(j_1)}, \quad g_\beta = \tilde{\chi}_{\mathcal{C}} \partial^b_\beta \phi^{(j_2)}. \tag{156}$$

From (5) and the definition of S_k from (20) and we have

$$\|\tau_x^{\frac{1}{2}} \tau_-^{\frac{1}{2}} f_u\|_{\ell^1_u \ell^1_r L^\infty} + \|\tau_x \tau_0 f_u\|_{\ell^1_r L^\infty} + \|\tau_x f_i\|_{\ell^1_r L^\infty} + \|\tau_-^{-\frac{1}{2}} f_i\|_{\ell^\infty_u \ell^\infty_r H^s} + \|\tau_x^{-\frac{1}{2}} f_\alpha\|_{\ell^\infty_r H^s} \lesssim \|\!|\phi|\!\|_{S_k[0,T]}, \tag{157}$$

where $s = 13 + 3(k-j)$, with an identical set of estimates for the components of g.

Case 3a.1: (*uu-components*) Using (19) it suffices to show

$$\|w_1 f_u g_u\|_{\mathrm{WLE}^{*,13+3(k-j)}} \lesssim \|\!|\phi|\!\|^2_{S_k[0,T]}, \tag{158}$$

where $|\partial^I w_1| \lesssim \tau_0$. This follows from (149) with $a = 1$ and (157).

Case 3a.2: (*ui- and iu-components*) Here we need the analog of (158) with w_1 replaced by weight w_0 with $|\partial^I w_0| \lesssim 1$. This follows from (148) and (157).

Case 3a.3: (*ij-components*) In this case the analog of (158) follows from (149) with $a = 0$ and (157).

Case 3b: ($\max\{j_1, j_2\} \ge k$) In this case from the constraint $j_1 + j_2 = j \le k+4$ we must have both $\min\{j_1, j_2\} \le 4$ and $13 + 3(k-j) \le 13$. Without loss of generality assume $\min\{j_1, j_2\} = j_1$. Then with the notation from (156) and setting $s = 13 + 3(k-j)$, we have if $k \ge 18$ the bounds

$$\|\tau_x^{\frac{1}{2}} \tau_-^{\frac{1}{2}} \partial^s f_u\|_{\ell^1_u \ell^1_r L^\infty} + \|\tau_x \tau_0 \partial^s f_\alpha\|_{\ell^1_r L^\infty} + \|\tau_x \partial^s f_i\|_{\ell^1_r L^\infty} + \|\tau_-^{-\frac{1}{2}} g_i\|_{\ell^\infty_u \ell^\infty_r H^s} + \|\tau_x^{-\frac{1}{2}} g_\alpha\|_{\ell^\infty_r H^s} \lesssim \|\!|\phi|\!\|_{S_k[0,T]}.$$

In particular (155) for this case follows from the bound above, (150), and (151). □

The proof of (144) and (145) largely boils down to (147) and the following lemma:

Lemma 8.4. *Let $\mathcal{N}$ be any quadratic form which satisfies (19), and define $\mathcal{N}(\phi, \psi) = \mathcal{N}^{\alpha\beta} \partial_\alpha \phi \partial_\beta \psi$. Suppose X^I is a product of vector fields in $\{\partial^b_u, \partial^b_i - \omega^i \partial^b_u, S, \Omega_{ij}, \tau_- \partial^b_u, \tau_x \partial^b_x\}$. Then if $|I| \le k-1$ we have the pointwise bound on* $[0, T]$

$$|\tau_0^{\frac{1}{2}} \mathcal{N}(X^I \phi, \psi)| \lesssim \|\!|\phi|\!\|_{S_k[0,T]} \cdot \tau_x^{-1} \tau_+^{-\frac{3}{2}} \tau_0^{-1} \sum_{i+|J|=1} |(\tau_0 \partial^b_u)^i (\partial^b_x)^J \psi|. \tag{159}$$

A similar bound holds with the roles of ϕ and ψ reversed.

Proof. Expanding the $\mathcal{N}$ using condition (19) we have

$$|\tau_0^{\frac{1}{2}}\mathcal{N}(X^I\phi,\psi)| \lesssim \tau_0^{\frac{3}{2}}|\partial_u^b X^I\phi|\cdot|\partial_u^b\psi| + \tau_0^{\frac{1}{2}}|\partial X^I\phi|\cdot|\partial_x^b\psi| + \tau_0^{\frac{1}{2}}|\partial_x^b X^I\phi|\cdot|\partial\psi|.$$

On the other hand inspection of the L^∞-term from the S_k norm defined in (20) shows for $|I|\leq k-1$

$$\tau_0^{\frac{3}{2}}|\partial X^I\phi| + \tau_0^{\frac{1}{2}}|\partial_x^b X^I\phi| \lesssim \tau_x^{-1}\tau_+^{-\frac{3}{2}}|||\phi|||_{S_k[0,T]}.$$

Taking the product of the last two inequalities yields (159). □

Proof of estimate (144). Using (147) to distribute derivatives, and splitting into interior and exterior bounds, it suffices to show that when $j_1+j_2=k+3$

$$\|\tau_+\mathcal{N}_{j_1,j_2}(\phi^{(j_1)},\phi^{(j_2)})\|_{\ell_t^1 H^1(r<R_0)[0,T]} \lesssim |||\phi|||^2_{S_k[0,T]}, \tag{160}$$

$$\|\mathcal{N}_{j_1,j_2}(\phi^{(j_1)},\phi^{(j_2)})\|_{N^{1/2}(r>R_0)[0,T]} \lesssim |||\phi|||^2_{S_k[0,T]}. \tag{161}$$

Note that (160) is stronger than what we need here, and for this bound we can even assume all vector fields are in the collection $\{\partial_u^b,\ \partial_i^b-\omega^i\partial_u^b,\ S,\ \Omega_{ij},\ \tau_-\partial_u^b,\ \tau_x\partial_x^b\}$. We will use this greater generality in a moment.

<u>Case 1</u>: (*interior estimate*) From the conditions $k\geq 18$ and $j_1+j_2=k+3$ we have $\min\{j_1,j_2\}\leq k-2$. Then (160) follows by taking the product of

$$\|\tau_+\phi^{(k)}\|_{\ell_t^1 L^\infty(r<R_0)[0,T]} \lesssim |||\phi|||_{S_k[0,T]}, \qquad \|\phi^{(k+4)}\|_{H^1(r<R_0)[0,T]} \lesssim |||\phi|||_{S_k[0,T]}.$$

<u>Case 2</u>: (*exterior estimate*) Using only $\min\{j_1,j_2\}\leq k-1$ estimate (161) follows from (159) and

$$\sum_{i+|J|=1}\|\tau_x^{-\frac{1}{2}}(\tau_0\partial_u^b)^i(\partial_x^b)^J\phi^{(k+3)}\|_{\ell_r^1\ell_u^1 L^2[0,T]} \lesssim \|\tau_x^{-\frac{1}{2}}\tau_+^{\frac{1}{2}}(\tau_0\partial_u^b\phi^{(k+3)},\partial_x^b\phi^{(k+3)})\|_{\ell_r^\infty L^2[0,T]} \lesssim |||\phi|||_{S_k[0,T]}. \ \square$$

Proof of estimate (145). Combining (147), (159), and the Klainerman–Sideris identity (84) we have

$$\sum_{i+|J|\leq k+2}\tau_0^{\frac{1}{2}}|(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\mathcal{N}|$$
$$\lesssim |||\phi|||_{S_k[0,T]}\cdot\Big(\sum_{i+|J|=1}\tau_x^{-\frac{1}{2}}\tau_+^{-\frac{3}{2}}\tau_0^{-\frac{1}{2}}|(\tau_0\partial_u^b)^i(\partial_x^b)^J\phi^{(k+2)}| + \sum_{i+|J|\leq k+2}\tau_0^{\frac{1}{2}}|(\tau_-\partial_u^b)^i(\tau_x\partial_x^b)^J\Box_g\phi|\Big). \tag{162}$$

The proof of (145) then follows by splitting into interior and exterior estimates as in the proof of (144) above. Note that (160) already handles the interior contribution. For the exterior contribution we use

$$\sum_{i+|J|=1}\|(\tau_0\partial_u^b)^i(\partial_x^b)^J\phi^{(k+2)}\|_{\ell_u^1\ell_t^1\ell_r^1 L^2[0,T]} \lesssim \|\tau_x^{-\frac{1}{2}}\tau_+(\tau_0\partial_u^b\phi^{(k+2)},\partial_x^b\phi^{(k+2)})\|_{\ell_t^\infty\ell_r^\infty L^2[0,T]} \lesssim |||\phi|||_{S_k[0,T]}. \ \square$$

Proof of (146). This follows by applying (162) at $t=0$ with index restriction $i+|J|\leq k$. □

Appendix A: Coordinates

In this appendix we discuss some basic consequences of Definition 1.1, as well as some simple conditions which guarantee the assumptions of Definition 1.1 hold.

Bounds between (t, x)- and (u, x)-coordinates.

Lemma A.1. *Let $u(t, x)$ be a function satisfying condition (I) of Definition 1.1. Then for any smooth function q and integer $N \geq 0$ one has the following equivalence of symbol-type bounds*:

$$\sum_{i+|J|\leq N} \|(\tau_-\partial_u^b)^i(\tau_x\tau_0\partial_x^b)^J q\|_{\ell_r^1 L^\infty} \lesssim_{N,q} 1 \quad \Longleftrightarrow \quad \sum_{i+|J|\leq N} \|(\tau_-\partial_t)^i(\tau_x\tau_0\partial_x)^J q\|_{\ell_r^1 L^\infty} \lesssim_{N,q} 1. \tag{163}$$

In addition the change of frame bounds (5) also hold.

Proof. Thanks to the change of variables formula

$$\partial_u^b = \frac{1}{u_t}\partial_t, \qquad \partial_i^b = \partial_i - \frac{u_i}{u_t}\partial_t, \tag{164}$$

the implication "∂_t, ∂_x bounds" $\Rightarrow$ "$\partial_u^b, \partial_i^b$ bounds" follows easily from (assuming $i + |J| \leq N$)

$$u_t > c, \qquad |(\tau_-\partial_t)^i(\tau_x\tau_0\partial_x)^J(u_t, u_i)| \lesssim_N 1,$$

where the inequality above itself holds thanks to part (I) of Definition 1.1. Applying this to $q = \partial u$ we have

$$|(\tau_-\partial_u^b)^i(\tau_x\tau_0\partial_x^b)^J(u_t, u_i)| \lesssim_N 1.$$

Finally, using this last inequality and formula (164) the implication "$\partial_u^b, \partial_i^b$ bounds" $\Rightarrow$ "∂_t, ∂_x bounds" becomes clear.

As a last step notice that (5) follows from the formulas (164) and estimate (163) applied to $q = \partial u$. □

Lemma A.2. *Let $u(t, x)$ be a function satisfying part (I) of Definition 1.1; then $\tau_+^{-1}(u + \tau_x - t) \in \ell_r^1 L^\infty$.*

Proof. We have $\tau_+^{-1}(u + \tau_x - t) = \tau_+^{-1}\int_0^{(t,x)} \partial(u + \tau_x - t) \cdot ds + O(\tau_+^{-1})$, where ds denotes the line integral along a straight ray from the origin to (t, x). Bounding the integral in absolute value gives

$$\sup_{r\approx 2^j} |\tau_+^{-1}(u + \tau_x - t)| \lesssim 2^{-j} + \sum_{k\leq j} 2^{k-j} \sup_{r\approx 2^k} |\partial(u + \tau_x - t)|.$$

The assumption $\partial(u + \tau_x - t) \in \ell_r^1 L^\infty$ and Young's convolution inequality finishes the proof. □

Lemma A.3. *Let $u(t, x)$ be a function satisfying condition (I) of Definition 1.1, and suppose that g is a metric satisfying condition (II). Then g is weakly asymptotically flat in (t, x)-coordinates in the sense that*

$$\|(\tau_-\partial_t)^i(\tau_x\tau_0\partial_x)^J(g - \eta)^{\alpha\beta}\|_{\ell_r^1 L^\infty} < \infty \quad \textit{for all } (i, J) \in \mathbb{N} \times \mathbb{N}^4, \tag{165}$$

where $\eta = \mathrm{diag}(-1, 1, 1, 1)$ is the Minkowski metric.

Proof. By Lemma A.1 is suffices to prove the bound

$$\|(\tau_-\partial_u^b)^i(\tau_x\tau_0\partial_x^b)^J(g-\eta)^{\alpha\beta}\|_{\ell_r^1L^\infty}<\infty \quad \text{for all } (i,J)\in\mathbb{N}\times\mathbb{N}^4, \tag{166}$$

where $(g-\eta)^{\alpha\beta}$ still denotes the components in (t,x)-coordinates. Such estimates for $(g-\eta)^{ij}$ follow at once from the first inclusion in (3) because g^{ij} is the same in either (t,x)- or (u,x)-coordinates.

For remaining components we compute

$$\begin{aligned} g^{ti} &= (g^{ui}+\omega^i)+\omega_j(g^{ij}-\delta^{ij})-g^{\alpha i}\partial_\alpha^b(u+\tau_x-t),\\ g^{tt}+1 &= g^{uu}+2(g^{ui}+\omega^i)\omega_i+\omega_i\omega_j(g^{ij}-\delta^{ij})-2(g^{u\alpha}+\omega_i g^{i\alpha})\partial_\alpha^b(u+\tau_x-t)\\ &\qquad + g^{\alpha\beta}\partial_\alpha^b(u+\tau_x-t)\partial_\beta^b(u+\tau_x-t)+\tau_x^{-2}, \end{aligned}$$

where all metric components on the right-hand side are now computed in (u,x)-coordinates, and where we are using the notation $\omega^i=\omega_i=x^i\tau_x^{-1}$. In addition to these formulas we also have the estimate

$$\|(\tau_-\partial_u^b)^i(\tau_x\tau_0\partial_x^b)^J\partial^b(u+\tau_x-t)\|_{\ell_r^1L^\infty}<\infty,$$

which itself is a consequence of (1), (164), and Lemma A.1. The remaining portion of estimate (166) follows from the last three displays above combined with assumption (3). □

Lemma A.4. *Fix $\delta>0$. Let u_1 be an approximate optical function satisfying the conditions of Definition 1.1 in the region $\langle t-r\rangle\geq\delta\langle t+r\rangle$, and let u_2 be an approximate optical function satisfying the conditions of Definition 1.1 in the region $\langle t-r\rangle\leq 2\delta\langle t+r\rangle$. Then if χ is any cutoff function with $\chi\equiv 1$ on $\langle t-r\rangle\geq 2\delta\langle t+r\rangle$, $\chi\equiv 0$ on $\langle t-r\rangle\leq\delta\langle t+r\rangle$, and $|(\tau_x\partial)^J\chi|\lesssim 1$, the function $u=\chi u_1+(1-\chi)u_2$ satisfies the conditions of Definition 1.1 globally. In particular, in Definition 1.1 we may always assume $u=t-\tau_x$ away from the region $\langle t-r\rangle\ll\langle t+r\rangle$.*

Proof. Using Lemma A.2 we have both

$$\begin{aligned} &\|(\langle t\rangle\partial_t)^i(\tau_x\partial_x)^J\tau_+^{-1}(u_1+\tau_x-t)\|_{\ell_r^1L^\infty(\delta\leq\langle t+r\rangle^{-1}\langle t-r\rangle\leq 2\delta)}<\infty,\\ &\|(\langle t\rangle\partial_t)^i(\tau_x\partial_x)^J\tau_+^{-1}(u_2+\tau_x-t)\|_{\ell_r^1L^\infty(\delta\leq\langle t+r\rangle^{-1}\langle t-r\rangle\leq 2\delta)}<\infty. \end{aligned}$$

Thus $\tau_+^{-1}(u_1-u_2)$ satisfies the same bound in $\delta\leq\langle t+r\rangle^{-1}\langle t-r\rangle\leq 2\delta$, and so $u=\chi u_1+(1-\chi)u_2$ satisfies (1) globally. Note that $\tau_+^{-1}\langle u_1\rangle\approx\tau_+^{-1}\langle u\rangle\approx 1$ in $\langle t-r\rangle\geq\delta\langle t+r\rangle$ and $\tau_+^{-1}\langle u_2\rangle\approx\tau_+^{-1}\langle u\rangle\approx 1$ in $\langle t-r\rangle\leq\delta\langle t+r\rangle$ thanks to Lemma A.2, so the definition of τ_0 is not affected by splicing u_1,u_2.

It remains to show the bounds (3) hold in (u,x)-coordinates. Because $u=u_2$ when $\langle t-r\rangle\leq\delta\langle t+r\rangle$, we concentrate on the complementary region. Here it suffices to show that if u is any function satisfying (1), and g is any metric satisfying (165), then one has automatically has the first inclusion in (3) restricted to the region $\langle t-r\rangle\geq\delta\langle t+r\rangle$. Notice that by combining (165) and (1), we see that (165) also holds for all Bondi coordinate components of $(g^{\alpha\beta}-\eta^{\alpha\beta})$. Therefore, adding and subtracting the tensor $h^{\alpha\beta}$ defined in (4), our task boils down to showing

$$\|(\langle t\rangle\partial_t)^i(\tau_x\partial)^J(\eta^{uu},\eta^{ui}+\omega^i)\|_{\ell_r^1L^\infty(\langle t+r\rangle^{-1}\langle t-r\rangle\geq\delta)}<\infty.$$

This last inequality follows from a few simple calculations and the estimate

$$\|(\langle t\rangle\partial_t)^i(\tau_x\partial)^J\partial(u+\tau_x-t)\|_{\ell^1_r L^\infty(\tau_+^{-1}\langle t-r\rangle\geq\delta)}<\infty,$$

which is an immediate consequence of (1). □

Constructions for nearly stationary/spherically symmetric metrics. In this section we discuss a simple situation where one can construct an approximate "optical function" $u(t,x)$ satisfying conditions (3). This is given by the following definitions.

Definition A.5. Let $g_{\alpha\beta}$ be a Lorentzian metric on $[0,\infty)\times(\mathbb{R}^3\setminus\mathcal{K})$, where $\mathcal{K}$ is a compact set. Then:

(i) g is called "weakly asymptotically flat and quasistationary" if

$$\|\ln^2(1+\tau_x)(t\partial_t)^i(\tau_x\partial)^J(g_{\alpha\beta}-\eta_{\alpha\beta})\|_{\ell^1_r L^\infty}<\infty\quad\text{for all } (i,J)\in\mathbb{N}\times\mathbb{N}^4. \tag{167}$$

Here $\eta=\mathrm{diag}(-1,1,1,1)$ is the Minkowski metric in $(t,x)\in\mathbb{R}\times\mathbb{R}^3$-coordinates.

(ii) g is called "quasispherical" if one can write $g=g_0+g_1$, where g_0 is a spherically symmetric in (t,x)-coordinates, and the remainder g_1 satisfies

$$\|\tau_x(\tau_x\partial)^J(g_1)_{\alpha\beta}\|_{\ell^1_r L^\infty}<\infty\quad\text{for all } J\in\mathbb{N}^4. \tag{168}$$

Proposition A.6. *Let $g_{\alpha\beta}$ be a Lorentzian metric on $[0,\infty)\times(\mathbb{R}^3\setminus\mathcal{K})$, where $\mathcal{K}$ is a compact set. Suppose that g is weakly asymptotically flat and quasistationary/spherical in the sense that (167) and (168) both hold. Then g satisfies the assumptions of Definition 1.1 (after a possible redefinition of the x^i-coordinates).*

Proof. We'll prove this in a series of steps.

Step 1: (*preliminary reduction*) By Lemma A.4 above it suffices to construct an approximate optical function $u(t,x)$ satisfying conditions (1) and (3) in the region $\langle t-r\rangle\ll\langle t+r\rangle$. Using a partition of unity we may extend g to be the Minkowski metric in the exterior $\langle t-r\rangle\gtrsim\langle t+r\rangle$. This extension will still satisfy (167) and (168).

Next, after a possible radial change of variables which preserves both (167) and (168), we may assume that the area of $t=\text{const}$ and $r=\text{const}$ with respect to the restriction of g_0 is $4\pi r^2$. In other words we may assume the spherically symmetric part g_0 can be written in polar coordinates as

$$g_0=(g_0)_{tt}\,dt^2+2(g_0)_{tr}\,dt\,dr+(g_0)_{rr}\,dr^2+r^2\,d\sigma^2, \tag{169}$$

where $d\sigma^2$ is the standard round metric on $\mathbb{S}^2$.

The goal now is to construct u in two pieces $u=u_0+u_1$, where $u_0=u_0(t,r)$ is radially symmetric and corresponds to g_0, while the remainder u_1 takes into account $g_1=g-g_0$. The requirements for these two functions will be

$$\|\ln(r)(r\partial)^J\partial(u_0+r-t)\|_{\ell^1_r L^\infty(r\geq R)}+\|r(r\partial)^J\partial u_1\|_{\ell^1_r L^\infty(r\geq R)}\lesssim_J 1 \tag{170}$$

for sufficiently large R, and in addition

$$g_0(du_0,du_0)=2g_0(du_0,du_1)+g_1(du_0,du_0)=0. \tag{171}$$

Notice that (170) and formulas (164) allow us to freely change $(r\partial)^J$ to $(r\partial^b)^J$ in any estimate we consider.

First suppose that we have achieved both (170) and (171). By the assumption (167) and (170), we have

$$\| \ln(r)(r\partial)^J(g^{\alpha\beta} - \eta^{\alpha\beta})\|_{\ell^1_r L^\infty(r\geq R)} + \| \ln(r)(r\partial)^J(\eta^{uu}, \eta^{ui} + \omega^i)\|_{\ell^1_r L^\infty(r\geq R)} < \infty,$$

where all components of $(g^{\alpha\beta} - \eta^{\alpha\beta})$ are computed in (u, x)-coordinates. This suffices to give the first inclusion in (3) (note we only need this for $\langle t-r\rangle \ll \langle t+r\rangle$). We remark that the convergence factor $\ln(r)$ is sufficient to sum in ℓ^1_u when $r \approx t \approx 2^j$.

Next, from the explicit form (169) and the identities in (171), we have both

$$\sqrt{|g_0|}g_0^{u_0 i} + r^{-1}x^i = 0, \quad g^{uu} = g(du_1, du_1) + 2g_1(du_0, du_1),$$

where $\sqrt{|g_0|}$ is computed in (u_0, x)-coordinates. Using (168) and (170) we have

$$\|r(r\partial)^J(\sqrt{|g_0|} - \sqrt{|g|})\|_{\ell^1_r L^\infty(r\geq R)} < \infty,$$

where $\sqrt{|g|}$ is computed in (u, x)-coordinates. Combining the last two displays and (170) again gives

$$\|r^2(r\partial)^J g^{uu}\|_{\ell^1_r L^\infty(r\geq R)} + \|r(r\partial)^J(\sqrt{|g|}g^{ui} + \omega^i, g^{ui} - \omega^i\omega_j g^{uj})\|_{\ell^1_r L^\infty(r\geq R)} < \infty,$$

which are sufficient to produce the remaining bounds in (3) (again for $\langle t-r\rangle \ll \langle t+r\rangle$).

It remains to construct u_0 and u_1 such that (170) and (171) hold.

Step 2: (*construction of* u_0) For g_0 we have the expression (169), where

$$\| \ln^2(r)(r\partial)^J(g_0^{tt} + 1, g_0^{rr} - 1, g_0^{tr})\|_{\ell^1_r L^\infty(r\geq 1)} < \infty \quad \text{for all } J \in \mathbb{N}^2.$$

For the remainder of the construction we only need to work in the (t, r)-coordinates.

Let $v = v(t, r)$ be any function which solves the radial eikonal equation $g_0^{\alpha\beta}\partial_\alpha v \partial_\beta v = 0$. From this we define the quantity $\zeta = \partial v - (1, -1)$, where ∂ denotes (t, r)-derivatives. As long as $|\zeta| < 1$ the coordinate change $(t, r) \mapsto (u, r)$ is well-defined and we have

$$\partial_r^b \zeta = G(t, r, \zeta), \quad \partial^b = q(\zeta)\partial, \tag{172}$$

where both G and q are smooth universally defined functions depending only on the (t, r)-components of g_0 and not on v. Moreover $q = q_0 + q_1(\zeta)$ with q_0 a constant invertible matrix and $q_1 = O(\zeta)$ when $|\zeta| \leq \frac{1}{2}$. Finally we have the uniform symbol bounds

$$\|r \ln^2(r)(r\partial)^J(\partial_\zeta)^k G\|_{\ell^1_r L^\infty(r\geq R)} \lesssim_{J,k} o_R(1) \quad \text{for } |\zeta| \leq \tfrac{1}{2}. \tag{173}$$

Now let $\tilde{u} = v$ in the previous construction with initial normalization $\partial_r \tilde{u} < 0$ and $\tilde{u}|_{r=R} = t - R$ for sufficiently large R, and set $\tilde{\zeta} = \partial\tilde{u} - (1, -1)$. Then $\tilde{u}$ is globally defined and smooth in $r > R$ because $(1+1)$ Lorentzian metrics have no caustics. We also have $|\tilde{\zeta}| \ll 1$ at least initially close to $r = R$. Commuting (172) with vector fields $\partial^b = q(\zeta)\partial$, and applying a straightforward bootstrapping argument, we may extend this to uniform bounds $|\tilde{\zeta}| \ll 1$ and $|\partial^J\tilde{\zeta}| \lesssim_J 1$ for all $r > R$.

Next, define the outgoing limit

$$f(\tilde{u}) = \lim_{r\to\infty} \partial_t \tilde{u} = 1 + \int_R^\infty G_t(t(\tilde{u}, r), r, \tilde{\zeta})\, dr,$$

where G_t denotes the t-component of G. By the previous paragraph we have both $|f-1| \ll 1$ and $|\partial_{\tilde{u}}^j f| \lesssim_j 1$. Let F solve $F' = 1/f$, and finally set $u_0 = F(\tilde{u})$. Again we let $\zeta = \partial u_0 - (1, -1)$, which we remind the reader solves (172) with conditions (173).

By construction we immediately have $g_0^{\alpha\beta}\partial_\alpha u_0 \partial_\beta u_0 = 0$, $|\zeta| \ll 1$, and $|\partial^J \zeta| \lesssim_J 1$. In addition to this and the fact that $\partial_t u_0 = (1/f)\tilde{u}_t$, we have $\lim_{r\to\infty} \partial_t u_0 = 1$. Combining this last piece of information with the $r\to\infty$ limits $g_0^{tt} \to -1$, $g_0^{rr} \to 1$, $g_0^{tr} \to 0$, as well as $\partial_r u_0 < 0$, we have the normalization $\zeta \to 0$ as $r \to \infty$. In particular we may write

$$\zeta(u_0, r) = -\int_r^\infty G(t(u_0, s), s, \zeta)\, ds.$$

Differentiating this identity any number of times with respect to $\partial^b = q(\zeta)\partial$, and using the weighted estimate (173) and a straightforward bootstrapping argument, gives the first bound in (170).

Step 3: (*construction of* u_1) Using the second identity from (171), we have that the correction u_1 solves the linear equation

$$g_0^{\alpha\beta}\partial_\alpha u_0 \partial_\beta u_1 = -\tfrac{1}{2} g_1^{\alpha\beta}\partial_\alpha u_0 \partial_\beta u_0, \qquad u_1 \to 0 \text{ as } r \to \infty.$$

Integrating this in (u_0, x)-coordinates we find that

$$u_1(u_0, x) = \frac{1}{2}\int_{|x|}^\infty (g_1^{u_0 u_0}/g_0^{r u_0})(u_0, sx/|x|)\, ds.$$

By the first estimate in (170) and assumptions (167)–(168) we have

$$\| r(r\partial^b)^J (g_1^{u_0u_0}/g_0^{ru_0}) \|_{\ell_r^1 L^\infty(r\geq R)} < \infty,$$

where ∂^b are computed in (u_0, x)-coordinates. An application of Young's convolution inequality to the integral above gives (170) for u_1. □

Appendix B: Local energy decay

In this appendix we discuss how assumption (7b) relates to assumption (7a) when the metric g enjoys structural properties similar to the Kerr family of metrics with angular momentum in a moderate range.

Let $\mathcal{T} \subseteq \mathbb{R}^3$ be a compact region contained in the exterior $\mathbb{R}^3 \setminus \mathcal{K}$. We first redefine the norms (6b) and (6d) so the regularity loss occurs only on $\mathcal{T}$:

$$\|\phi\|_{\mathrm{WLE}^s_{\mathrm{class}}[0,T]} = \sum_{|J|\leq s} (\|\tau_x^{-1}\partial^J\phi\|_{\mathrm{LE}[0,T]} + \|\partial\partial^J\phi\|_{\mathrm{LE}(\mathcal{T}^c)[0,T]}), \tag{174}$$

$$\|F\|_{\mathrm{WLE}^{*,s}[0,T]} = \sum_{|J|\leq s} (\|\partial^J F\|_{\mathrm{LE}^*[0,T]} + \|\partial\partial^J F\|_{L^2(\mathcal{T})[0,T]}). \tag{175}$$

With respect to these modified norms we have:

Proposition B.1. *Let the norms* $\mathrm{WLE}^s_{\text{class}}$ *and* $\mathrm{WLE}^{*,s}_{\text{class}}$ *be defined as in* (174) *and* (175). *Suppose in addition that* ∂_t *is uniformly timelike on* $[0,\infty)\times\mathcal{T}$. *Then estimate*

$$\sup_{0\le t\le T}\|\partial\phi(t)\|_{H^s_x}+\|\phi\|_{\mathrm{WLE}^s_{\text{class}}[0,T]}\lesssim\|\partial\phi(0)\|_{H^s_x}+\|\Box_g\phi\|_{\mathrm{WLE}^{*,s}[0,T]} \tag{176}$$

implies estimate

$$\sup_{0\le t\le T}\|\partial\phi(t)\|_{H^s_x}+\|\phi\|_{\mathrm{LE}^s_{\text{class}}[0,T]} \lesssim\|\partial\phi(0)\|_{H^s_x}+\|\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]}+\|\partial_t\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]}+\|\Box_g\phi\|_{\mathrm{LE}^{*,s}[0,T]}, \tag{177}$$

where $\widetilde{\mathcal{T}}$ *is any compact neighborhood of* $\mathcal{T}$ *(here the implicit constant depends on* $\widetilde{\mathcal{T}}$*).*

Remark B.2. We do not need to make other assumptions on the metric g aside from boundedness of its derivatives and ∂_t being uniformly timelike on $[0,\infty)\times\mathcal{T}$. In particular the size of the time variation of g plays no role in establishing (177) from (176).

Note that (177) implies (7b) when $s=0$. For $s>0$ a simple induction allows us to reduce the second term in the right-hand side of (177) to $\|\phi\|_{L^2(\widetilde{\mathcal{T}})[0,T]}$.

Proof. Let $\mathcal{T}\Subset\mathcal{T}'\Subset\widetilde{\mathcal{T}}$, where $\mathcal{T}'$ is an intermediate compact neighborhood. Without loss of generality we may assume $\widetilde{\mathcal{T}}$ is a small enough that ∂_t is still uniformly timelike on it. Estimate (177) will be shown by adding together the pair of bounds

$$\sup_{0\le t\le T}\|\partial\phi(t)\|_{H^s_x((\mathcal{T}')^c)}+\|\phi\|_{\mathrm{LE}^s_{\text{class}}((\mathcal{T}')^c)[0,T]}\lesssim\|\partial\phi(0)\|_{H^s_x}+\|\phi\|_{H^{s+1}(\mathcal{T}')[0,T]}+\|\Box_g\phi\|_{\mathrm{LE}^{*,s}[0,T]}, \tag{178}$$

$$\sup_{0\le t\le T}\|\partial\phi(t)\|_{H^s_x(\mathcal{T}')}+\|\phi\|_{H^{s+1}(\mathcal{T}')[0,T]}\lesssim\|\partial\phi(0)\|_{H^s_x}+\|\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]} +\|\partial_t\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]}+\|\Box_g\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]}. \tag{179}$$

The first estimate (178) follows directly by applying (176) to $(1-\chi_{\mathcal{T}})\phi$, where $\chi_{\mathcal{T}}=1$ on $\mathcal{T}$ and $\chi_{\mathcal{T}}=0$ on $(\mathcal{T}')^c$, and then using the Hardy estimate (184) with $a=0$ for the boundary term where all derivatives fall on the cutoff.

The second estimate (179) is slightly more involved so we do it in a series of steps. For the demonstration it will be convenient to fix an additional pair of intermediate spatial regions $\mathcal{T}'\Subset\mathcal{T}''\Subset\mathcal{T}'''\Subset\widetilde{\mathcal{T}}$, and an $\epsilon>0$ sufficiently small that the domain of dependence of $\mathcal{T}'''$ is contained in $\widetilde{\mathcal{T}}$ for each fixed time in the slab $[0,\epsilon]$, and the domain of dependence of $\mathcal{T}'$ is contained in $\mathcal{T}''$ for each fixed time in the slab $[T-\epsilon,T]$.

<u>Step 1</u>: (*elliptic bound in* $[\epsilon,T-\epsilon]\times\mathcal{T}'''$) Our assumption is that the spatial part of $\Box_g$, i.e., the operator $(1/\sqrt{|g|})\partial_i\sqrt{|g|}g^{ij}\partial_j$ for $i,j=1,2,3$, is uniformly elliptic in $\widetilde{\mathcal{T}}$. In particular the operator $P(t,x,D)=\lambda\partial_t^2+\Box_g$ is uniformly elliptic in $\widetilde{\mathcal{T}}$ for sufficiently large $\lambda\in\mathbb{R}$. It follows that for any $s\ge0$ one has the spacetime elliptic bound

$$\|\chi\phi\|_{H^{s+1}}\lesssim\|\chi\phi\|_{H^s}+\|\partial_t(\chi\phi)\|_{H^s}+\|\Box_g(\chi\phi)\|_{H^{s-1}},$$

where $\chi\equiv1$ on $[\epsilon,T-\epsilon]\times\mathcal{T}'''$ and $\chi\equiv0$ outside of $[0,T]\times\widetilde{\mathcal{T}}$. Writing $\Box_g(\chi\phi)=2\nabla^\alpha(\phi\nabla_\alpha\chi)-\phi\Box_g\chi+\chi\Box_g\phi$ and removing the cutoff, the previous inequality implies

$$\|\phi\|_{H^{s+1}(\mathcal{T}''')[\epsilon,T-\epsilon]}\lesssim\|\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]}+\|\partial_t\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]}+\|\Box_g\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]}. \tag{180}$$

Step 2: (*filling in the bottom time slab*) Fixed time local energy estimates based on the domain of dependence of $\mathcal{T}'''$ in the slab $[0,\epsilon]$ give us

$$\sup_{0\leq t\leq\epsilon} \|\partial\phi(t)\|_{H^s_x(\mathcal{T}''')} \lesssim \|\partial\phi(0)\|_{H^s_x} + \|\Box_g\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]}.$$

Integrating these with respect to L^2_t and adding them to (180) yields the improvement

$$\|\phi\|_{H^{s+1}(\mathcal{T}''')[0,T-\epsilon]} \lesssim \|\partial\phi(0)\|_{H^s_x} + \|\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]} + \|\partial_t\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]} + \|\Box_g\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]}. \tag{181}$$

Step 3: (*fixed-time energies on the slab* $[0,T-\epsilon]\times\mathcal{T}''$) Next, we let χ be a spatial cutoff with $\chi\equiv 1$ on $\mathcal{T}''$ and $\chi\equiv 0$ outside $\mathcal{T}'''$. Using the multiplier $\chi\partial_t$ and the left-hand side of (181) to control the spacetime error we have

$$\sup_{0\leq t\leq T-\epsilon} \|\partial\phi(t)\|_{H^s_x(\mathcal{T}'')} \lesssim \|\partial\phi(0)\|_{H^s_x} + \|\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]} + \|\partial_t\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]} + \|\Box_g\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]}. \tag{182}$$

Step 4: (*fixed-time energies on the slab* $[T-\epsilon,T]\times\mathcal{T}'$) Using the domain of dependence of $\mathcal{T}'$ in $[T-\epsilon,T]$ and the left-hand side of (182) at time $T-\epsilon$, we have

$$\sup_{T-\epsilon\leq t\leq T} \|\partial\phi(t)\|_{H^s_x(\mathcal{T}')} \lesssim \|\partial\phi(0)\|_{H^s_x} + \|\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]} + \|\partial_t\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]} + \|\Box_g\phi\|_{H^s(\widetilde{\mathcal{T}})[0,T]}. \tag{183}$$

Finally, adding together (181), (182), (183), and the L^2_t integral of the left-hand side of (183), we have (179). □

Appendix C: Hardy and trace inequalities

Lemma C.1 (Hardy inequalities). *Let $\mathcal{K}\subseteq\mathbb{R}^3$ be compact with connected complement, and for any other $Q\subseteq\mathbb{R}^3$ define $L^2_x(Q)$ where the domain of integration is $Q\setminus\mathcal{K}$, and $L^2(Q)[0,T]$ where the domain of integration is $[0,T]\times(Q\setminus\mathcal{K})\subseteq\mathbb{R}^4$. Then for test functions ϕ one has the following:*

(I) *For all $R\geq 0$ there hold uniformly*

$$\|\tau_x^{a-1}\phi\|_{L^2_x(r>R)} \lesssim_a \|\tau_x^a\partial_x\phi\|_{L^2_x(r>R)}, \quad \text{when } -\tfrac{1}{2}<a<\infty, \tag{184}$$

$$\|\tau_x^{-\frac{3}{2}}\phi\|_{\ell^\infty_r L^2_x(r>R)} \lesssim \|\tau_x^{-\frac{3}{2}}\phi\|_{L^2_x(\frac{1}{2}R<r<R)} + \|\tau_x^{-\frac{1}{2}}\partial_x\phi\|_{\ell^1_r L^2_x(r>\frac{1}{2}R)}. \tag{185}$$

(II) *When $R\geq 1$ is large enough that $\mathcal{K}\subseteq\{r<\frac{1}{2}R\}$ there is the fixed-time estimate*

$$\|\phi\|_{L^2_x(r>R)} \lesssim \|\tau_-\tau_x^{-1}\partial_x(\tau_x\phi)\|_{L^2_x(r>R)} + R^{-\frac{1}{2}}\|\tau_-^{\frac{1}{2}}\phi\|_{L^2_x(\frac{1}{2}R<r<R)} + R^{\frac{1}{2}}\|\tau_-^{\frac{1}{2}}\partial\phi\|_{L^2_x(\frac{1}{2}R<r<R)}. \tag{186}$$

(III) *Again for $R\geq 1$ large enough that $\mathcal{K}\subseteq\{r<\frac{1}{2}R\}$ there is the spacetime estimate*

$$\begin{aligned}&\|\tau_x^{a-1}\phi(T)\|_{L^2_x(r>R)} + \|\tau_x^{a-\frac{3}{2}}\phi\|_{L^2(r>R)[0,T]}\\ &\lesssim_a \|\tau_x^{a-\frac{3}{2}}\partial_r^b(\tau_x\phi)\|_{L^2(r>\frac{1}{2}R)[0,T]} + \|\tau_x^{a-\frac{3}{2}}\phi\|_{L^2(\frac{1}{2}R<r<R)[0,T]} + \|\tau_x^a\partial\phi(0)\|_{L^2_x(r>\frac{1}{2}R)}, \quad \text{when } a<1.\end{aligned} \tag{187}$$

Lemma C.2 (weighted trace inequalities). *For $T \geq 0$ and any $R \geq 1$ one has the uniform bound*

$$\|\tau_-^{\frac{1}{2}}\phi(T)\|_{L_x^2(\frac{1}{2}R<r<R)} \lesssim \|(\tau_-\partial_u^b\phi, \phi)\|_{L^2(\frac{1}{2}R<r<R)[\frac{1}{2}T,T]}, \qquad R \lesssim T, \tag{188a}$$

$$\|\tau_-^{\frac{1}{2}}\phi(T)\|_{L_x^2(\frac{1}{2}R<r<R)} \lesssim \|(\tau_-\partial_u^b\phi, \phi)\|_{L^2(\frac{1}{2}R<r<R)[0,T]} + \|\tau_x^{\frac{1}{2}}\phi(0)\|_{L_x^2(\frac{1}{2}R<r<R)}, \qquad R \gg T, \tag{188b}$$

$$\|\tau_-^{\frac{1}{2}}\phi(T)\|_{L_x^2(r<R)} \lesssim \|(\tau_-\partial_u^b\phi, \phi)\|_{L^2(r<R)[0,T]} + \|\tau_x^{\frac{1}{2}}\phi(0)\|_{L_x^2(r<R)}. \tag{188c}$$

We note that estimate (188a) also holds with the restriction $\frac{1}{2}R < r < R$ replaced by $r < R$.

Proof of (184). Let $R_1 \geq 1$ be chosen so that $\mathcal{K} \subseteq \{r < R_1\}$. First we prove the bound assuming $R \geq R_1$. We have $I = \int_R^\infty \int_{\mathbb{S}^2} \partial_r(r^{2a+1}\phi^2)\, d\omega\, dr \leq 0$, and also

$$(2a+1)\|r^{a-1}\phi\|^2_{L_x^2(r>R)} \leq I + 2\|r^{a-1}\phi\|_{L_x^2(r>R)}\|r^a\partial\phi\|_{L_x^2(r>R)},$$

which concludes the proof in this case because $2a+1 > 0$ and $r^\alpha \approx \tau_x^\alpha$ in $r > 1$.

Now suppose $0 \leq R < R_1$, where R_1 is as above. A standard compactness argument shows

$$\|\phi\|_{L_x^2(R<r<R_1)} \lesssim \|\phi\|_{L_x^2(R_1<r<2R_1)} + \|\partial\phi\|_{L^2(R<r<2R_1)}, \tag{189}$$

where the implicit constant is uniform in R. Combining this bound with estimate (184) in $r > R_1$ completes the proof. □

Proof of (185). Using estimates of the form (189) it suffices to show for $R \geq R_1$, where R_1 is as above, that

$$\|\tau_x^{-\frac{3}{2}}\phi\|^2_{\ell^\infty L_x^2(r>2R)} \lesssim \|\tau_x^{-\frac{3}{2}}\phi\|_{\ell^\infty L_x^2(r>R)}\|\tau_x^{-\frac{1}{2}}\partial\phi\|_{\ell^1 L_x^2(r>R)}.$$

To prove it choose $h_k(r)$ so that $h_k(R) = 0$ and $h_k' = 2^{-k}\chi_k$, where $\chi_k = 1$ when $R2^k \leq r \leq R2^{k+1}$ for $k \geq 1$, and $\chi_k = 0$ when either $r \leq R2^{k-1}$ or $r \geq R2^{k+2}$. Then computing the integral $I = \int_R^\infty \int_{\mathbb{S}^2} \partial_r(h_k\phi^2)\, d\omega\, dr = 0$ and taking $\sup_{k\geq 1}$ of the result yields the estimate in the display above. □

Proof of (186). Thanks to Remark 1.3 and the conditions (1) we can assume the coordinate u is chosen so that $-C \leq \partial_r u \leq -1/C$ for some fixed $C > 0$. Then (186) follows from integration of $\partial_r(\chi_{>R}u(\tau_x\phi)^2)$ with respect to $dr\, d\omega$. □

Proof of (187). This boils down to integration of the quantity $\partial_r^b(\chi_{r>R}r^{2a-2}(\tau_x\phi)^2)$ with respect to the measure $du\, dr\, d\omega$ over the slab $0 \leq t \leq T$. By Remark 1.3 and the conditions (1) we have the pair of bounds $1/C \leq \partial_r^b t,\ \partial_t u \leq C$. In particular

$$\int_0^T\int_0^\infty\int_{\mathbb{S}^2} \partial_r^b(\chi_{r>R}r^{2a-2}(\tau_x\phi)^2)\alpha\, d\omega\, dr\, dt = \int_0^\infty\int_{\mathbb{S}^2} \chi_{r>R}r^{2a-2}(\tau_x\phi)^2\beta\, d\omega\, dr\Big|_0^T \tag{190}$$

for some pair of smooth functions $\alpha, \beta \approx 1$. This yields (187) using the condition $a < 1$. □

Proof of (188). To prove (188a) let $\chi_T(t) = \chi(t/T)$, where $\chi(s) \in C_0^\infty\big(s > \frac{1}{2}\big)$ with $\chi(1) = 1$. Then (188a) follows from integration of $\partial_u^b(\tau_-\chi_T\phi^2)$ with respect to $du\, dx$ on the cylinder $0 \leq t \leq T$ and $\frac{1}{2}R < r < R$. In the case of (188b) and (188c) we use a similar procedure but simply drop the cutoff χ_T. □

References

[Alinhac 2001] S. Alinhac, "The null condition for quasilinear wave equations in two space dimensions, I", *Invent. Math.* **145**:3 (2001), 597–618. MR Zbl

[Bieri 2010] L. Bieri, "An extension of the stability theorem of the Minkowski space in general relativity", *J. Differential Geom.* **86**:1 (2010), 17–70. MR Zbl

[Blue and Sterbenz 2006] P. Blue and J. Sterbenz, "Uniform decay of local energy and the semi-linear wave equation on Schwarzschild space", *Comm. Math. Phys.* **268**:2 (2006), 481–504. MR Zbl

[Bony and Häfner 2010] J.-F. Bony and D. Häfner, "The semilinear wave equation on asymptotically Euclidean manifolds", *Comm. Partial Differential Equations* **35**:1 (2010), 23–67. MR Zbl

[Christodoulou and Klainerman 1993] D. Christodoulou and S. Klainerman, *The global nonlinear stability of the Minkowski space*, Princeton Math. Series **41**, Princeton Univ. Press, 1993. MR Zbl

[Dafermos and Rodnianski 2009] M. Dafermos and I. Rodnianski, "The red-shift effect and radiation decay on black hole spacetimes", *Comm. Pure Appl. Math.* **62**:7 (2009), 859–919. MR Zbl

[Dafermos and Rodnianski 2010] M. Dafermos and I. Rodnianski, "A new physical-space approach to decay for the wave equation with applications to black hole spacetimes", pp. 421–432 in *XVIth International Congress on Mathematical Physics* (Prague, 2009), edited by P. Exner, World Sci. Publ., Hackensack, NJ, 2010. MR Zbl

[Dafermos et al. 2013] M. Dafermos, G. Holzegel, and I. Rodnianski, "A scattering theory construction of dynamical vacuum black holes", preprint, 2013. To appear in *J. Diff. Geom.* arXiv

[Dafermos et al. 2016] M. Dafermos, I. Rodnianski, and Y. Shlapentokh-Rothman, "Decay for solutions of the wave equation on Kerr exterior spacetimes, III: The full subextremal case $|a| < M$", *Ann. of Math.* (2) **183**:3 (2016), 787–913. MR Zbl

[Keel et al. 2002] M. Keel, H. F. Smith, and C. D. Sogge, "Almost global existence for some semilinear wave equations", *J. Anal. Math.* **87** (2002), 265–279. MR Zbl

[Klainerman 1985] S. Klainerman, "Uniform decay estimates and the Lorentz invariance of the classical wave equation", *Comm. Pure Appl. Math.* **38**:3 (1985), 321–332. MR Zbl

[Klainerman 1986] S. Klainerman, "The null condition and global existence to nonlinear wave equations", pp. 293–326 in *Nonlinear systems of partial differential equations in applied mathematics, I* (Santa Fe, NM, 1984), edited by B. Nicolaenko et al., Lectures in Appl. Math. **23**, Amer. Math. Soc., Providence, RI, 1986. MR Zbl

[Klainerman and Sideris 1996] S. Klainerman and T. C. Sideris, "On almost global existence for nonrelativistic wave equations in 3D", *Comm. Pure Appl. Math.* **49**:3 (1996), 307–321. MR Zbl

[Lindblad 2017] H. Lindblad, "On the asymptotic behavior of solutions to the Einstein vacuum equations in wave coordinates", *Comm. Math. Phys.* **353**:1 (2017), 135–184. MR Zbl

[Lindblad and Rodnianski 2010] H. Lindblad and I. Rodnianski, "The global stability of Minkowski space-time in harmonic gauge", *Ann. of Math.* (2) **171**:3 (2010), 1401–1477. MR Zbl

[Lindblad and Sterbenz 2006] H. Lindblad and J. Sterbenz, "Global stability for charged-scalar fields on Minkowski space", *Int. Math. Res. Pap.* **2006** (2006), art. id. 52976. MR Zbl

[Lindblad and Tohaneanu 2018] H. Lindblad and M. Tohaneanu, "Global existence for quasilinear wave equations close to Schwarzschild", *Comm. Partial Differential Equations* **43**:6 (2018), 893–944. MR Zbl

[Luk 2012] J. Luk, "A vector field method approach to improved decay for solutions to the wave equation on a slowly rotating Kerr black hole", *Anal. PDE* **5**:3 (2012), 553–625. MR Zbl

[Luk 2013] J. Luk, "The null condition and global existence for nonlinear wave equations on slowly rotating Kerr spacetimes", *J. Eur. Math. Soc.* **15**:5 (2013), 1629–1700. MR Zbl

[Marzuola et al. 2010] J. Marzuola, J. Metcalfe, D. Tataru, and M. Tohaneanu, "Strichartz estimates on Schwarzschild black hole backgrounds", *Comm. Math. Phys.* **293**:1 (2010), 37–83. MR Zbl

[Metcalfe and Sogge 2005] J. Metcalfe and C. D. Sogge, "Hyperbolic trapped rays and global existence of quasilinear wave equations", *Invent. Math.* **159**:1 (2005), 75–117. MR Zbl

[Metcalfe and Sogge 2007] J. Metcalfe and C. D. Sogge, "Global existence of null-form wave equations in exterior domains", *Math. Z.* **256**:3 (2007), 521–549. MR Zbl

[Metcalfe et al. 2012] J. Metcalfe, D. Tataru, and M. Tohaneanu, "Price's law on nonstationary space-times", *Adv. Math.* **230**:3 (2012), 995–1028. MR Zbl

[Metcalfe et al. 2017] J. Metcalfe, J. Sterbenz, and D. Tataru, "Local energy decay for scalar fields on time dependent non-trapping backgrounds", preprint, 2017. To appear in *Amer. J. Math.* arXiv

[Moschidis 2016] G. Moschidis, "The r^p-weighted energy method of Dafermos and Rodnianski in general asymptotically flat spacetimes and applications", *Ann. PDE* **2**:1 (2016), art. id. 6. MR Zbl

[Oliver 2013] J. Oliver, *A vector field method for non-trapping spacetimes*, Ph.D. thesis, University of California, San Diego, 2013, available at https://search.proquest.com/docview/1417765203.

[Oliver 2016] J. Oliver, "A vector field method for non-trapping, radiating spacetimes", *J. Hyperbolic Differ. Equ.* **13**:4 (2016), 735–790. MR Zbl

[Rodnianski and Sterbenz 2010] I. Rodnianski and J. Sterbenz, "On the formation of singularities in the critical O(3) σ-model", *Ann. of Math.* (2) **172**:1 (2010), 187–242. MR Zbl

[Tataru 2013] D. Tataru, "Local decay of waves on asymptotically flat stationary space-times", *Amer. J. Math.* **135**:2 (2013), 361–401. MR Zbl

[Tataru and Tohaneanu 2011] D. Tataru and M. Tohaneanu, "A local energy estimate on Kerr black hole backgrounds", *Int. Math. Res. Not.* **2011**:2 (2011), 248–292. MR Zbl

[Wang and Yu 2014] C. Wang and X. Yu, "Global existence of null-form wave equations on small asymptotically Euclidean manifolds", *J. Funct. Anal.* **266**:9 (2014), 5676–5708. MR Zbl

[Yang 2013] S. Yang, "Global solutions of nonlinear wave equations in time dependent inhomogeneous media", *Arch. Ration. Mech. Anal.* **209**:2 (2013), 683–728. MR Zbl

Received 30 May 2017. Revised 24 Sep 2018. Accepted 29 Dec 2018.

JESÚS OLIVER: jesus.oliver@csueastbay.edu
Department of Mathematics, California State University, East Bay, Hayward, CA, United States

JACOB STERBENZ: jsterben@math.ucsd.edu
Department of Mathematics, University of California at San Diego, La Jolla, CA, United States

ANALYSIS AND PDE
Vol. 13, No. 1, 2020
dx.doi.org/10.2140/apde.2020.13.93

STABLE ODE-TYPE BLOWUP FOR SOME QUASILINEAR WAVE EQUATIONS WITH DERIVATIVE-QUADRATIC NONLINEARITIES

JARED SPECK

We prove a constructive stable ODE-type blowup-result for open sets of solutions to a family of quasilinear wave equations in three spatial dimensions. The blowup is driven by a Riccati-type derivative-quadratic semilinear term, and the singularity is more severe than a shock in that the solution itself blows up like the log of the distance to the blowup-time. We assume that the quasilinear terms satisfy certain structural assumptions, which in particular ensure that the "elliptic part" of the wave operator vanishes precisely at the singular points. The initial data are compactly supported and can be small or large in L^∞ in an absolute sense, but we assume that their spatial derivatives satisfy a nonlinear smallness condition relative to the size of the time derivative. The first main idea of the proof is to construct a quasilinear integrating factor, which allows us to reformulate the wave equation as a first-order system whose solutions remain regular, all the way up to the singularity. Using the integrating factor, we construct quasilinear vector fields adapted to the nonlinear flow. The second main idea is to exploit some crucial monotonic terms in various estimates, especially the energy estimates, that feature the integrating factor. The availability of the monotonicity is tied to our size assumptions on the initial data and on the structure of the nonlinear terms. The third main idea is to propagate the relative smallness of the spatial derivatives all the way up to the singularity so that the solution behaves, in many ways, like an ODE solution. As a corollary of our main results, we show that there are quasilinear wave equations that exhibit two distinct kinds of blowup: the formation of shocks for one nontrivial set of data, and ODE-type blowup for another nontrivial set.

1. Introduction 93
2. Mathematical setup and the evolution equations 116
3. Assumptions on the initial data and bootstrap assumptions 118
4. Energy identities 121
5. A priori estimates 125
6. Local well-posedness and continuation criteria 137
7. The main theorem 138
References 142

1. Introduction

A fundamental issue surrounding the study of solutions to nonlinear hyperbolic PDEs is that singularities can form in finite time, starting from smooth initial data. For a given singularity-forming solution, perhaps

Speck gratefully acknowledges support from NSF grant # DMS-1162211, from NSF CAREER grants # DMS-1454419 and # DMS-1414537, from a Sloan Research Fellowship provided by the Alfred P. Sloan foundation, and from a Solomon Buchsbaum grant administered by the Massachusetts Institute of Technology.
MSC2010: primary 35L67; secondary 35L05, 35L45, 35L52, 35L72.
Keywords: blowup, ODE blowup, shocks, singularities, stable blowup.

the most basic question one can ask is whether or not the singularity formation is stable under perturbations of its initial data. Our main result provides a constructive, affirmative answer to this question for some solutions to a class of quasilinear wave equations. Specifically, in three spatial dimensions, we provide a sharp proof of stable ODE-type blowup for solutions corresponding to an *open* set (in a suitable Sobolev topology in which there are no radial weights in the norms) of initial data for a class of quasilinear wave equations that are well-modeled by

$$-\partial_t^2 \Phi + \frac{1}{1+\partial_t \Phi} \Delta \Phi = -(\partial_t \Phi)^2. \tag{1.1}$$

It is only for concreteness that we restrict our attention to three spatial dimensions; our approach can be applied to any number of spatial dimensions, with only slight modifications needed. See Sections 1B and 2A for a precise description of the class of equations that we treat, Section 1C1 for a summary of our main results, and Section 7 for the detailed statement of our main theorem.

There are many results on stable breakdown for solutions to wave equations, some of which we review in Section 1D. The "theory" of stable breakdown is quite fragmented in that the techniques that have been employed vary wildly between different classes of equations. In particular, the techniques that have been developed do not seem to apply to the equations that we study in this article. This will become more clear after we describe the main ideas of our proof (see Section 1C) and review prior works on stable breakdown. Although ODE-type blowup is arguably the simplest blowup scenario, there do not seem to be any prior constructive stable blowup-results of this type for scalar wave equations with derivative-quadratic nonlinearities, in any number of spatial dimensions. We mention, however, that in [Rodnianski and Speck 2018b], we proved, using rather different techniques specialized to Einstein's equations, a singularity formation result for Einstein's equations that can be interpreted as a stable ODE-type blowup-result for the first derivatives[1] of a solution to a quasilinear elliptic-hyperbolic system with derivative-quadratic nonlinearities.

Our proof of stable blowup is based on constructing a dynamic integrating factor $\mathcal{I}$, depending on Φ, that allows us to transform the problem of proving singularity formation into the problem of showing that a certain renormalized first-order system (involving Φ and $\mathcal{I}$) has regular solutions that exist for a sufficiently long time. This strategy of "renormalizing" the problem has been employed in other contexts, such as [Rodnianski and Sterbenz 2010; Raphaël and Rodnianski 2012] on stable blowup for equivariant wave maps, [Collot 2018] on the existence of singularity-forming solutions to energy-supercritical semilinear wave equations, and [Alinhac 1999a; 1999b; 2001; Christodoulou 2007; Christodoulou and Miao 2014; Speck 2016; 2019a; 2019b; 2018; Speck et al. 2016; Miao and Yu 2017; Miao 2018; Luk and Speck 2016; 2018] on stable shock formation in multiple spatial dimensions; later we will further describe these works. We stress that the issue of constructing an appropriately renormalized system lies at the heart of the difficulty of understanding the singular dynamics, and that the details behind the renormalization vary considerably between the different works. However, there is a common theme tying together many of

[1] For the solutions studied in [Rodnianski and Speck 2018b], relative to a geometrically defined coordinate system, the second fundamental form of the metric blows up, while the metric components do not; this can be viewed as the blowup of the first derivatives of the metric.

these singularity formation results (and also others that we describe in Section 1D): the idea of studying the dynamics with the help of a modulation parameter, which solves a corresponding modulation equation. Roughly, a modulation equation is an ODE-type equation that is coupled to the PDE of interest (i.e., an ODE with "PDE source terms"), and the role of the modulation parameter is to describe the "singular portion" of the PDE dynamics as being driven by the ODE-type equation (that is, the modulation equation). In the present article, the modulation parameter is $\mathcal{I}$, and it solves a transport equation (see (1C.1)) that allows us, in the renormalized system (see Proposition 2.5), to cancel the singularity-driving Riccati term $-(\partial_t \Phi)^2$ on the right-hand side of (1.1).

As in works cited in the previous paragraph, in the present article, the renormalized system that we construct becomes degenerate as Φ blows up, and for this reason, it is difficult to close the energy estimates. To achieve this, we crucially rely on some hidden monotonicity, which becomes active near the singularity and whose availability is tied to the monotonicity of the background ODE solutions (whose perturbations we study) and to our assumptions on the factor multiplying the Laplacian, e.g., $1/(1+\partial_t \Phi)$ in the model equation (1.1).

For the solutions featured in our main results, Φ itself blows up in L^∞. As of present, there is no known exhaustive classification of the kinds of singularities that can form in general solutions to quasilinear wave equations. Thus, in principle, for a different set of initial data compared to the set treated in our main theorem, other kinds of singularities could form in solutions to the equations under study. We highlight that the blowup of Φ is a much more drastic singularity compared to the formation of a shock, which for wave equations whose quasilinear terms are of the form $f(\partial\Phi)\cdot\partial^2\Phi$ (as in (1.1)), is a particular kind of singularity in which certain second-order derivatives of Φ blow up in L^∞ due to the intersection of a family of characteristic hypersurfaces, while Φ and $\partial\Phi$ remain bounded in L^∞. The blowup of Φ for the solutions under study here is philosophically important because it dashes any hope of uniquely weakly continuing the solution past the singularity, at least in a sense analogous to what might be achievable in the case of shock singularities in multiple spatial dimensions; see Remark 1.5 for further discussion on the blowup of Φ and Section 1D for further discussion on the formation of shock singularities in the context of multiple spatial dimensions and for discussion of the problem of (weakly) continuing[2] the solution past a shock.

There are a variety of singularity formation results for semilinear wave equations (whose nonlinearities satisfy appropriate assumptions) in which solutions exhibit a "universal blowup-profile" near the singularity, which roughly means that the solutions $\Phi(t,\underline{x})$ are asymptotic to $\kappa(t) f(\lambda(t)(\underline{x}-\underline{x}_0))$ as t approaches the blowup-time, where $f=f(\underline{x})$ is the universal profile, $\underline{x}_0$ is the spatial blowup-point, and κ and λ are functions that blow up as the singular time is approached; see Section 1D for a discussion of some of these results. In contrast, our proof does not rely on exhibiting a universal blowup-profile near

[2]The most significant weak continuation result in more than one spatial dimension is Christodoulou's recent solution [2019] of the restricted shock development problem in compressible fluid mechanics, which, roughly speaking, is a local well-posedness result for weak solutions and their corresponding hypersurfaces of discontinuity, starting from the first shock, whose formation from smooth initial conditions was described in detail in his breakthrough work for relativistic fluids [Christodoulou 2007] and in the follow-up work [Christodoulou and Miao 2014] on nonrelativistic compressible fluids. The term "restricted" means that the jump in entropy across the shock hypersurface was ignored. See Section 1D for further discussion.

singularities, and moreover, *we expect that there should not be any such universal blow-up profile for the set of singularity-forming solutions exhibited by our main theorem* (although we do not prove this). Let us briefly explain the reasons behind our expectation. Our proof is based on constructing the dynamic function $\mathcal{I} > 0$ mentioned above, whose evolution equation is coupled to Φ, such that $\mathcal{I}$ and $\mathcal{I}\partial_t\Phi$ remain rather smooth all the way up to the singularity. To prove that a singularity develops, we show that $\mathcal{I}$ vanishes in finite time in a region where $\mathcal{I}\partial_t\Phi$ is strictly positive; this shows that $\partial_t\Phi$ blows up "like $F(\underline{x})/\mathcal{I}(t,\underline{x})$", where the regular function F is positive in regions where $\mathcal{I}$ is small, while the blowup of Φ itself then follows from simple arguments given in Remark 1.5. Our proof allows for the possibility that at the time of first blowup, $\mathcal{I}$ and $\mathcal{I}\partial_t\Phi$ (where $\mathcal{I}\partial_t\Phi = F$ along the constant-time hypersurface of first blowup) could be arbitrary smooth functions of the spatial variables, aside from the facts that $\partial_t\mathcal{I} = -\mathcal{I}\partial_t\Phi$ (see (1C.1)), that $\mathcal{I}\partial_t\Phi$ is nonzero at points where $\mathcal{I}$ vanishes, and that the spatial derivatives of $\mathcal{I}$ and $\mathcal{I}\partial_t\Phi$ are small (the latter two features are consequences of our assumptions on the initial data, which we state in Section 3A); it is because of this *arbitrariness* that a universal blowup-profile is not featured in our proofs. See the last point of Theorem 1.2 and Remark 1.4 for further discussion of these points. We also remark that for related reasons, the proofs of the shock formation results described in Section 1D do not make any reference to universal blowup-profiles; see the next two paragraphs for further discussion.

We expect that for the solutions treated by our main theorem, generically, the constant-time hypersurface of first blowup will feature only isolated singularities; i.e., we expect that generic solutions have the property that the zeros of $\mathcal{I}$ within the constant-time hypersurface of first blowup are isolated.[3] For such isolated singularities, it should be possible to provide a sharper description of the blowup (compared to the description outlined in the previous paragraph) in a neighborhood of each singularity, essentially through Taylor expansions of the smooth functions $\mathcal{I}$ and $\mathcal{I}\partial_t\Phi$ at each blowup-point. Although (for brevity) we do not pursue this issue in the present article, deriving such a sharp description would be an important precursor step to studying the boundary of the maximal development of the initial data; see also Remark 1.3. This style of analysis (involving, in particular, Taylor expansions) was carried out in [Christodoulou 2007, Chapter 15] in three spatial dimensions in his sharp proof of shock formation for the wave equations of irrotational relativistic fluid mechanics, which are quasilinear. In particular, Christodoulou's work yielded a sharp description of a portion of the boundary of the maximal development in a neighborhood of nondegenerate shock singularities, which is a class of shock singularities that includes many that are isolated in the constant-time hypersurface of first blowup; see [Christodoulou 2007, Chapter 15] for the detailed nondegeneracy assumptions made in his study of the boundary of the maximal development. Christodoulou's results substantially extended those of Alinhac [1999a; 1999b; 2001], who, for various quasilinear wave equations in multiple spatial dimensions, gave a precise description of the formation of nondegenerate shock singularities, but without uncovering the structure of the boundary of the maximal development; see Section 1D for further discussion.

The arguments given in [Christodoulou 2007, Chapter 15] suggest (although they do not definitively prove) that there should not be any universal blowup-profile that captures the asymptotics of solutions

[3]We believe this because $\mathcal{I}$, when viewed as a function of the spatial variables at fixed first blowup time, has a minimum wherever it vanishes, and we believe that these minima should generically be nondegenerate critical points.

to the wave equations that he studied near shock singularities, even singularities that are isolated in the constant-time hypersurface of first blowup. The reason is that, much like in the present work, the blowup of the solution was shown in [Christodoulou 2007] to occur precisely along the zero level set of a smooth "dynamically constructed" function,[4] somewhat analogous to $\mathcal{I}$, and the arguments of [Christodoulou 2007, Chapter 15] allowed for the possibility that the function is arbitrary (aside from satisfying Christodoulou's nondegeneracy assumptions, which one might expect to generically hold). It is interesting to contrast this freedom against the rigidity that occurs in the study of blowup for semilinear heat equations with focusing power-law nonlinearities in multiple spatial dimensions, i.e., $\partial_t \Phi = \Delta \Phi + |\Phi|^{p-1}\Phi$ with $p > 1$: there is a classical result [Merle and Zaag 1997] showing that, under suitable assumptions on p, there exists an open set of initial data whose corresponding solutions form an isolated singularity in finite time and exhibit a universal blowup-profile. This work can be viewed as an extension of [Bricmont and Kupiainen 1994], in which the authors constructed (without proving stability) an infinite number of families of solutions containing isolated singularities within the constant-time hypersurface of first blowup such that the solutions exhibit universal blowup-profiles near the singularities, where each family corresponds to a distinct profile.

As a corollary of our main results, we show (see Section 1E) that there are quasilinear wave equations that exhibit *two distinct kinds of blowup*: ODE-type blowup for one nontrivial (but not necessarily open) set of initial data, and the formation of a shock for a different nontrivial set of data. We view this as a parable highlighting two key phenomena that would have to be accounted for in any sufficiently broad theory of singularity formation in solutions to quasilinear wave equations; i.e., in principle, a quasilinear wave equation can admit radically different types of singularity-forming solutions. The phenomenon of distinct types of singularities is well known for certain *semilinear* equations, but this issue has not been substantially explored for quasilinear equations. For example, for the nonlinear Schrödinger equation with a suitable L^2-supercritical semilinear term, there is a stable regime of self-similar blowup [Merle et al. 2010], while other radially symmetric solutions exhibit collapsing ring singularities [Merle et al. 2014]. In Section 1D, we describe some known results for semilinear wave equations exhibiting distinct types of singularity-forming solutions, but we mention already that there can be type-I blowup, in which the L^∞ norm of the solution itself blows up, as well as type-II blowup, in which the solution remains bounded in an appropriate Sobolev norm. We also highlight that in the quasilinear case, the phenomenon of distinct singularity types can be exhibited in the much simpler setting of quasilinear transport equations. For example, the inhomogeneous Burgers equation $\partial_t \Psi + \Psi \partial_x \Psi = \Psi^2$ admits the T-parametrized family of spatially homogeneous singularity-forming solutions $\Psi_{(\mathrm{ODE});T} := (T-t)^{-1}$ as well as solutions that form shocks, i.e., $\partial_x \Psi$ blows up but Ψ remains bounded (at least up to the singularity in $\partial_x \Psi$).

The precise algebraic details of the weight $1/(1+\partial_t \Phi)$ in front of the Laplacian term in (1.1) are not important for our proof. What is important is that the weight decays at an appropriate rate as $\partial_t \Phi \to \infty$, that is, as the singularity forms; see Section 2A for our assumptions on the weight. As we will explain,

[4]The function appearing in [Christodoulou 2007], denoted by μ, measures the inverse density of a family of characteristic hypersurfaces. It is a three-space-dimensional analog of the function, also denoted by μ, that we use in Section 1E to study shock formation in one spatial dimension.

this decay yields a friction-type spacetime integral that is important for closing the energy estimates, and it also helps us to prove that spatial derivative terms remain small relative to the time derivative terms, up to the singularity. The problem of providing a sharp description of blowup for solutions to derivative-quadratic semilinear wave equations, such as $-\partial_t^2\Phi + \Delta\Phi = -(\partial_t\Phi)^2$, remains open, even though [John 1981] showed, via proof by contradiction, that all nontrivial, smooth, compactly supported solutions to the equation $-\partial_t^2\Phi + \Delta\Phi = -(\partial_t\Phi)^2$ in three spatial dimensions blow up in finite time.

Our results show, in part due to the weight in front of the Laplacian, that the spatial-derivative-involving nonlinearities in (1.1) (and the other equations that we study) exhibit a subcritical[5] blowup-rate relative to the pure time derivative terms. However, as we explain below, this *subcritical behavior does not seem detectable relative to the standard partial derivatives* ∂_α; to detect the behavior, we will use a combination of "quasilinear vector field derivatives" $\mathcal{I}\partial_\alpha$ and standard derivatives ∂_α, where $\mathcal{I}$ is the "quasilinear integrating factor" mentioned above. As we have already alluded to above, our proof is based on showing that $\mathcal{I}\partial_\alpha\Phi$ remains bounded up to the singularity and that the singularity formation coincides with the vanishing of $\mathcal{I}$. In total, our approach allows us to treat the equations under study as quasilinear perturbations of the Riccati ODE $\frac{d^2}{dt^2}\Phi = \left(\frac{d}{dt}\Phi\right)^2$. By "perturbation of the Riccati ODE", we mean in particular that the singularity formation is similar to the one that occurs in the following T-parametrized family of ODE solutions to (1.1):

$$\Phi_{(\mathrm{ODE});T}(t) := \ln((T-t)^{-1}), \tag{1.2}$$

where $T \in \mathbb{R}$ is the blowup-time. Our methods are tailored to the quadratic term on the right-hand side of (1.1) in that they do not apply, at least in their current form, to semilinear terms of type $(\partial_t\Phi)^p$ for $p \neq 2$. However, derivative-quadratic terms are of particular interest in view of the fact that they commonly arise in nonlinear field theories (though the derivative-quadratic terms in such theories are often not Riccati-type, like the one featured on the right-hand side of (1.1)).

1A. ***Paper outline.*** • In the remainder of Section 1, we summarize our results, outline their proofs, place our work in context by discussing prior works on the breakdown of solutions, discuss a corollary (see Section 1E) of our main results, and summarize our notation.

• In Section 2, we define the quantities that play a role in our analysis and derive various evolution equations.

• In Section 3, we state our assumptions on the initial data and state bootstrap assumptions that are useful for studying the solution.

• In Section 4, we derive energy identities.

• In Section 5, which is the main section of the article, we derive a priori estimates that in particular yield strict improvements of the bootstrap assumptions.

[5]In contrast, for the semilinear equation $-\partial_t^2\Phi + \Delta\Phi = -(\partial_t\Phi)^2$, our approach suggests, but does not prove, that the blowup-rate for the Laplacian term $\Delta\Phi$ might be critical with respect to the expected blowup-rate for the other two terms in the equation, i.e., that all terms might blow up at the same rate.

- In Section 6, we state a standard local well-posedness result and continuation criteria for the equations under study.
- In Section 7, we prove the main theorem.

1B. ***The class of wave equations under study.*** Our main theorem concerns the following Cauchy problem for a quasilinear wave equation in three spatial dimensions:

$$-\partial_t^2\Phi + \mathscr{W}(\partial_t\Phi)\Delta\Phi = -(\partial_t\Phi)^2, \tag{1B.1a}$$

$$(\partial_t\Phi|_{\Sigma_0}, \partial_1\Phi|_{\Sigma_0}, \partial_2\Phi|_{\Sigma_0}, \partial_3\Phi|_{\Sigma_0}) = (\mathring{\Psi}_0, \mathring{\Psi}_1, \mathring{\Psi}_2, \mathring{\Psi}_3), \tag{1B.1b}$$

where throughout, Σ_t denotes the hypersurface of constant time t. We assume that $\partial_i\mathring{\Psi}_j = \partial_j\mathring{\Psi}_i$ for $i, j = 1, 2, 3$, which by Poincaré's lemma is equivalent to the existence of a function $\mathring{\Phi}$ on $\mathbb{R}^3$ such that $\mathring{\Psi}_i = \partial_i\mathring{\Phi}$ for $i = 1, 2, 3$; see Remark 1.1. Our use of the notation "$\mathring{\Psi}_\alpha$" for the data functions is tied to our use of the "renormalized solutions variables" Ψ_α that we will use in studying solutions; see Definition 2.3.

Remark 1.1 (viewing (1B.1a) as an equation in $\partial\Phi$). Since Φ itself is not featured in (1B.1a) (only its derivatives appear), we only need to prescribe the first derivatives of Φ along Σ_0 (subject to the constraint $\partial_i\mathring{\Psi}_j = \partial_j\mathring{\Psi}_i$ for $i, j = 1, 2, 3$ mentioned above) in order to solve for $\{\partial_\alpha\Phi\}_{\alpha=0,1,2,3}$. This is relevant in that we do not bother to derive estimates for Φ itself (see, however, Remark 1.5).

In (1B.1a), $\Delta := \sum_{a=1}^3 \partial_a^2$ is the standard Euclidean Laplacian on $\mathbb{R}^3$ and $\mathscr{W} = \mathscr{W}(\partial_t\Phi)$ is a nonlinear "weight function" satisfying certain technical conditions stated below, specifically (2A.1)–(2A.5). Prototypical examples of weights satisfying (2A.1)–(2A.5) are the functions

$$\mathscr{W}(y) = \frac{1}{1+y^M} \quad \text{or} \quad \mathscr{W}(y) = \frac{1}{(1+y)^M}, \tag{1B.2}$$

where $M \geq 1$ is an integer, and the function

$$\mathscr{W}(y) = \exp(-y). \tag{1B.3}$$

1C. ***Rough summary of the results and discussion of the proof.***

1C1. *Rough summary of the results.* We now briefly summarize the main results; see Theorem 7.1 for precise statements.

Theorem 1.2 (stable ODE-type blowup, rough version). *Under suitable assumptions (stated in Section 2A) on the weight $\mathscr{W}(\partial_t\Phi)$, there exists an open set of compactly supported initial data $\{\mathring{\Psi}_\alpha\}_{\alpha=0,1,2,3}$ (see (1B.1b)) for (1B.1a), with $\mathring{\Psi}_\alpha \in H^5(\mathbb{R}^3)$, such that the solution blows up in finite time in a manner similar to the ODE solutions $\Phi_{(\text{ODE});T}$ from (1.2). In particular, there exists a time $0 < T_{(\text{Lifespan})} < \infty$ such that $\|\partial_t\Phi\|_{L^\infty(\Sigma_t)}$ and[6] $\|\Phi\|_{L^\infty(\Sigma_t)}$ blow up as $t \uparrow T_{(\text{Lifespan})}$. The data functions $\{\mathring{\Psi}_\alpha\}_{\alpha=0,1,2,3}$ are allowed to be large or small as measured by a Sobolev norm without radial weights, but $\{\mathring{\Psi}_a\}_{a=1,2,3}$, $\nabla\mathring{\Psi}_0$, and their spatial derivatives up to top order must satisfy a nonlinear smallness condition relative to*[7] $\max_{\Sigma_0}[\mathring{\Psi}_0]_+$.

[6] More precisely, one can conclude that $\|\Phi\|_{L^\infty(\Sigma_t)}$ blows up at $t = T_{(\text{Lifespan})}$ if the initial datum for Φ itself is prescribed; Remark 1.1.

[7] Here and throughout, $[f]_+ := \max\{f, 0\}$.

Moreover, let the integrating factor $\mathcal{I}$ *be the solution to the initial value problem*

$$\partial_t \mathcal{I} = -\mathcal{I}\partial_t \Phi, \quad \mathcal{I}|_{\Sigma_0} = 1. \tag{1C.1}$$

Then $\mathcal{I}$, *the variables*

$$\Psi_\alpha := \mathcal{I}\partial_\alpha \Phi, \tag{1C.2}$$

and their partial derivatives with respect to the Cartesian coordinates remain regular all the way up to time $T_{(\text{Lifespan})}$, *except possibly at the top derivative level due to the vanishing of* $\mathscr{W}(\partial_t\Phi)$ (*which appears as a weight in the energies*) *as* $\partial_t \Phi \uparrow \infty$. *Moreover*,

Along $\Sigma_{T_{(\text{Lifespan})}}$, *the set of points at which a singularity forms is exactly*

$$\Sigma^{\text{Blowup}}_{T_{(\text{Lifespan})}} := \{(T_{(\text{Lifespan})}, \underline{x}) \mid \mathcal{I}(T_{(\text{Lifespan})}, \underline{x}) = 0\}.$$

More precisely, *there exists a data-dependent function*[8] F *on* $\mathbb{R}^3$ *such that*, *for any real number* $N < 5$, *we have* $F \in H^N(\mathbb{R}^3)$, *such that* $\lim_{t\uparrow T_{(\text{Lifespan})}} \|[\mathcal{I}\partial_t\Phi](t,\cdot) - F\|_{H^N(\mathbb{R}^3)} = 0$, *and such that*

$$\min_{\Sigma^{\text{Blowup}}_{T_{(\text{Lifespan})}}} F > 0.$$

In particular, *in view of* (1C.1), *we see that*, *for spatial points* $\underline{x}$ *with* $(T_{(\text{Lifespan})}, \underline{x}) \in \Sigma^{\text{Blowup}}_{T_{(\text{Lifespan})}}$, $\mathcal{I}(t,\underline{x})$ *vanishes* ***linearly***[9] *as* $t \uparrow T_{(\text{Lifespan})}$, *and* $\partial_t\Phi(t,\underline{x})$ *blows up like* $F(\underline{x})/\mathcal{I}(t,\underline{x})$ *as* $t \uparrow T_{(\text{Lifespan})}$.

Remark 1.3 (maximal development). We anticipate that the sharp results of Theorem 1.2 should be useful for obtaining detailed information about the solution not just up to the first singular time, but also up to the boundary of the maximal development.[10] In the context of shock formation for irrotational solutions to the compressible Euler equations, [Christodoulou 2007, Chapter 15] used similar sharp estimates to follow the solution up to boundary. Broadly similar results were obtained in [Merle and Zaag 2012], in which, in the case of one spatial dimension, they gave a sharp description of the boundary of the maximal development for *any* singularity-forming solution to the semilinear focusing wave equation $-\partial_t^2\Psi + \partial_x^2\Psi = -|\Psi|^{p-1}\Psi$ with $p > 1$ and showed in particular that characteristic points on the boundary are isolated. In related work [Merle and Zaag 2015], in $n \geq 1$ spatial dimensions, the authors studied the focusing wave equation $-\partial_t^2\Psi + \Delta\Psi = -|\Psi|^{p-1}\Psi$ with data $(\Psi|_{t=0}, \partial_t\Psi|_{t=0}) \in H^1 \times L^2$ in the subconformal case, i.e., the case $1 < p$ for $n = 1$ and $1 < p < (n+3)/(n-1)$ for $n \geq 2$. The authors showed that a subset[11] of the noncharacteristic portion of the blowup-surface is open, C^1, and stable under perturbations of the data.

[8] In Theorem 7.1, we did not explicitly mention the function F, but the existence of F and its properties follow easily from the results stated in the theorem, in particular from (7.6a) and (7.7).

[9] That is, for $(T_{(\text{Lifespan})}, \underline{x}) \in \Sigma^{\text{Blowup}}_{T_{(\text{Lifespan})}}$, we have $-\infty < \partial_t\mathcal{I}(T_{(\text{Lifespan})}, \underline{x}) < 0$.

[10] The maximal development of the data is, roughly, the largest possible classical solution that is uniquely determined by the data. Readers can consult [Sbierski 2016; Wong 2013] for further discussion.

[11] The subset is the collection of noncharacteristic points on the boundary of the maximal development such that near that point, the singularity-forming solution has the asymptotic profile of a Lorentz transform of a member of a certain family of equilibria.

Remark 1.4 (remarks on the set of points where blowup occurs). Theorem 1.2 shows in particular that at time $T_{(\text{Lifespan})}$, the singularity occurs precisely along the zero level set of the smooth[12] function $\mathcal{I}(T_{(\text{Lifespan})}, \cdot)$ on $\mathbb{R}^3$. Since any closed subset of $\mathbb{R}^3$ can be realized as the 0 level set of a smooth function, Theorem 1.2 allows for the possibility that any compact[13] subset of $\Sigma_{T_{(\text{Lifespan})}}$ could in principle be the set of points at which some initially smooth solution first blows up. We conjecture that this is in fact the case, that is, that, for any time $T_{(\text{Lifespan})} \in \mathbb{R}$ and any compact subset $\mathfrak{K}$ of $\Sigma_{T_{(\text{Lifespan})}}$, there is a solution Ψ to (1B.1a) that is smooth in a slab of the form $[T_{(\text{Lifespan})} - \delta, T_{(\text{Lifespan})}) \times \mathbb{R}^3$ (for some $\delta > 0$) and that blows up along $\{T_{(\text{Lifespan})}\} \times \mathfrak{K}$ (which is a subset of $\Sigma_{T_{(\text{Lifespan})}}$). One might approach proving this conjecture by adopting a "backwards approach", that is, by prescribing smooth initial data (more precisely, initial data belonging to a sufficiently high-order Sobolev space) for $\mathcal{I}$ and Ψ_α along $\Sigma_{T_{(\text{Lifespan})}}$ such that $\mathfrak{K} = \{\underline{x} \in \mathbb{R}^3 \mid \mathcal{I}(T_{(\text{Lifespan})}, \underline{x}) = 0\}$ and[14] $\Psi_0(T_{(\text{Lifespan})}, \underline{x}) > 0$ for $\underline{x} \in \mathfrak{K}$, and then trying to solve (2B.1) and the equations of Proposition 2.5 backwards, that is, for $t < T_{(\text{Lifespan})}$. Such an approach has successfully been employed in related contexts, for example in the construction of singular solutions to the Einstein equations of general relativity [Beyer and LeFloch 2010]. We also recall the aforementioned works [Bricmont and Kupiainen 1994; Merle and Zaag 1997] on focusing semilinear heat equations, in which families of singular solutions (with, however, singularities that are *isolated*) exhibiting prescribed asymptotic behavior were constructed.

Remark 1.5 (the blowup of Φ). We now make some remarks on the blowup of Φ itself since, as we highlighted in Remark 1.1, one does not need to prescribe the initial datum of Φ itself (and since in the rest of the paper we do not assume that initial data for Φ itself are prescribed). If one does prescribe its initial data, then the results of Theorem 7.1 can easily be used to show that Φ itself blows up at time $T_{(\text{Lifespan})}$ (such a result is not stated in Theorem 7.1). To deduce the blowup for Φ, one can first use (1C.1) and the fundamental theorem of calculus to deduce that $\ln \mathcal{I}(t, \underline{x}) + \Phi(t, \underline{x}) = \Phi(0, \underline{x})$, where $\Phi(0, \cdot)$ is a regular function that by assumption satisfies $\|\Phi(0, \cdot)\|_{L^\infty} < \infty$. Since the singularity formation for $\partial_t \Phi$ yielded by Theorem 7.1 coincides with the vanishing of $\mathcal{I}$ for the first time at $t = T_{(\text{Lifespan})}$, it follows that $\lim_{t \uparrow T_{(\text{Lifespan})}} \sup_{s \in [0,t)} \|\Phi\|_{L^\infty(\Sigma_s)} = \infty$, as is claimed in Theorem 1.2.

1C2. *The main ideas behind the proof of Theorem 7.1.* The initial data that we consider are such that the spatial derivatives of Φ up to top order are initially small relative to $\partial_t \Phi$. We also assume that the spatial derivatives of $\partial_t \Phi$ up to top order are initially small. The smallness assumptions that we need to close the proof are nonlinear in nature,[15] for reasons described just below (1C.5); see Section 3C for our precise smallness assumptions and Section 3D for a simple proof that such data exist. In our analysis, we propagate certain aspects of this smallness all the way up to the singularity. As we mentioned earlier, this

[12] More precisely, in view of our assumptions on the initial data, we have $\mathcal{I}(T_{(\text{Lifespan})}, \cdot) - 1 \in H^5(\mathbb{R}^3)$; see (7.6c).

[13] Since we are assuming that the initial data for $\partial_\alpha \Phi$ are compact, it is easy to show that the zero level set of $\mathcal{I}(T_{(\text{Lifespan})}, \cdot)$ is also compact; by virtue of finite speed of propagation for the wave equation satisfied by Ψ, one can show that there exists a compact subset K of $\mathbb{R}^3$ such that if $t \in [0, T_{(\text{Lifespan})}]$ and $\underline{x} \notin K$, then $\mathcal{I}(t, \underline{x}) = 1$.

[14] Note that by (1C.1)–(1C.2), this latter condition implies that $\partial_t \mathcal{I}(T_{(\text{Lifespan})}, \underline{x}) < 0$ for $\underline{x} \in \mathfrak{K}$.

[15] In particular, our smallness assumptions on the data (1B.1b) are *not* generally invariant under rescalings of the form $(\mathring{\Psi}_0, \mathring{\Psi}_1, \mathring{\Psi}_2, \mathring{\Psi}_3) \to \lambda^{-1}(\mathring{\Psi}_0, \mathring{\Psi}_1, \mathring{\Psi}_2, \mathring{\Psi}_3)$ if λ is too large.

allows us to effectively treat (1B.1a) as a perturbation of the Riccati ODE $\frac{d^2}{dt^2}\Phi = \left(\frac{d}{dt}\Phi\right)^2$. We again stress that the vanishing of the coefficient $\mathscr{W}(\partial_t\Phi)$ of the Laplacian term in (1B.1a) as the singularity forms is important for our estimates, in particular for showing that spatial derivative terms remain relatively small.

A key point is that it does not seem possible to follow the solution all the way to the singularity by studying the wave equation in the form (1B.1a). To caricature the situation, let us pretend that the singularity occurs at $t = 1$. Our proof shows, roughly, that, for $k \geq 1$, $\|\partial^k\Phi\|_{L^\infty(\Sigma_t)}$ blows up like $c_k(1-t)^{-k}$, where c_k is a data-dependent constant and ∂^k denotes k-th-order Cartesian coordinate partial derivatives. This means, in particular, that commuting (1B.1a) with more and more spatial derivatives makes the singularity strength of the nonlinear terms worse and worse, which is a serious obstacle to closing nonlinear estimates. For this reason, as our statement of Theorem 1.2 already makes clear, our proof is fundamentally based on the solution to (1C.1), that is, the integrating factor $\mathcal{I}$ solving the transport equation $\partial_t\mathcal{I} = -\mathcal{I}\partial_t\Phi$ with initial conditions $\mathcal{I}|_{\Sigma_0} = 1$. We again stress that the finite-time blowup $\partial_t\Phi \uparrow \infty$ would follow from the finite-time vanishing of $\mathcal{I}$ and thus, to prove that a singularity forms, *we will show that $\mathcal{I}$ vanishes in finite time*. Using $\mathcal{I}$, we are able to transform the wave equation into a "renormalized system" that is equivalent to (1B.1a) up to the singularity. We then analyze the renormalized system and show that the weighted derivatives $\{\Psi_\alpha := \mathcal{I}\partial_\alpha\Phi\}_{\alpha=0,1,2,3}$, $\mathcal{I}$, and the *Cartesian* spatial partial derivatives of these quantities remain bounded, in appropriate norms (some with $\mathcal{I}$ weights), all the way up to the singularity. In particular, our proof relies on a combination of the derivatives $\{\mathcal{I}\partial_\alpha\}_{\alpha=0,1,2,3}$ and $\{\partial_\alpha\}_{\alpha=0,1,2,3}$, where the weighted derivatives $\mathcal{I}\partial_\alpha$ act first. Here and throughout, $\partial_0 = \partial_t$ and $\{\partial_i\}_{i=1,2,3}$ are the standard Cartesian coordinate spatial partial derivatives.

In Proposition 2.5, we derive the renormalized system of equations satisfied by $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$. Here we only note that the system is first-order hyperbolic and that a *seemingly dangerous factor of $\mathcal{I}^{-1}$ appears in the equations* (recall that $\mathcal{I}$ vanishes at the singularity). However, the factor $\mathcal{I}^{-1}$ is multiplied by the weight $\mathscr{W} = \mathscr{W}(\mathcal{I}^{-1}\Psi_0)$ from (1B.1a), and due to our assumptions on $\mathscr{W}$, we are able to show that the product $\mathcal{I}^{-1}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)$ (which is equal to $\mathcal{I}^{-1}\mathscr{W}(\partial_t\Phi)$) remains uniformly bounded up to the singularity. Moreover, the spatial derivatives of the product $\mathcal{I}^{-1}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)$ also are controllable up to the singularity; it is in this sense that the equations satisfied by $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ can be viewed as a "renormalization" (i.e., a "regularization") of the original problem. The proof (see Lemma 5.6) of these bounds for the product $\mathcal{I}^{-1}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)$ constitutes the most technical analysis of the article and is based on separately treating regions where $\mathcal{I}$ is large and $\mathcal{I}$ is small.

To prove that $\partial_t\Phi$ blows up, we derive, in an appropriate *localized* region of spacetime, a pointwise bound for Ψ_0 of the form $\Psi_0 \gtrsim 1$. In view of the evolution equation (1C.1) for $\mathcal{I}$, we see that such a bound is strong enough to drive $\mathcal{I}$ to 0 in finite time (note that the right-hand side of the evolution equation in (1C.1) can be expressed as $-\Psi_0$). To prove that $\Psi_0 \gtrsim 1$, we of course rely on the size assumptions described in the first paragraph of this subsubsection, which in particular include the assumption that $\Psi_0|_{\Sigma_0} \gtrsim 1$ (in a localized region). If we caricature the situation by assuming the estimate[16] $\Psi_0 \sim \delta$ for some $\delta > 0$, then it follows from the evolution equation for $\mathcal{I}$ that $\mathcal{I} \sim 1-\delta t$, $\partial_t\Phi \sim (1-\delta t)^{-1}$, $\ln\mathcal{I} + \Phi \sim \text{data}$, and thus $\Phi \sim \ln(1-\delta t)^{-1} + \text{data}$, where data is a smooth function determined by the

[16]Here we use the notation "$A \sim B$" to imprecisely indicate that A is well-approximated by B.

initial data. Note that $\ln(1-\delta t)^{-1}$ is one of the ODE blowup solutions (1.2). It is in this sense that our results yield the stability of ODE-type blowup.

In reality, to close the proof sketch described above, we must overcome several major difficulties. The first is that the blowup-time is not known in advance. However, we are able to make a good approximate guess for it, which is sufficient for closing a bootstrap argument. We now describe what we mean by this. The discussion in the previous paragraph suggests that the (future) blowup-time is approximately $1/\mathring{A}_*$, where $\mathring{A}_* := \max_{\Sigma_0}[\mathring{\Psi}_0]_+$ (where $\mathring{A}_* > 0$ by assumption). Indeed, if we set all spatial derivative terms equal to zero in (1B.1a), then the blowup-time is precisely $1/\mathring{A}_*$. Our main theorem confirms that, for data with small spatial derivatives, the blowup-time is a small perturbation of $1/\mathring{A}_*$. This is conceptually important in that it enables us to use a bootstrap argument in which we only aim to control the solution for times less than $2/\mathring{A}_*$; the factor of 2 gives us a sufficient margin of error to show that the singularity does form, and it allows us, in most cases, to soak factors of $1/\mathring{A}_*$ into the constants "C" in our estimates; see Section 5A for further discussion on our conventions for constants.

The second and main difficulty that we encounter in our proof is that we need to derive energy estimates for $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ that hold up to the singularity and, at the same time, to control the integrating factor $\mathcal{I}$; most of our work in this paper is towards this goal. Our energies are roughly of the following form, where $V = (V_0, V_1, V_2, V_3)$ should be thought of as some k-th Cartesian spatial derivative of $(\Psi_0, \Psi_1, \Psi_2, \Psi_3)$:

$$\mathbb{E}[V] = \mathbb{E}[V](t) := \int_{\Sigma_t} \left\{ V_0^2 + \sum_{a=1}^{3} \mathscr{W}(\mathcal{I}^{-1}\Psi_0) V_a^2 \right\} d\underline{x}. \tag{1C.3}$$

For the data under study, $\mathbb{E}[V](0)$ is small. Since $\mathcal{I}$ is small near the singularity and Ψ_0 is order-unity, our assumptions on $\mathscr{W}$ imply that the factor $\mathscr{W}(\mathcal{I}^{-1}\Psi_0)$ on the right-hand side of (1C.3) is small near the singularity; i.e., the energy provides only weak control over $\{V_a\}_{a=1,2,3}$. This makes it difficult to control certain terms in the energy identities, which arise from commutator error terms (that are generated upon commuting the evolution equations for $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ with spatial derivatives) and from the basic integration by parts argument that we use to derive the energy identities. To control the most difficult error integrals, we exploit the following spacetime integral, which also appears in the energy identities (roughly it is generated when ∂_t falls on the weight $\mathscr{W}(\mathcal{I}^{-1}\Psi_0)$ on the right-hand side of (1C.3)):

$$\mathfrak{I}[V](t) := \sum_{a=1}^{3} \int_{s=0}^{t} \int_{\Sigma_s} (\mathcal{I}^{-1}\Psi_0)^2 \mathscr{W}'(\mathcal{I}^{-1}\Psi_0)(V_a)^2 \, d\underline{x}\, ds, \tag{1C.4}$$

where $\mathscr{W}'(y) = \frac{d}{dy}\mathscr{W}(y)$. A good model scenario to keep in mind is the case $\mathscr{W} = 1/(1+\partial_t\Phi)$ in regions where $\partial_t\Phi$ is large (and thus the energy (1C.3) is weak), in which case $\mathscr{W}' = -1/(1+\partial_t\Phi)^2$, and the factor $(\mathcal{I}^{-1}\Psi_0)^2\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)$ on the right-hand side of (1C.4) can be expressed as $-(\partial_t\Phi)^2/(1+\partial_t\Phi)^2$. In view of our assumptions on $\mathscr{W}$, the term $\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)$ has a *quantitatively negative* sign in the difficult regions where $\mathcal{I}$ is small (which is equivalent to the largeness of $\partial_t\Phi$). More precisely, (1C.4) has a *friction-type* sign. This is important because the difficult error integrals mentioned above can be bounded by $\lesssim \varepsilon\mathfrak{I}[V](t)$, where, roughly, ε is the small L^∞ size of the solution's spatial derivatives. For this reason, the integral (1C.4) can be used to absorb the difficult error integrals. In total, this allows us to prove a priori energy estimates, roughly of the form

$$\mathbb{E}[V](t) + \mathfrak{I}[V](t) \leq \text{data} \times C \exp(Ct), \tag{1C.5}$$

where "data" is, roughly, the small size of the spatial derivatives of the initial data. For our proof to close, the right-hand side of (1C.5) must be sufficiently small. Thanks to our bootstrap assumption that $t < 2/\mathring{A}_*$, it suffices to choose that initial data so that data $\times\, C \exp(C/\mathring{A}_*)$ is sufficiently small. This is one example of the *nonlinear smallness* of the spatial derivatives, relative to $\mathring{A}_*$, that we impose to close the proof. In reality, to make this procedure work, we must separately treat regions where $\mathcal{I}$ is small and $\mathcal{I}$ is large; see Proposition 5.8 and its proof for the details. We stress that absorbing the difficult error integrals into the friction integral (1C.4) *is crucial for showing that the energies remain bounded up to the vanishing of* $\mathcal{I}$, which is in turn central for our approach. In the model case $\mathscr{W}(\partial_t \Phi) = 1/(1 + \partial_t \Phi)$, if we had instead tried to directly control the difficult error integrals by the energy, then we would have obtained the inequality $\mathbb{E}[V](t) \leq C \int_{s=0}^{t} \|\partial_t \Phi\|_{L^\infty(\Sigma_s)} \mathbb{E}[V](s)\, ds + \cdots$. Since $\|\partial_t \Phi\|_{L^\infty(\Sigma_s)}$ goes to infinity at a nonintegrable rate[17] as the blowup-time is approached, this would have led to a priori energy estimates allowing for the possibility that the energies blow up at the singularity, which would have completely invalidated our philosophy of obtaining nonsingular estimates for the $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$. We also highlight that the regularity theory of $\mathcal{I}$ is somewhat subtle at top order: our proof requires that we show that $\mathcal{I}$ and Φ have the same degree of differentiability and that the estimates for $\mathcal{I}$ do not involve any dangerous factors of $\mathcal{I}^{-1}$; these features are not immediately apparent from (1C.1).

The circle of ideas tied to the "regularization approach" that we have taken here seems to be new in the context of proving the stability of ODE-type blowup for a quasilinear wave equation. However, our approach has some parallels with the known proofs of stable shock formation in multiple spatial dimensions, which we describe in Section 1D. In those problems, the crux of the proofs also involves quasilinear integrating factors that "hide" the singularity. In shock formation problems, the integrating factor (which, earlier in the introduction, we referred to as a "modulation parameter") is tied to nonlinear geometric optics,[18] and its top-order regularity theory is very difficult, at least in multiple spatial dimensions (much more so than the top-order regularity theory of the integrating factor $\mathcal{I}$ employed in the present article). The proofs of shock formation also crucially rely on friction-type spacetime integrals, in analogy with (1C.4), that are available because the integrating factor has a negative derivative (in an appropriate direction) in regions near the singularity. However, in multidimensional shock formation problems, the top-order energy identities feature dangerous terms, analogous to terms of strength $\mathcal{I}^{-1}$, which lead to a priori energy estimates allowing for the possibility that the high-order energies might blow up like $\mathcal{I}^{-P}$ for some large universal constant P. This makes it difficult to derive the nonsingular estimates at the lower derivative levels, which are central for closing the proof. In contrast, in our work here, the difficult factors of $\mathcal{I}^{-1}$ are always multiplied by the term $\mathscr{W}$, which effectively ameliorates them, making it easier

[17] With a bit of additional effort, Theorem 7.1 could be sharpened to show that $\|\partial_t \Phi\|_{L^\infty(\Sigma_t)}$ blows up like $c/(T_{(\text{Lifespan})} - t)$, where c is a positive data-dependent constant.

[18] In the shock formation problems described in Section 1D, the integrating factor is the inverse foliation density of a family of characteristic hypersurfaces, which are the level sets of an eikonal function. We also encounter the inverse foliation density (which we denote by μ) in Section 1E, where we derive a shock formation result in one spatial dimension.

to close the energy estimates. On the other hand, the singularities that form in the solutions from our main results are much more severe in that Φ and $\partial_t \Phi$ blow up; in contrast, in the shock formation results (see Section 1D) for equations whose principal part is similar to that of (1.1), $|\Phi|$ and $\max_{\alpha=0,1,2,3} |\partial_\alpha \Phi|$ remain bounded up to the singularity, while $\max_{\alpha,\beta=0,1,2,3} |\partial_\alpha \partial_\beta \Phi|$ blows up in finite time. Our approach to proving Theorem 1.2 also has some parallels with Kichenassamy's stable blowup-results [1996] for semilinear wave equations with exponential nonlinearities, but we delay further discussion of this point until the next subsection.

1D. ***Our results in the context of prior breakdown-results.*** There are many prior breakdown-results for solutions to various hyperbolic equations, especially of wave type. Here we give a nonexhaustive account of some of these works, which is meant to give the reader some feel for the kinds of results that are known and how they compare with/contrast against our main results. In particular, we aim to expose how the proof techniques vary considerably between different types of breakdown-results. We separate the results into seven classes.

(1) (proofs of blowup by contradiction) For various hyperbolic systems, there are proofs of blowup by contradiction, based on showing that, for smooth solutions, certain spatially averaged quantities satisfy ordinary differential inequalities that force them to blow up in finite time, contradicting the assumed smoothness. Notable contributions of this type are [John 1979; 1981] on several classes of nonlinear wave equations with signed nonlinearities and Sideris' proof [1984a] of blowup for various hyperbolic systems, for semilinear wave equations in higher dimensions [Sideris 1984b] (which improved upon Kato's result [1980]), and for the compressible Euler equations in three spatial dimensions [Sideris 1985]. See also [Guo and Tahvildar-Zadeh 1999] for similar results in the case of the relativistic Euler and Euler–Maxwell equations. None of these results yield constructive information about the nature of the blowup, nor do they apply to the wave equations under study here.

(2) (blowup for semilinear wave equations with power-law nonlinearities) There are many interesting constructive blowup-results, in various spatial dimensions, for focusing semilinear wave equations of the form $\square_m \Phi = -|\Phi|^{p-1}\Phi$, where $\square_m := -\partial_t^2 + \Delta$ is the wave operator of the Minkowski metric m. A notable difference between these works and our work here is that these works relied on a careful analysis of the spectrum of a linearized operator. We now discuss some specific examples.

In one spatial dimension, sharp results are known about the structure of singularities. For example, [Merle and Zaag 2007] showed that, for $p \in (1, \infty)$, for general initial data, there is a one-parameter family of functions that serve as the blowup-profiles relative to self-similar variables at noncharacteristic points belonging to the boundary of the maximal development; see also Remark 1.3.

In more than one spatial dimension, much less is known. Under the assumption of radial symmetry, Donninger [2010] proved the nonlinear stability of the ODE blowup solutions $\Phi_{(\mathrm{ODE});T} := c_p (T-t)^{-2/(p-1)}$, where[19] $p = 3, 5, 7, \ldots$. More precisely, using "similarity coordinates", he proved stability only in the interior of the backward light cone emanating from the singularity. In three spatial dimensions, in the

[19] Actually, for convenience, Donninger [2010] considered the semilinear term $-\Phi^p$. However, as he noted there, his work could be extended to apply to the term $-|\Phi|^{p-1}\Phi$ for $p > 1$.

subcritical cases $p \in (1, 3]$, [Donninger and Schörkhuber 2012] proved an asymptotic stability result for $\Phi_{(\mathrm{ODE});T}$ under radially symmetric perturbations of the data in the energy space, again only in the interior of the backward light cone emanating from the singularity. The result [Donninger and Schörkhuber 2012] sharpened (in the near-ODE case) the works [Merle and Zaag 2003; 2005], which showed that *all* solutions that form a singularity cannot blow up faster than the ODE rate $(T-t)^{-2/(p-1)}$, but which did not yield any information about the profile of the solution near the singularity. Donninger and Schörkhuber [2014] extended their stability results (still within radial symmetry) to the supercritical cases $p > 3$, but they assumed additional regularity on the initial data (which they believed to be essential for closing the proof). In [Donninger and Schörkhuber 2016], the authors proved results similar to those of [Donninger and Schörkhuber 2014], but without the assumption of radial symmetry. See also [Chatzikaleas and Donninger 2019] for extensions of the results of [Donninger and Schörkhuber 2016] to the case of the cubic wave equations in spatial dimensions belonging to the set $\{5, 7, 9, 11, 13\}$. In [Donninger 2017], in three spatial dimensions under the assumption of radial symmetry, Donninger established Strichartz estimates for solutions to wave equations featuring a self-similar potential, and as an application, in the critical case $p = 5$, he showed the stability of the ODE blowup in the interior of the backward light cone emanating from the singularity. In $n \geq 1$ spatial dimensions without symmetry assumptions, under the assumptions $1 < p$ if $n = 1$ and $1 < p < 1 + 4/(n-1)$ if $n \geq 2$, [Merle and Zaag 2016] used similarity coordinates to show that near a set of equilibria, solutions are either nonglobal, converge to 0, or converge to an explicit equilibrium solution. The authors also tied the various possibilities to the nature of the boundary of the maximal development (i.e., whether or not a certain point on the boundary is characteristic can affect which of the possibilities can occur there).

In three spatial dimensions, in the critical case $p = 5$, there are many blowup-results tied to the ground state solution $W(r) := (1 + r^2/3)^{-1/2}$. For solutions with (conserved) energy below that of the ground state, [Kenig and Merle 2008] established a sharp dichotomy showing that solutions blow up in finite time to the past and future if $\|\Phi\|_{\dot{H}^1(\Sigma_0)} > \|W\|_{\dot{H}^1(\Sigma_0)}$, while they exist globally and scatter if $\|\Phi\|_{\dot{H}^1(\Sigma_0)} < \|W\|_{\dot{H}^1(\Sigma_0)}$. For the same equation, the authors of [Krieger et al. 2009] proved the existence of radially symmetric "slow" type-II blowup solutions $\Phi(t, r) = \lambda^{1/2}(t) W(\lambda(t) r) + w(t, r)$, where w is a small error term, $\lambda(t) := t^{-1-\nu}$, $\nu > \frac{1}{2}$, and the singularity occurs at $t = 0$. In this context, a type-II singularity is such that the solution remains uniformly bounded in the energy space (which is critical) up to the time of first blowup. The results were extended to $\nu > 0$ in [Krieger and Schlag 2014]. In [Donninger et al. 2014], the results were extended to cases in which $\lambda(t)$ does not behave like a power law. Hillairet and Raphaël [2012] constructed type-II blowup solutions for the critical focusing wave equation in four spatial dimensions. Jendrej [2017] treated the case of five spatial dimensions. For the radial critical focusing wave equation in three spatial dimensions, [Duyckaerts et al. 2011] yielded that if the blowup-time T is finite and if the quantitative type-II condition $\sup_{t\in[0,T)}\{\|\partial_t\Phi\|^2_{L^2(\Sigma_t)} + \|\nabla\Phi\|^2_{L^2(\Sigma_t)}\} \leq \|\nabla W\|^2_{L^2} + \eta_0$ holds, where W is the ground state and $\eta_0 > 0$ is a small constant, then the blowup asymptotics are of the type exhibited by the solutions constructed in [Krieger et al. 2009]. The results were extended to the nonradial case in three and five spatial dimensions in [Duyckaerts et al. 2012b]. Similar results were obtained in the case of four spatial dimensions in [Côte et al. 2018] in the radial case. In [Duyckaerts et al. 2013], the authors gave a detailed description of the possible large-time behaviors of all finite-energy radial solutions to the focusing critical wave equation in three spatial dimensions, extending the work [Duyckaerts et al. 2012a], where information along a sequence of times was obtained. For $n \in \{3, 4, 5\}$ spatial dimensions, [Jendrej 2016] proved an upper bound for the blowup-rate $\lambda(t)$ for type-II blowup solutions whose asymptotics are $\Phi(t, r) = [\lambda(t)]^{(n-2)/2} W(\lambda(t)r) + w(t, r)$, assuming that w is sufficiently regular. In 11 or more spatial dimensions, [Collot 2018] considered a range of p-values that are energy-supercritical. For each sufficiently large integer ℓ, he constructed a codimension-$(\ell-1)$ Lipschitz manifold of spherically symmetric solutions that blow up like

$$\frac{1}{\lambda^{2/(p-1)}}(t) Q\left(\frac{r}{\lambda(t)}\right),$$

where $Q(r)$ is the ground state profile, $\lambda \sim c(T-t)^{\ell/\alpha}$, and $\alpha > 2$ is a constant that depends on p and the number of spatial dimensions.

(3) (constructive blowup-results for wave maps) There are related blowup-results for some wave maps whose targets admit a nontrivial harmonic map. For example, for the critical case of the wave maps equation $\square_m \Phi = \Phi(|\partial_t \Phi|^2 - |\nabla \Phi|^2)$, where $\Phi : \mathbb{R}^{1+2} \to \mathbb{S}^2$, under the equivariant symmetry assumption $\Phi(t, r, \theta) = (k\theta, \phi(t, r))$, where the first and second entries on the right-hand side are Euler angles parametrizing $\mathbb{S}^2$ and $k \in \mathbb{Z}_+$, there are blowup-results tied to the ground state $Q(r) := 2\arctan(r^k)$. Rodnianski and Sterbenz [2010] gave a sharp description of *stable* blowup when $k \geq 4$. They showed that (under the symmetry assumptions) there is an open set of data, with energy slightly larger than the ground state, such that the corresponding solutions blow up at a time $T < \infty$. Moreover, the asymptotics can be described as $\phi(t, r) = Q(r/\lambda(t)) + q(t, r)$, where $\lambda(t) \to 0$ as $t \uparrow T$, $\lambda(t) \gtrsim (T-t)/|\ln(T-t)|^{1/4}$, and $(q, \partial_t q)$ is small in $\dot{H}^1 \times L^2$. In particular, Q is the universal blowup-profile. As in the works cited above involving $\lambda(t)$, a key point of the proof is to derive and analyze an appropriate *modulation equation*, that is, the ODE (which is coupled to the PDE) that governs the evolution of $\lambda(t)$. The function λ is somewhat analogous to the integrating factor $\mathcal{I}$ that we use in our work here. Raphaël and Rodnianski [2012] extended the results to all cases $k \geq 1$, proving *stable* blowup with

$$\lambda(t) = c_k(1 + o(1)) \frac{T-t}{|\ln(T-t)|^{1/(2k-2)}}$$

as $t \uparrow T$ in the cases $k \geq 2$, and, in the case $k = 1$, $\lambda(t) = (T-t)\exp(-\sqrt{\ln|T-t|} + O(1))$ as $t \uparrow T$. In [Krieger et al. 2008], in the case $k = 1$, the authors proved the existence of a continuum of related solutions (believed to be nongeneric) exhibiting the blowup-rates $\lambda(t) = (T-t)^\nu$, where $\nu > \frac{3}{2}$. The results were extended to $\nu > 1$ in [Gao and Krieger 2015]. In [Côte et al. 2015], in the equivariance class $k = 1$, the authors proved that within the subclass of degree-0 maps (i.e., in radial coordinates (t, r), one assumes $\phi(0, 0) = \phi(0, \infty) = 0$), there exist solutions with energy above but arbitrarily close to twice the energy of the ground state that blow up in finite time. Within the subclass of degree-1 maps (i.e., $\phi(0, 0) = 0$ and $\phi(0, \infty) = \pi$), for maps with energy bigger than that of the ground state but less than three times the energy of the ground state, the authors show that if a singularity forms, then the solution has asymptotics whose blowup-profiles are the same as those from the works [Krieger et al. 2008; Rodnianski and Sterbenz 2010; Raphaël and Rodnianski 2012]. For equivariant wave maps $\Phi : \mathbb{R} \times \mathbb{S}^2 \to \mathbb{S}^2$, in the class $k = 1$, [Shahshahani 2016] proved the existence of a continuum of blowup solutions. In [Donninger 2011; Donninger et al. 2012], in the supercritical context of equivariant wave maps from $\mathbb{R}^{1+3}$ into $\mathbb{S}^3$, the authors proved the stability of self-similar blowup solutions $\phi_T(t, r) := 2\arctan(r/(T - r))$. More precisely, those results relied on some mode stability results that were later proved in [Costin et al. 2016]; see also [Costin et al. 2017] for similar results in a more general context.

(4) (blowup for semilinear wave equations with exponential nonlinearities) For the focusing semilinear wave equation $\square_m \Phi = -e^{\Phi}$ in three spatial dimensions, [Kichenassamy 1996] proved that the singular solution $\ln(2/t^2)$ is stable under perturbations of the data along the constant-time hypersurface $\{t = -1\}$. Moreover, he showed that the blowup-surface is of the form $\{t = f(x)\}$, where $f(x)$ loses Sobolev regularity compared to the initial data. It would be interesting to see if our main results could be extended to show a similar result for the equations under study here. More precisely, we conjecture that a portion of the blowup-surface (which would be a subset of the boundary of the maximal development mentioned in Remark 1.3) is $\{\mathcal{I} = 0\}$ for the solutions under study here. Kichenassamy's work has one key feature in common with ours: he devised a reformulation of the wave equation in which no singularity was visible, in his case by constructing a singular change of coordinates adapted to the singularity; this is broadly similar to the approach taken by authors who have proved shock formation results, as we describe just below. However, unlike the "forwards approach" that we take in this article, Kichenassamy used a "backwards approach" in which he first solved problems in which the singularity was *prescribed* along blowup-surfaces and then showed that the map from the singularity to the Cauchy data along $\{t = -1\}$ is invertible; see also Remark 1.4 concerning backwards approaches. His proof of the invertibility of the map from the singularity data to the Cauchy data was based on studying appropriately linearized versions of the equations and on using Fuchsian techniques. The linearized equations exhibited derivative loss, and Kichenassamy used a Nash–Moser approach to handle the loss.

(5) (shock formation for quasilinear equations) Roughly speaking, a shock singularity is a "mild" type of singularity such that the solution remains bounded but one of its derivatives blows up. More precisely, shock singularities are tied to the intersection of a family of characteristics (which, in one spatial dimension, are curves), and the blowup occurs only for derivatives of the solution in directions transversal to the characteristics. There are many classical shock formation results in one spatial dimension, based on the method of characteristics, with important contributions coming from [Riemann 1860; Lax 1964; 1972; 1973; John 1974], among others. Even in one spatial dimension, the field is still active, as is evidenced by the recent work [Christodoulou and Perez 2016], which significantly sharpened [John 1974], giving a complete description of shock formation for electromagnetic plane waves in nonlinear crystals.

Shock singularities are of interest in particular because, at least for systems with suitable structure, there is hope of uniquely continuing the solution past the shock in a weak sense, subject to appropriate selection criteria (typically tied to Rankine–Hugoniot jump conditions across the shock curve, that is, the curve across which the solution exhibits a jump discontinuity). In one spatial dimension, there is an adequate rigorous mathematical theory, at least for strictly hyperbolic quasilinear PDE systems, that is able to accommodate the formation of shocks, the evolution of the solution after shocks, and subsequent interactions of the shocks. We stress that the one-dimensional theory fundamentally relies on the availability of PDE estimates in the space of functions of bounded variation. We refer readers to the comprehensive work [Dafermos 2000] for a detailed discussion of the theory of shock waves in one spatial dimension.

We now turn our attention to the case of more than one spatial dimension. As of the present, in more than one spatial dimension, there is no broad, rigorous well-posedness theory for solutions to quasilinear hyperbolic PDE systems that is able to accommodate the formation of shocks, the evolution of the solution after shocks, and subsequent interactions of the shocks. The main difficulty is that, in view of the fundamental result [Rauch 1986], bounded variation estimates typically fail for quasilinear hyperbolic systems. This prevents one from being able to directly extend the theory described in the previous paragraph to the case of multiple spatial dimensions. For this reason, in multiple spatial dimensions, it seems that one is forced to work with Sobolev spaces and to derive energy estimates up to top order, a task that has proven to be exceptionally difficult in the neighborhood of a shock singularity. Although in multiple spatial dimensions the theory of what happens *after* shocks form is in its infancy,[20] the rigorous theory of the *formation* of a shock, starting from smooth initial conditions, has undergone dramatic advancements in recent years. We refer readers to the survey article [Holzegel et al. 2016] for discussion concerning the history of the subject and for an overview of some recent shock formation results in the context of quasilinear wave equations in three spatial dimensions.

In more than one spatial dimension, the basic picture of what a shock singularity is remains unchanged from the case of one spatial dimension: it is a singularity such that the solution remains bounded but one of its derivatives blows up, and the blowup is tied to the intersection of a family of characteristic hypersurfaces, as in the case of one spatial dimension. In the case of more than one spatial dimension, the characteristic hypersurfaces are levels sets of a solution u to the eikonal equation, which is a (typically) fully nonlinear transport-type equation in u whose coefficients depend on the solution to the original PDE of interest; see the next paragraph for further discussion on eikonal functions in the context of quasilinear wave equations. The use of eikonal functions to analyze solutions to quasilinear hyperbolic PDEs is tantamount to the implementation of nonlinear geometric optics; again, we refer the reader to [Holzegel et al. 2016] for further discussion of eikonal functions and their role in the study of quasilinear wave equations.

We now summarize some important results in the theory of shock formation in more than one spatial dimension. Alinhac [1999a; 1999b; 2001] obtained the first results on shock formation without symmetry assumptions in more than one spatial dimension. The main new difficulty compared to the case of one spatial dimension is that the method of characteristics (implemented via eikonal functions) must be supplemented with energy estimates, which leads to enormous technical complications. Alinhac's work

[20] We mention, however, that Majda [1981; 1983a; 1983b] has solved, in appropriate Sobolev spaces, the *shock front problem*. That is, he proved a local existence result starting from an initial discontinuity given across a smooth hypersurface contained in the Cauchy hypersurface. The data must satisfy suitable jump conditions, entropy conditions, and higher-order compatibility conditions. This is different than [Christodoulou 2019] on the restricted shock development problem, which we describe below. A key difference is that [Christodoulou 2007; 2019] together describe the *emergence* of the shock hypersurface from an initially smooth solution, whereas Majda *prescribed* the initial singularity.

applied to small-data solutions to a class of scalar quasilinear wave equations of the form

$$(g^{-1})^{\alpha\beta}(\partial\Phi)\partial_\alpha\partial_\beta\Phi = 0 \tag{1D.1}$$

that fail to satisfy the null condition. In (1D.1), $g = g(\partial\Phi)$ is a Lorentzian metric that depends on the solution. Alinhac showed that, for a set of "nondegenerate" small data, Φ and $\partial\Phi$ remain bounded, while $\partial^2\Phi$ blows up in finite time due to the intersection of the characteristics. As we alluded to above, his proof fundamentally relied on nonlinear geometric optics, that is, on an eikonal function, which in the case of scalar quasilinear wave equations is a solution to the eikonal equation

$$(g^{-1})^{\alpha\beta}(\partial\Phi)\partial_\alpha u\partial_\beta u = 0, \tag{1D.2}$$

supplemented with appropriate initial data. The level sets of u are characteristic hypersurfaces[21] for (1D.1). As it turns out, the intersection of the level sets of u is tied to the formation of a singularity in the solution to (1D.1), just as in the case of one spatial dimension. Much like in the present work, the main estimates in Alinhac's proof did not concern singularities; the crux of his proof was to construct a system of geometric coordinates, one of which is u, and to prove that relative to them, the solution remains very smooth, except possibly at the very high derivative levels. He then showed that a singularity forms in the standard second-order derivatives $\partial^2\Phi$ as a consequence of a finite-time degeneracy between the geometric coordinates and the standard coordinates; roughly, the level sets of u intersect and cause the blowup, much like in the classical example of the blowup of solutions to Burgers' equation. The main challenge in the proof is that to derive energy estimates relative to the geometric coordinates, one must control the eikonal function, whose top-order regularity properties are difficult to obtain; naive estimates lead to the loss of a derivative.

The intricate regularity properties of eikonal functions had previously been understood in certain problems for quasilinear wave equations in which singularities did not form. For example, the first quasilinear wave problem in which the regularity properties of eikonal functions were fully exploited was the celebrated proof [Christodoulou and Klainerman 1993] of the stability of Minkowski spacetime as a solution to the Einstein vacuum equations. Eikonal functions have also played a central role in proofs of low-regularity well-posedness for quasilinear wave equations [Klainerman and Rodnianski 2003; 2005; Smith and Tataru 2005; Klainerman et al. 2015]. However, unlike in these problems, in the problem of shock formation, the top-order geometric energy estimates feature a degenerate weight that vanishes as the shock forms, which leads to a priori estimates allowing for the possibility that the high-order energies might blow up; note that this possible geometric energy blowup is distinct from the formation of the shock, which happens at the low derivative levels relative to the standard coordinates. The "degenerate weight" mentioned above is the inverse foliation density[22] of the level sets of u. The inverse foliation density vanishes when the characteristics intersect, and it is in some ways analogous to the integrating factor $\mathcal{I}$ that we use in our work here. Alinhac closed his singular top-order energy estimates with a Nash–Moser iteration scheme that was adapted to the singularity and that handled the issue of

[21] In the context of wave equations, characteristic hypersurfaces are often referred to as "null hypersurfaces" due to their intimate connection to the Lorentzian notion of a null vector field.

[22] In Section 1E, we encounter the inverse foliation density (we denote it by μ), although we do not need to derive energy estimates in that subsection since we consider only the case of one spatial dimension there.

the regularity theory of u in a different way than [Christodoulou and Klainerman 1993; Klainerman and Rodnianski 2003; 2005; Smith and Tataru 2005; Klainerman et al. 2015]. He then used a "descent scheme" to show that the top-order geometric energy blowup does not propagate down too far to the lower derivative levels. Consequently, the solution remains highly differentiable relative to the geometric coordinates. The solution's high degree of smoothness relative to the geometric coordinates is not just an interesting artifact of the approach, but rather is fundamental to all aspects of the proof.

Due to his reliance on the Nash–Moser iteration scheme, Alinhac's proof applied only to "nondegenerate" initial data such that the first singularity is isolated in the constant-time hypersurface of first blowup, and his framework breaks down precisely at the time of first blowup. For this reason, his approach is inadequate for following the solution to the boundary of the maximal development of the data (see footnote 10), which intersects the future of the first singular time. The breakthrough work [Christodoulou 2007] overcame this drawback and significantly sharpened Alinhac's results for the subset of quasilinear wave equations that arise in the study of irrotational relativistic fluid mechanics. More precisely, Christodoulou's proof yielded a sharp description of the solution up to the boundary of the maximal development. This information was essential even for setting up the shock development problem, which, roughly speaking, is the problem of uniquely extending the solution past the singularity in a weak sense, subject to appropriate jump conditions. We note that the shock development problem in relativistic fluid mechanics was solved in spherical symmetry in [Christodoulou and Lisibach 2016] and, in the recent breakthrough work [Christodoulou 2019] for the nonrelativistic compressible Euler equations and the relativistic Euler equations without symmetry assumptions in a restricted case (known as the restricted shock development problem) such that the jump in entropy across the shock hypersurface is ignored.

The wave equations studied in [Christodoulou 2007] form a subclass of the ones (1D.1) studied by Alinhac. They enjoy special properties that Christodoulou used in his proofs, notably an Euler–Lagrange formulation such that the Lagrangian is invariant under a symmetry group. The main technical improvement afforded by Christodoulou's framework is that in closing the energy estimates, he avoided using a Nash–Moser iteration scheme and instead used an approach similar to the one employed in the aforementioned works [Christodoulou and Klainerman 1993; Klainerman and Rodnianski 2003]. This approach is more robust and is capable of accommodating solutions such that the blowup occurs along a hypersurface, which, in the problem of shock formation, is what typically occurs along a portion of the boundary of the maximal development.[23] Christodoulou's results have since been extended in many directions, including to apply to other wave equations [Christodoulou and Miao 2014; Speck 2016], different sets of initial data [Speck et al. 2016; Miao and Yu 2017; Miao 2018], the compressible Euler equations with nonzero vorticity [Luk and Speck 2016; 2018; Speck 2019b], systems of wave equations with multiple speeds [Speck 2018], and quasilinear systems in which a solution to a transport equation forms a shock [Speck 2019a]. Some of the earlier extensions are explained in detail in the survey article [Holzegel et al. 2016].

(6) (breakdown-results for Einstein's equations) The Einstein equations of general relativity have many remarkable properties and as such, it is not surprising that there are stable breakdown-results that are

[23] Roughly, a portion of the boundary is equal to the zero level set of the inverse foliation density.

specialized to these equations. Here we simply highlight the following constructive results in three spatial dimensions without symmetry assumptions: Christodoulou's breakthrough results [2009] on the formation of trapped surfaces and the stable singularity formation results [Rodnianski and Speck 2018a; 2018b; Luk 2018]. The work [Rodnianski and Speck 2018b] can be viewed as a stable ODE-type blowup-result for Einstein's equations in which the wave speed became infinite at the singularity. Note that in contrast, for (1B.1a), the wave speed vanishes when $\partial_t \Phi$ blows up.

(7) (finite-time degeneration of hyperbolicity) In [Speck 2017], we studied the wave equations

$$-\partial_t^2 \Psi + (1+\Psi)^P \Delta \Psi = 0$$

in three spatial dimensions for $P = 1, 2$. We showed that there exists an open set of initial data such that Ψ is initially small but $1+\Psi$ vanishes in finite time, corresponding to a breakdown in the hyperbolicity of the equation, but without any singularity forming. The difficult part of the proof is closing the energy estimates in regions where $1+\Psi$ is small. The proof has some features in common with the proof of the main results of this paper. For example, the proof relies on monotonicity tied to the sign of $\partial_t \Psi$ and the small size of $\nabla \Psi$. In particular, this leads to the availability of a friction-type integral in the energy identities, analogous to the one (1C.4), which is crucially important for controlling certain error terms.

1E. ***Different kinds of singularity formation within the same quasilinear hyperbolic system.*** In this subsection, we show that there are quasilinear wave equations, closely related to the wave equation (1B.1a), that can exhibit two distinct kinds of blowup: ODE-type blowup for one set of data, and the formation of shocks for another set. The ODE-type blowup is provided by our main results, so in this subsection, most of the discussion is centered on shock formation. Our discussion is based on ideas and techniques found in [Christodoulou 2007; Speck 2016].

To initiate the discussion, we define

$$\Phi_0 := \partial_t \Phi. \tag{1E.1}$$

For convenience, we will restrict our discussion to the specific weight

$$\mathscr{W} = \frac{1}{1+\partial_t \Phi} = \frac{1}{1+\Phi_0},$$

though similar results hold for any weight that satisfies the assumptions of Section 2A. To proceed, we differentiate (1B.1a) with ∂_t to obtain the following closed equation in Φ_0:

$$\partial_t^2 \Phi_0 - \frac{1}{1+\Phi_0} \Delta \Phi_0 = -\frac{1}{1+\Phi_0} (\partial_t \Phi_0)^2 + \frac{2\Phi_0}{1+\Phi_0} \partial_t \Phi_0 + \frac{3\Phi_0^2}{1+\Phi_0} \partial_t \Phi_0. \tag{1E.2}$$

In the remainder of our discussion of shock formation, we will only consider plane-symmetric solutions, that is, solutions that depend only on t and x^1. Note that in (1E.2), $\Delta = \partial_1^2$ for plane-symmetric solutions.

To study plane-symmetric solutions to (1E.2), we will use the characteristic vector fields

$$L := \partial_t + \frac{1}{\sqrt{1+\Phi_0}} \partial_1, \quad \underline{L} := \partial_t - \frac{1}{\sqrt{1+\Phi_0}} \partial_1. \tag{1E.3}$$

We next define the characteristic coordinate u to be the solution to the following initial value problem for a transport equation:

$$Lu = 0, \quad u|_{\Sigma_0} = 1 - x^1. \tag{1E.4}$$

Above and throughout, $Xf := X^\alpha \partial_\alpha f$ denotes the derivative of the scalar function f in the direction of the vector field X. Our choice of the vector field L in (1E.4) is adapted to the problem of detecting "right-moving" shock formation. Similar results could be proved in the "left-moving" case; one could analyze such solutions by constructing a characteristic coordinate $\underline{u}$ that solves $\underline{L}\underline{u} = 0$. We view u as a new coordinate adapted to the characteristics, and we will use the "geometric" coordinate system (t, u) when analyzing solutions. In particular, the level sets of u are characteristic for (1E.2). In fact, it is not difficult to see that as a consequence of (1E.4), u is a solution to the eikonal equation (1D.2), where the inverse metric $(g^{-1})^{\alpha\beta}$ is determined by the principal-order coefficients in the wave equation (1E.2).

We now define $\Sigma_t^{u'}$, relative to the geometric coordinates, to be the subset $\Sigma_t^{u'} := \{(t, u) \mid 0 \le u \le u'\}$. Note that Σ_0^1 can be identified with an orientation-reversed version of the unit x^1 interval [0, 1]. Associated to u, we define the *inverse foliation density* $\mu > 0$ by

$$\mu := \frac{1}{\partial_t u}. \tag{1E.5}$$

Then $1/\mu$ is a measure of the density of the level sets of u, and $\mu = 0$ corresponds to the intersection of the characteristics, that is, the formation of a shock. From (1E.4), it follows that $\mu|_{\Sigma_0} = \sqrt{1+\Phi_0} = 1 + \mathcal{O}(\Phi_0)$ (for Φ_0 small). One can check that from the above definitions, we have, in addition to (1E.4), the identities $Lt = 1$, $\mu \underline{L} t = \mu$, and $\mu \underline{L} u = 2$. In particular, $L = \frac{d}{dt}$ along the integral curves of L and $\mu \underline{L} = 2\frac{d}{du}$ along the integral curves of $\mu \underline{L}$.

Next, we differentiate (1B.1a) and (1E.4) with ∂_t and carry out tedious but straightforward calculations to obtain the following system in Φ_0 and μ:

$$L(\mu\underline{L}\Phi_0) = -\frac{1}{2(1+\Phi_0)}(L\Phi_0)(\mu\underline{L}\Phi_0) + \mu\frac{\Phi_0}{1+\Phi_0}\left\{1+\tfrac{3}{2}\Phi_0\right\}L\Phi_0 + \frac{\Phi_0}{1+\Phi_0}\left\{1+\tfrac{3}{2}\Phi_0\right\}(\mu\underline{L}\Phi_0), \tag{1E.6a}$$

$$\begin{aligned}\mu\underline{L}L\Phi_0 = &-\frac{\mu}{4(1+\Phi_0)}(L\Phi_0)^2 - \frac{3}{4(1+\Phi_0)}(L\Phi_0)(\mu\underline{L}\Phi_0) \\ &+ \mu\frac{\Phi_0}{1+\Phi_0}\left\{1+\tfrac{3}{2}\Phi_0\right\}L\Phi_0 + \frac{\Phi_0}{1+\Phi_0}\left\{1+\tfrac{3}{2}\Phi_0\right\}(\mu\underline{L}\Phi_0),\end{aligned} \tag{1E.6b}$$

$$L\mu = \frac{1}{4(1+\Phi_0)}\mu L\Phi_0 + \frac{1}{4(1+\Phi_0)}(\mu\underline{L}\Phi_0). \tag{1E.6c}$$

For convenience, we will prove shock formation only for a restricted class of initial data supported in Σ_0^1; as can easily be inferred from our proof, the shock-forming solutions that we will construct are stable under plane-symmetric perturbations, and our approach could be applied to a much larger set of plane-symmetric initial data. Specifically, we will prove shock formation for solutions corresponding to initial data such that

$$\sup_{\Sigma_0^1} |\Phi_0| \le \varepsilon, \quad L\Phi_0|_{\Sigma_0} = 0, \quad \sup_{\Sigma_0^1} |\underline{L}\Phi_0| = 4, \tag{1E.7}$$

such that $\underline{L}\Phi_0|_{\Sigma_0^1}$ is *negative* at some maximum of $|\underline{L}\Phi_0|$ on Σ_0^1, and such that ε is small. The negativity of $\underline{L}\Phi_0$ will drive the vanishing of μ. To show the existence of such data, it is convenient to refer to the Cartesian coordinate x^1. Specifically, we fix a smooth nontrivial function $f = f(x^1)$ supported in Σ_0^1 and set $\Phi_0|_{\Sigma_0^1}(x^1) := \kappa f(\lambda x^1)$, where κ and λ are real parameters. Note that $\partial_1 \Phi_0|_{\Sigma_0}(x^1) = \kappa\lambda f'(\lambda x^1)$. We then set

$$\partial_t \Phi_0|_{\Sigma_0^1} := -\frac{1}{\sqrt{1+\Phi_0|_{\Sigma_0^1}}} \partial_1 \Phi_0|_{\Sigma_0^1},$$

which implies that $L\Phi_0|_{\Sigma_0^1} = 0$ and

$$\underline{L}\Phi_0|_{\Sigma_0^1}(x^1) = -2\kappa\lambda \frac{1}{\sqrt{1+\kappa f(\lambda x^1)}} f'(\lambda x^1).$$

We now choose $|\kappa|$ sufficiently small and λ sufficiently large, which allows us to achieve (1E.7) with $\varepsilon > 0$ arbitrarily small. Moreover, by adjusting the sign of κ, we can ensure that $\underline{L}\Phi_0|_{\Sigma_0^1}$ is negative at some maximum of $|\underline{L}\Phi_0|$ on Σ_0^1. We also note that from domain of dependence considerations, it follows that in terms of the geometric coordinates, solutions with data supported in Σ_0^1 vanish when $u \leq 0$, and that the level set $\{u = 0\}$ can be described in Cartesian coordinates as $\{(t, x^1) \mid 1 - x^1 + t = 0\}$.

To derive estimates, we make the following bootstrap assumptions on any region of classical existence such that $0 \leq t \leq 2$ and $0 \leq u \leq 1$:

$$0 < \mu \leq 3, \quad |\Phi_0| \leq \sqrt{\varepsilon}, \quad |L\Phi_0| \leq \sqrt{\varepsilon}, \quad |\underline{L}\Phi_0| \leq 5. \tag{1E.8}$$

Note also that the solution satisfies $\Phi_0(t, u = 0) = 0$ and that the assumptions (1E.8) are consistent with the initial data when ε is small.

We now derive estimates. In the rest of the subsection, we will silently assume that $\varepsilon > 0$ is sufficiently small. We define

$$Q(t,u) := \sup_{(t',u') \in [0,t] \times [0,u]} \{|\Phi_0|(t',u') + |L\Phi_0|(t',u')\}. \tag{1E.9}$$

Note that $Q(0,u) \lesssim \varepsilon$, while our data-support assumptions and finite speed of propagation imply that $Q(t,0) = 0$. Using the evolution equation (1E.6b), the bootstrap assumptions, the fact that $L = \frac{d}{dt}$ along the integral curves of L, and the fact that $\mu\underline{L} = 2\frac{d}{du}$ along the integral curves of $\mu\underline{L}$, we deduce $Q(t,u) \leq CQ(0,u) + c\int_{t'=0}^t Q(t',u)\,dt' + c\int_{u'=0}^u Q(t,u')\,du'$. From this estimate and Gronwall's inequality (in two variables), we deduce that there are constants $C > 0$ and $c' > c$ such that, for $0 \leq t \leq 2$ and $0 \leq u \leq 1$, we have

$$Q(t,u) \leq CQ(0,u)e^{c't}e^{c'u} \leq CQ(0,u)e^{3c'} \leq Ce^{3c'}\varepsilon \lesssim \varepsilon. \tag{1E.10}$$

Using the estimate (1E.10) and the bootstrap assumptions for μ and $\mu\underline{L}\Phi_0$ to control the terms on the right-hand side of (1E.6a), we deduce $|L(\mu\underline{L}\Phi_0)| \lesssim \varepsilon$. Integrating this estimate along the integral curves of L and using that $\mu(0,u) = 1 + \mathcal{O}(\varepsilon)$, we find that, for $0 \leq t \leq 2$ and $0 \leq u \leq 1$, we have $[\mu\underline{L}\Phi_0](t,u) = [\mu\underline{L}\Phi_0](0,u) + \mathcal{O}(\varepsilon) = \underline{L}\Phi_0(0,u) + \mathcal{O}(\varepsilon)$. Inserting this information into (1E.6c) and using (1E.10), we deduce $L\mu = \frac{1}{4}\underline{L}\Phi_0(0,u) + \mathcal{O}(\varepsilon)$. Integrating in time and using the initial condition $\mu(0,u) = 1 + \mathcal{O}(\varepsilon)$, we deduce that $\mu(t,u) = 1 + \frac{1}{4}\underline{L}\Phi_0(0,u)t + \mathcal{O}(\varepsilon) = 1 + \frac{1}{4}[\mu\underline{L}\Phi_0](t,u)t + \mathcal{O}(\varepsilon)$.

We now note that if ε is sufficiently small, then the above estimates yield strict improvements of the bootstrap assumptions (1E.8). By a standard continuity argument in t and u, this justifies the bootstrap assumptions and shows that the solution exists on regions of the form $0 \le t \le 2$ and $0 \le u \le 1$, as long as μ remains positive; the positivity of μ and the above estimates guarantee that $|\Phi_0| + \max_{\alpha=0,1} |\partial_\alpha \Phi_0|$ is finite. Moreover, since (by construction) $\sup_{\Sigma_0^1} |\underline{L}\Phi_0| = 4$ and since there is a value $u_* \in (0, 1)$ such that $\underline{L}\Phi_0(0, u_*) = -4$, the above estimates for $\mu\underline{L}\Phi_0$ and μ guarantee that $\min_{\Sigma_t^1} \mu = 1 + \mathcal{O}(\varepsilon) - t$ and that at points $(t, u) \in [0, 2] \times [0, 1]$ of classical existence with $\mu(t, u) \le \frac{1}{4}$, we have $\mu\underline{L}\Phi_0(t, u) \le -1$. It follows that $\min_{\Sigma_t^1} \mu$ cannot remain positive for times larger than $1 + \mathcal{O}(\varepsilon)$ and that

$$\min_{\Sigma_t^1} \mu \le \tfrac{1}{4} \quad \Longrightarrow \quad \sup_{\Sigma_t^1} |\underline{L}\Phi_0| \ge \frac{1}{\min_{\Sigma_t^1} \mu}.$$

In total, these arguments yield that $\sup_{\Sigma_t^1} |\underline{L}\Phi_0|$ blows up at some time $t_{(\text{Shock})} = 1 + \mathcal{O}(\varepsilon)$, while $|\Phi_0|$ and $|L\Phi_0|$ remain uniformly bounded by $\lesssim \varepsilon$. We have thus shown that a shock forms. Readers can consult [Holzegel et al. 2016] for further discussion of the style of proof of shock formation given here and extensions to the case of more than one spatial dimension.

We now revisit the solutions from our main results under the weight $\mathscr{W}(\partial_t\Phi) := 1/(1 + \partial_t\Phi)$. Notice that, for such solutions, Φ_0 also solves (1E.2) but is such that $|\Phi_0|$ blows up at the singularity. This is *different blowup behavior* compared to the shock-forming solutions to (1E.2) constructed above, in which $|\Phi_0|$ remained bounded. Notice also that our main theorem requires, roughly, that $\Phi_0|_{\Sigma_0}$ should not be too small, which is in contrast to the initial data for the shock-forming formation solutions described above. To close this subsection, we clarify that it could be, in principle, that the ODE-type blowup solutions that we have constructed are *unstable* when viewed as solutions to (1E.2), even though they are stable solutions of the original wave equation (1B.1a). The key point is that to solve (1E.2) (viewed as a wave equation for Φ_0), we need to prescribe the data functions $\Phi_0|_{\Sigma_0}$ and $\partial_t\Phi_0|_{\Sigma_0}$, whereas for the ODE-type blowup solutions we have constructed, we can freely prescribe (in plane symmetry) only $\Phi_0|_{\Sigma_0}$; the quantity $\partial_t\Phi_0|_{\Sigma_0}$ is not "free", but rather is uniquely determined from $\Phi_0|_{\Sigma_0}$ and $\partial_1\Phi|_{\Sigma_0}$ via the wave equation (1B.1a). Put differently, the ODE-type blowup solutions that we have constructed yield "special" solutions to (1E.2) that are constrained by the fact that Φ_0 is the time derivative of a solution to the original wave equation (1B.1a). In contrast, we expect that the methods of [Speck et al. 2016] could be used to show that the plane-symmetric shock-forming solutions to (1E.2) that we constructed in this subsection are stable under perturbations that break the plane symmetry.

1F. ***Notation.*** In this subsection, we summarize some notation that we use throughout.

- $\{x^\alpha\}_{\alpha=0,1,2,3}$ are the standard Cartesian coordinates on $\mathbb{R}^{1+3} = \mathbb{R}\times\mathbb{R}^3$ and $\partial_\alpha := \frac{\partial}{\partial x^\alpha}$ are the corresponding coordinate partial derivative vector fields; $x^0 \in \mathbb{R}$ is the time coordinate and $\underline{x} := (x^1, x^2, x^3) \in \mathbb{R}^3$ are the spatial coordinates.
- We often use the alternative notation $x^0 = t$ and $\partial_0 = \partial_t$.
- $\Sigma_t := \{(t, \underline{x}) \mid \underline{x} \in \mathbb{R}^3\}$ is the standard flat hypersurface of constant time.

• Greek "spacetime" indices such as α vary over 0, 1, 2, 3, while Latin "spatial" indices such as a vary over 1, 2, 3. We use primed indices, such as a', in the same way that we use their nonprimed counterparts. We use Einstein's summation convention in that repeated indices are summed over their respective ranges.

• We sometimes omit the arguments of functions appearing in pointwise inequalities. For example, we sometimes write $|f| \leq C\mathring{\epsilon}$ instead of $|f(t,\underline{x})| \leq C\mathring{\epsilon}$.

• $\nabla^k\Psi$ denotes the array comprising all k-th-order derivatives of Ψ with respect to the Cartesian spatial coordinate vector fields. We often use the alternative notation $\nabla\Psi$ in place of $\nabla^1\Psi$. For example, $\nabla^1\Psi = \nabla\Psi := (\partial_1\Psi, \partial_2\Psi, \partial_3\Psi)$.

• $|\nabla^{\leq k}\Psi| := \sum_{k'=0}^{k} |\nabla^{k'}\Psi|$.

• $|\nabla^{[a,b]}\Psi| := \sum_{k'=a}^{b} |\nabla^{k'}\Psi|$.

• $H^N(\Sigma_t)$ denotes the standard Sobolev space of functions on Σ_t. If $N \geq 0$ is an integer, then the corresponding norm is

$$\|f\|_{H^N(\Sigma_t)} := \left\{ \sum_{a_1+a_2+a_3\leq N} \int_{\underline{x}\in\mathbb{R}^3} |\partial_1^{a_1}\partial_2^{a_2}\partial_3^{a_3} f(t,\underline{x})|^2 \, d\underline{x} \right\}^{1/2}.$$

In the case $N = 0$, we use the standard notation "L^2" in place of "H^0".

• Above and throughout, $d\underline{x} := dx^1\,dx^2\,dx^3$ denotes standard Lebesgue measure on Σ_t.

• $L^\infty(\Sigma_t)$ denotes the standard Lebesgue space of functions on Σ_t with corresponding norm $\|f\|_{L^\infty(\Sigma_t)} := \operatorname{ess\,sup}_{\underline{x}\in\mathbb{R}^3} |f(t,\underline{x})|$.

• If A and B are two quantities, then we often write $A \lesssim B$ to indicate that "there exists a constant $C > 0$ such that $A \leq CB$".

• We sometimes write $A = \mathcal{O}(B)$ to signify that there exists a constant $C > 0$ such that $|A| \leq C|B|$.

• Explicit and implicit constants are allowed to depend on the data-size parameters $\mathring{A}$ and $\mathring{A}_*^{-1}$ from Section 3A, in a manner that we more fully explain in Section 5A.

2. Mathematical setup and the evolution equations

In this section, we state our assumptions on the nonlinearities, define the quantities that we will study in the rest of the paper, and derive evolution equations.

2A. ***Assumptions on the weight.*** Let $\mathscr{W}$ be the scalar function from (1B.1a). We assume that there are constants $C_k > 0$ such that

$$\mathscr{W}(y) > 0, \quad y \in \left(-\tfrac{1}{2}, \infty\right), \tag{2A.1}$$

$$\mathscr{W}(0) = 1, \tag{2A.2}$$

$$\mathscr{W}'(y) \leq 0, \quad y \in [0, \infty), \tag{2A.3}$$

$$\left| \left\{ (1+y)^2 \frac{d}{dy} \right\}^k [(1+y)\mathscr{W}(y)] \right| \leq C_k, \quad 0 \leq k \leq 5, \ y \in \left(-\tfrac{1}{2}, \infty\right). \tag{2A.4}$$

We also assume that there is a constant $\alpha > 0$ such that

$$\mathscr{W}(y) \leq \alpha |\mathscr{W}'(y)|^{1/2}, \quad y \in [1, \infty). \tag{2A.5}$$

Note that (2A.1), (2A.3), and (2A.5) imply in particular that

$$\mathscr{W}'(y) < 0, \quad y \in [1, \infty). \tag{2A.6}$$

2B. *The integrating factor and the renormalized solution variables.*

2B1. *Definitions.* As we described in Section 1C2, our analysis fundamentally relies on the following integrating factor.

Definition 2.1 (the integrating factor). Let Φ be the solution to the wave equation (1B.1a). We define $\mathcal{I} = \mathcal{I}(t, \underline{x})$ to be the solution to the transport equation

$$\partial_t \mathcal{I} = -\mathcal{I} \partial_t \Phi, \quad \mathcal{I}|_{\Sigma_0} = 1. \tag{2B.1}$$

Moreover, we define

$$\mathcal{I}_\star(t) := \min_{\Sigma_t} \mathcal{I}. \tag{2B.2}$$

Remark 2.2 (the vanishing of $\mathcal{I}$ implies singularity formation). It is straightforward to see from (2B.1) that if $\mathcal{I}(T, \underline{x}) = 0$ for some $T > 0$ and for one or more $\underline{x} \in \mathbb{R}^3$, then at such values of $\underline{x}$ we have $\lim_{t \uparrow T} \sup_{s \in [0,t)} \partial_t \Phi(s, \underline{x}) = \infty$. In fact, it follows that $\int_{s=0}^{t} |\partial_t \Phi(s, \underline{x})| \, ds = \infty$.

Most of our effort will go towards analyzing the following "renormalized" solution variables. We will show that they remain regular up to the singularity.

Definition 2.3 (renormalized solution variables). Let Φ be the solution to the wave equation (1B.1a) and let $\mathcal{I}$ be as in Definition 2.1. For $\alpha = 0, 1, 2, 3$, we define

$$\Psi_\alpha := \mathcal{I} \partial_\alpha \Phi. \tag{2B.3}$$

2B2. *A crucial identity for $\mathcal{I}$ and the $\mathcal{I}$-weighted evolution equations.* Our main goal in this subsubsection is to derive evolution equations for the renormalized solution variables; see Proposition 2.5. As a preliminary step, we first provide a lemma that shows that $\partial_i \mathcal{I}$ can be controlled in terms Ψ_i and the initial data, and that no singular factors of $\mathcal{I}^{-1}$ appear in the relationship. Though simple, the lemma is crucial for the top-order regularity theory of $\mathcal{I}$.

Lemma 2.4 (identity for the spatial derivatives of the integrating factor). *The following identity holds for $i = 1, 2, 3$, where $\{\mathring{\Psi}_a\}_{a=1,2,3}$ are the wave equation initial data from* (1B.1b):

$$\partial_i \mathcal{I} = -\Psi_i + \mathcal{I} \mathring{\Psi}_i. \tag{2B.4}$$

Proof. Dividing (2B.1) by $\mathcal{I}$ and then applying ∂_i, we compute that

$$\partial_t \left\{ \frac{\partial_i \mathcal{I} + \Psi_i}{\mathcal{I}} \right\} = 0. \tag{2B.5}$$

Integrating (2B.5) with respect to time and using the initial conditions $\mathcal{I}|_{\Sigma_0} = 1$ and $\Psi_i|_{\Sigma_0} = \mathring{\Psi}_i$, we arrive at (2B.4). □

We now derive the main evolution equations that we will study in the remainder of the paper.

Proposition 2.5 (the renormalized first-order system: $\mathcal{I}$-weighted evolution equations). *For solutions to (1B.1a)–(1B.1b), the renormalized solution variables of Definition 2.3 satisfy the system*

$$\partial_t \Psi_0 = \mathscr{W}(\mathcal{I}^{-1}\Psi_0)\sum_{a=1}^{3}\partial_a\Psi_a + \mathcal{I}^{-1}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\sum_{a=1}^{3}(\Psi_a)^2 - \mathscr{W}(\mathcal{I}^{-1}\Psi_0)\sum_{a=1}^{3}\mathring{\Psi}_a\Psi_a, \tag{2B.6a}$$

$$\partial_t \Psi_i = \partial_i\Psi_0 - \mathring{\Psi}_i\Psi_0. \tag{2B.6b}$$

Proof. We first prove (2B.6a). From (1B.1a) and (2B.1), we deduce

$$\partial_t(\mathcal{I}\partial_t\Phi) = \mathcal{I}\mathscr{W}(\partial_t\Phi)\Delta\Phi = \mathscr{W}(\partial_t\Phi)\sum_{a=1}^{3}\partial_a(\mathcal{I}\partial_a\Phi) - \mathscr{W}(\partial_t\Phi)\sum_{a=1}^{3}(\partial_a\mathcal{I})\partial_a\Phi.$$

Using (2B.4) to substitute for $\partial_a\mathcal{I}$ and appealing to Definition 2.3, we arrive at the desired (2B.6a).

To prove (2B.6b), we first use Definition 2.3 and the symmetry property $\partial_t\partial_i\Phi = \partial_i\partial_t\Phi$ to obtain $\partial_t\Psi_i = (\partial_t \ln\mathcal{I})\Psi_i + \partial_i\Psi_0 - (\partial_i\ln\mathcal{I})\Psi_0$. Using (2B.1) to replace $\partial_t\ln\mathcal{I}$ with $-\mathcal{I}^{-1}\Psi_0$ and (2B.4) to replace $-\partial_i\ln\mathcal{I}$ with $\mathcal{I}^{-1}\Psi_i - \mathring{\Psi}_i$, we arrive at (2B.6b). □

3. Assumptions on the initial data and bootstrap assumptions

In this section, we state our size assumptions on the initial data for the wave equation (1B.1a), i.e., for $(\partial_t\Phi|_{\Sigma_0}, \partial_1\Phi|_{\Sigma_0}, \partial_2\Phi|_{\Sigma_0}, \partial_3\Phi|_{\Sigma_0}) = (\mathring{\Psi}_0, \mathring{\Psi}_1, \mathring{\Psi}_2, \mathring{\Psi}_3)$, and formulate bootstrap assumptions that are convenient for studying the solution. We also precisely describe the smallness assumptions that we need to close our estimates and show the existence of initial data that satisfy the smallness assumptions.

3A. ***Assumptions on the initial data.*** We assume that the initial data are compactly supported and satisfy the following size assumptions for $i = 1, 2, 3$:

$$\|\nabla^{\leq 2}\mathring{\Psi}_i\|_{L^\infty(\Sigma_0)} + \|\nabla^{[1,3]}\mathring{\Psi}_0\|_{L^\infty(\Sigma_0)} + \|\mathring{\Psi}_i\|_{H^5(\Sigma_0)} + \mathring{\epsilon}^{3/2}\|\nabla\mathring{\Psi}_0\|_{L^2(\Sigma_0)} + \|\nabla^2\mathring{\Psi}_0\|_{H^3(\Sigma_0)} \leq \mathring{\epsilon}, \tag{3A.1a}$$

$$\|\mathring{\Psi}_0\|_{L^\infty(\Sigma_0)} \leq \mathring{A}, \tag{3A.1b}$$

$$-\tfrac{1}{4} \leq \min_{\Sigma_0}\mathring{\Psi}_0, \tag{3A.1c}$$

where $\mathring{\epsilon} > 0$ and $\mathring{A} > 0$ are two data-size parameters that we will discuss below (roughly, $\mathring{\epsilon}$ will have to be small for our proofs to close). Roughly, in our analysis, we will propagate the above size assumptions during the solution's classical lifespan, except for the top-order spatial derivatives of Ψ_i; we are not able to control these top-order derivatives uniformly in the norm $\|\cdot\|_{L^2(\Sigma_t)}$ because of the presence of the weight $\mathscr{W}$ in our energy, which can go to 0 as the singularity forms (see Definition 4.2).

We now introduce the crucial parameter $\mathring{A}_*$ that controls the time of first blowup; our analysis shows that, for $\mathring{\epsilon}$ sufficiently small, the time of first blowup is $\{1 + \mathcal{O}(\mathring{\epsilon})\}\mathring{A}_*^{-1}$; see also Remark 3.2.

Definition 3.1 (the parameter that controls the time of first blowup). We define the data-dependent parameter $\mathring{A}_*$ as follows:

$$\mathring{A}_* := \max_{\Sigma_0}[\mathring{\Psi}_0]_+, \tag{3A.2}$$

where $[\mathring{\Psi}_0]_+ := \max\{\mathring{\Psi}_0, 0\}$.

Our main results concern solutions such that $\mathring{A}_* > 0$, so we will assume in the rest of the article that this is the case.

Remark 3.2 (the relevance of $\mathring{A}_*$). The solutions that we study are such that[24] $\partial_t \mathcal{I} = -\Psi_0$ and $\partial_t \Psi_0 \sim 0$ (throughout the evolution). Hence, by the fundamental theorem of calculus, we have $\Psi_0(t, \underline{x}) \sim \mathring{\Psi}_0(\underline{x})$ and $\mathcal{I}(t, \underline{x}) \sim 1 - t\mathring{\Psi}_0(\underline{x})$. From this last expression, we see that $\mathcal{I}$ is expected to vanish for the first time at approximately $t = \mathring{A}_*^{-1}$ which, since $\partial_t \mathcal{I} = -\mathcal{I}\partial_t \Phi$, implies the blowup of $\partial_t \Phi$ (see Remark 2.2). See Lemmas 5.1 and 5.2 for the precise statements.

3B. *Bootstrap assumptions.* To prove our main results, we find it convenient to rely on a set of bootstrap assumptions, which we provide in this subsection.

- *The size of $T_{(\text{Boot})}$.* We assume that $T_{(\text{Boot})}$ is a bootstrap time with

$$0 < T_{(\text{Boot})} \leq 2\mathring{A}_*^{-1}, \tag{3B.1}$$

where $\mathring{A} > 0$ is the data-size parameter from Definition 3.1. The assumption (3B.1) gives us a sufficient margin of error to prove that finite-time blowup occurs (see Remark 3.2).

- *Blowup has not yet occurred.* Recall that, for the solutions under study, the vanishing of $\mathcal{I}$ will coincide with the formation of a singularity in $\partial_t \Phi$. For this reason, we assume that, for $t \in [0, T_{(\text{Boot})})$, we have

$$\mathcal{I}_\star(t) > 0, \tag{3B.2}$$

where $\mathcal{I}_\star$ is defined in (2B.2).

- *The solution is contained in the regime of hyperbolicity.*[25] We assume that, for $(t, \underline{x}) \in [0, T_{(\text{Boot})}) \times \mathbb{R}^3$,

$$\frac{\Psi_0(t, \underline{x})}{\mathcal{I}(t, \underline{x})} > -\frac{1}{2}. \tag{3B.3}$$

- *Smallness of $\mathcal{I}$ implies largeness of Ψ_0.* We assume that, for $(t, \underline{x}) \in [0, T_{(\text{Boot})}) \times \mathbb{R}^3$,

$$\mathcal{I}(t, \underline{x}) \leq \tfrac{1}{8} \quad \Longrightarrow \quad \Psi_0(t, \underline{x}) \geq \tfrac{1}{8}\mathring{A}_*. \tag{3B.4}$$

- *L^∞ bootstrap assumptions.* We assume that, for $t \in [0, T_{(\text{Boot})})$, we have

$$\|\Psi_0\|_{L^\infty(\Sigma_t)} \leq \mathring{A} + \varepsilon, \tag{3B.5a}$$

$$\|\nabla^{[1,3]}\Psi_0\|_{L^\infty(\Sigma_t)} \leq \varepsilon, \tag{3B.5b}$$

[24] Here "$A \sim B$" imprecisely indicates that A is well-approximated by B.

[25] In particular, the assumptions of Section 2A guarantee that $\mathscr{W}(\mathcal{I}^{-1}\Psi_0) > 0$ whenever (3B.3) holds. This inequality is needed in order to guarantee that (1B.1a) is a wave equation.

$$\|\nabla^{\leq 2}\Psi_i\|_{L^\infty(\Sigma_t)} \leq \varepsilon, \tag{3B.5c}$$

$$\|\mathcal{I}\|_{L^\infty(\Sigma_t)} \leq 1 + 2\mathring{A}_*^{-1}\mathring{A} + \varepsilon, \tag{3B.5d}$$

where $\varepsilon > 0$ is a small bootstrap parameter; we describe our smallness assumptions in the next subsection.

Remark 3.3 (the solution remains compactly supported in space). From the bootstrap assumptions and the assumptions of Section 2A on $\mathscr{W}$, we see that the wave speed $\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}^{1/2}$ associated to (1B.1a) remains uniformly bounded from above by a positive constant on the slab $(t, \underline{x}) \in [0, T_{(\text{Boot})}) \times \mathbb{R}^3$. It follows that there exists a large, data-dependent ball $B \subset \mathbb{R}^3$ such that $\Psi_\alpha(t, \underline{x})$ and $\mathcal{I} - 1$ vanish for $(t, \underline{x}) \in [0, T_{(\text{Boot})}) \times B^c$.

3C. *Smallness assumptions.* For the rest of the article, when we say that "A is small relative to B", we mean that $B > 0$ and that there exists a continuous increasing function $f : (0, \infty) \to (0, \infty)$ such that $A < f(B)$. For brevity, we typically do not specify the form of f.

In the rest of the article, we make the following relative smallness assumptions. We continually adjust the required smallness in order to close the estimates.

- The bootstrap parameter ε from Section 3B is small relative to 1 (i.e., in an absolute sense, without regard for the other parameters).
- ε is small relative to $\mathring{A}^{-1}$, where $\mathring{A}$ is the data-size parameter from (3A.1b).
- ε is small relative to the data-size parameter $\mathring{A}_*$ from (3A.2).
- We assume that

$$\varepsilon^{4/3} \leq \mathring{\epsilon} \leq \varepsilon, \tag{3C.1}$$

where $\mathring{\epsilon}$ is the data-smallness parameter from (3A.1a).

The first two of the above assumptions will allow us to control error terms that, roughly speaking, are of size $\varepsilon\mathring{A}^k$ for some integer $k \geq 0$. The third assumption is relevant because the expected blowup-time is approximately $\mathring{A}_*^{-1}$ (see Remark 3.2); the assumption will allow us to show that various error products, specifically ones featuring a small factor ε, remain small for $t \leq 2\mathring{A}_*^{-1}$, which is plenty of time for us to show that $\mathcal{I}$ vanishes and $\partial_t\Phi$ blows up. (3C.1) is convenient for closing our bootstrap argument.

3D. *Existence of initial data satisfying the smallness assumptions.* It is easy to construct initial data such that the parameters $\mathring{\epsilon}$, $\mathring{A}$, and $\mathring{A}_*$ satisfy the size assumptions stated in Section 3C. For example, we can start with *any* smooth compactly supported data $(\mathring{\Psi}_0, \mathring{\Psi}_1, \mathring{\Psi}_2, \mathring{\Psi}_3)$ such that $\max_{\Sigma_0} \mathring{\Psi}_0 > 0$ and $-\frac{1}{4} \leq \min_{\Sigma_0} \mathring{\Psi}_0$. We then consider the one-parameter family (for $i = 1, 2, 3$)

$$({}^{(\lambda)}\mathring{\Psi}_0(\underline{x}), {}^{(\lambda)}\mathring{\Psi}_i(\underline{x})) := (\mathring{\Psi}_0(\lambda^{-1}\underline{x}), \lambda^{-1}\mathring{\Psi}_i(\underline{x})).$$

It is straightforward to check that, for $\lambda > 0$ sufficiently large, all of the size assumptions of Section 3C are satisfied by the rescaled data (where, roughly speaking, the role of $\mathring{\epsilon}$ is played by λ^{-1}), as is (3A.1c).

The proof relies on the simple scaling identities

$$\nabla^{k\,(\lambda)}\mathring{\Psi}_0(\underline{x}) = \lambda^{-k}(\nabla^k\mathring{\Psi}_0)(\lambda^{-1}\underline{x}),$$
$$\nabla^{k\,(\lambda)}\mathring{\Psi}_i(\underline{x}) = \lambda^{-1}(\nabla^k\mathring{\Psi}_i)(\underline{x}),$$
$$\|\nabla^{k\,(\lambda)}\mathring{\Psi}_0\|_{L^2(\Sigma_0)} = \lambda^{3/2-k}\|\nabla^k\mathring{\Psi}_0\|_{L^2(\Sigma_0)},$$
$$\|\nabla^{k\,(\lambda)}\mathring{\Psi}_i\|_{L^2(\Sigma_0)} = \lambda^{-1}\|\nabla^k\mathring{\Psi}_i\|_{L^2(\Sigma_0)}.$$

4. Energy identities

In this section, we define the energies that we use to control the solution in L^2 up to top order. We then derive energy identities.

4A. *Definitions.*

The following energy functional serves as a building block for our energies.

Definition 4.1 (basic energy functional). To any array-valued function $V = V(t,\underline{x}) := (V_0, V_1, V_2, V_3)$, we associate the following energy, where $\mathcal{I}$ is as in Definition 2.1 and Ψ_0 is as in Definition 2.3:

$$\mathbb{E}[V] = \mathbb{E}[V](t) := \int_{\Sigma_t}\left\{V_0^2 + \sum_{a=1}^{3}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)(V_a)^2\right\}d\underline{x}. \tag{4A.1}$$

We now define $\mathbb{Q}_{(\mathring{\epsilon})}(t)$, which is the main L^2-type quantity that we use to control the solution up to top order.

Definition 4.2 (the L^2-controlling quantity). Let $\mathring{\epsilon} > 0$ be the data-size parameter from Section 3A. We define the L^2-controlling quantity $\mathbb{Q}_{(\mathring{\epsilon})}$ as follows:

$$\mathbb{Q}_{(\mathring{\epsilon})}(t) := \sum_{k=2}^{5}\int_{\Sigma_t}\left\{|\nabla^k\Psi_0|^2 + \sum_{a=1}^{3}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)|\nabla^k\Psi_a|^2\right\}d\underline{x} + \sum_{k=1}^{4}\int_{\Sigma_t}|\nabla^k\Psi_a|^2\,d\underline{x} + \mathring{\epsilon}^3\int_{\Sigma_t}\left\{|\nabla\Psi_0|^2 + \sum_{a=1}^{3}(\Psi_a)^2\right\}d\underline{x}. \tag{4A.2}$$

Remark 4.3 (the $\mathring{\epsilon}$-weight in the definition of $\mathbb{Q}_{(\mathring{\epsilon})}$). Our main a priori energy estimate shows that $\mathbb{Q}_{(\mathring{\epsilon})}(t) \lesssim \mathring{\epsilon}^2$ up to the singularity. The small coefficient of $\mathring{\epsilon}^3$ in front of the last integral on the right-hand side of (4A.2) is needed to ensure the $\mathcal{O}(\mathring{\epsilon}^2)$ smallness of $\mathbb{Q}_{(\mathring{\epsilon})}$. However, the small coefficient of $\mathring{\epsilon}^3$ implies that $\mathbb{Q}_{(\mathring{\epsilon})}(t)$ provides only weak L^2-control of $\nabla\Psi_0$ and Ψ_a; i.e., their L^2 norms can be as large as $\mathcal{O}(\mathring{\epsilon}^{-1/2})$. We clarify that the possible $\mathcal{O}(\mathring{\epsilon}^{-1/2})$-largeness of $\nabla\Psi_0$ is consistent with the construction of initial data described in Section 3D, where the largeness comes from the scaling identity

$$\|\nabla^{(\lambda)}\mathring{\Psi}_0\|_{L^2(\Sigma_0)} = \lambda^{1/2}\|\nabla\mathring{\Psi}_0\|_{L^2(\Sigma_0)},$$

and the large parameter λ can be viewed, roughly, as a size-$\mathcal{O}(\mathring{\epsilon}^{-1})$ quantity. Despite the possible $\mathcal{O}(\mathring{\epsilon}^{-1/2})$-largeness of $\nabla\Psi_0$ and Ψ_a in the norm $\|\cdot\|_{L^2(\Sigma_t)}$, we will nonetheless be able to show, through a separate argument, the following crucial bounds: $\nabla\Psi_0$ and Ψ_a are bounded in the norm $\|\cdot\|_{L^\infty(\Sigma_t)}$ by $\lesssim \mathring{\epsilon}$, up to the singularity; see Proposition 5.8.

4B. *Basic energy identity.* We aim to derive an energy identity for the L^2-controlling quantity $\mathbb{Q}_{(\mathring{\epsilon})}$ defined in (4A.2). As a preliminary step, in this subsection, we derive a standard energy identity for the building-block energy defined in (4A.1).

Lemma 4.4 (basic energy identity). *Let $\mathcal{I}$ be as in Definition 2.1, and assume that $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ is a solution to (2B.6a)–(2B.6b) with initial data $\{\mathring{\Psi}_\alpha\}_{\alpha=0,1,2,3}$. Let $\mathbb{E}[V](t)$ be the building-block energy defined in (4A.1). Let $\{F_\alpha\}_{\alpha=0,1,2,3}$ be scalar functions. Then for spatially compactly supported solutions $V := (V_0, V_1, V_2, V_3)$ to the inhomogeneous linear PDE system*

$$\partial_t V_0 = \sum_{a=1}^{3} \mathscr{W}(\mathcal{I}^{-1}\Psi_0)\partial_a V_a + F_0, \tag{4B.1a}$$

$$\partial_t V_i = \partial_i V_0 + F_i, \tag{4B.1b}$$

the following energy identity holds:

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}[V](t) = {} & \sum_{a=1}^{3}\int_{\Sigma_t} (\mathcal{I}^{-1}\Psi_0)^2 \mathscr{W}'(\mathcal{I}^{-1}\Psi_0)(V_a)^2 \, d\underline{x} \\
& + \sum_{a=1}^{3}\sum_{b=1}^{3}\int_{\Sigma_t} \mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)(\partial_a\Psi_a)(V_b)^2 \, d\underline{x} \\
& + \sum_{a=1}^{3}\sum_{b=1}^{3}\int_{\Sigma_t} \mathcal{I}^{-2}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)(\Psi_a)^2(V_b)^2 \, d\underline{x} \\
& - \sum_{a=1}^{3}\sum_{b=1}^{3}\int_{\Sigma_t} \mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\mathring{\Psi}_a\Psi_a(V_b)^2 \, d\underline{x} \\
& - 2\sum_{a=1}^{3}\int_{\Sigma_t} \mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)(\partial_a\Psi_0)V_a V_0 \, d\underline{x} \\
& - 2\sum_{a=1}^{3}\int_{\Sigma_t} \mathcal{I}^{-2}\Psi_0\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\Psi_a V_a V_0 \, d\underline{x} \\
& + 2\sum_{a=1}^{3}\int_{\Sigma_t} \mathcal{I}^{-1}\Psi_0\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathring{\Psi}_a V_a V_0 \, d\underline{x} \\
& + 2\int_{\Sigma_t} V_0 F_0 \, d\underline{x} + 2\sum_{a=1}^{3}\int_{\Sigma_t} \mathscr{W}(\mathcal{I}^{-1}\Psi_0) V_a F_a \, d\underline{x}.
\end{aligned} \tag{4B.2}$$

Proof. First, using (2B.1) and (2B.6a), we compute that

$$\begin{aligned}
\partial_t\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\} &= \mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)(\partial_t\Psi_0) + (\mathcal{I}^{-1}\Psi_0)^2\mathscr{W}'(\mathcal{I}^{-1}\Psi_0) \\
&= \sum_{a=1}^{3}\mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)(\partial_a\Psi_a) + \sum_{a=1}^{3}\mathcal{I}^{-2}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)(\Psi_a)^2 \\
&\quad - \sum_{a=1}^{3}\mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\mathring{\Psi}_a\Psi_a + (\mathcal{I}^{-1}\Psi_0)^2\mathscr{W}'(\mathcal{I}^{-1}\Psi_0).
\end{aligned} \tag{4B.3}$$

Next, taking the time derivative of (4A.1), using (4B.3), and using (4B.1a)–(4B.1b) to substitute for $\partial_t V_\alpha$, we obtain

$$
\begin{aligned}
\frac{d}{dt}\mathbb{E}[V](t) &= 2\sum_{a=1}^{3}\int_{\Sigma_t}\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)V_0\partial_a V_a + \mathscr{W}(\mathcal{I}^{-1}\Psi_0)V_a\partial_a V_0\}\,d\underline{x} \\
&\quad + 2\int_{\Sigma_t}\left\{V_0F_0 + \sum_{a=1}^{3}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)V_aF_a\right\}d\underline{x} \\
&\quad + \sum_{a=1}^{3}\sum_{b=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)(\partial_a\Psi_a)(V_b)^2\,d\underline{x} \\
&\quad + \sum_{a=1}^{3}\sum_{b=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-2}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)(\Psi_a)^2(V_b)^2\,d\underline{x} \\
&\quad - \sum_{a=1}^{3}\sum_{b=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\mathring{\Psi}_a\Psi_a(V_b)^2\,d\underline{x} \\
&\quad + \sum_{a=1}^{3}\int_{\Sigma_t}(\mathcal{I}^{-1}\Psi_0)^2\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)(V_a)^2\,d\underline{x}.
\end{aligned}
\tag{4B.4}
$$

Integrating by parts in the first integral on the right-hand side of (4B.4) and using the identity (2B.4), we obtain

$$
\begin{aligned}
&2\sum_{a=1}^{3}\int_{\Sigma_t}\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)V_0\partial_a V_a + \mathscr{W}(\mathcal{I}^{-1}\Psi_0)V_a\partial_a V_0\}\,d\underline{x} \\
&\quad = -2\sum_{a=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)(\partial_a\Psi_0)V_aV_0\,d\underline{x} \\
&\qquad - 2\sum_{a=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-2}\Psi_0\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\Psi_aV_aV_0\,d\underline{x} + 2\sum_{a=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-1}\Psi_0\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathring{\Psi}_aV_aV_0\,d\underline{x}.
\end{aligned}
\tag{4B.5}
$$

Using (4B.5) to substitute for the first integral on the right-hand side of (4B.4), we arrive at (4B.2). □

4C. *Integral identity for the fundamental L^2-controlling quantity.* With the help of Lemma 4.4, we now derive an energy identity for the controlling quantity $\mathbb{Q}_{(\mathring{\epsilon})}$.

Lemma 4.5 (integral identity for the L^2-controlling quantity). *Consider the following inhomogeneous PDE system, obtained by commuting (2B.6a)–(2B.6b) with the k-th-order spatial derivative operator ∇^k (where $\mathcal{I}$ is as in Definition 2.1 and the precise form of the inhomogeneous terms $F_\alpha^{(k)}$ in (4C.1a)–(4C.1b) is not important for the purposes of this lemma)*:

$$
\partial_t\nabla^k\Psi_0 = \mathscr{W}(\mathcal{I}^{-1}\Psi_0)\sum_{a=1}^{3}\partial_a\nabla^k\Psi_a + F_0^{(k)}, \tag{4C.1a}
$$

$$
\partial_t\nabla^k\Psi_i = \partial_i\nabla^k\Psi_0 + F_i^{(k)}. \tag{4C.1b}
$$

Then for solutions $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ *to* (2B.6a)–(2B.6b), *the* L^2*-controlling quantity* $\mathbb{Q}_{(\mathring{\epsilon})}$ *of Definition 4.2 satisfies the following integral identity*:

$$\begin{aligned}
\mathbb{Q}_{(\mathring{\epsilon})}(t) = \mathbb{Q}_{(\mathring{\epsilon})}(0) &+ \sum_{k=2}^{5}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}(\mathcal{I}^{-1}\Psi_0)^2\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|\nabla^k\Psi_a|^2\,d\underline{x}\,ds \\
&+\sum_{k=2}^{5}\sum_{a=1}^{3}\sum_{b=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)(\partial_a\Psi_a)|\nabla^k\Psi_b|^2\,d\underline{x}\,ds \\
&+\sum_{k=2}^{5}\sum_{a=1}^{3}\sum_{b=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathcal{I}^{-2}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)(\Psi_a)^2|\nabla^k\Psi_b|^2\,d\underline{x}\,ds \\
&-\sum_{k=2}^{5}\sum_{a=1}^{3}\sum_{b=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\mathring{\Psi}_a\Psi_a|\nabla^k\Psi_b|^2\,d\underline{x}\,ds \\
&-2\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)(\partial_a\Psi_0)\nabla^k\Psi_a\cdot\nabla^k\Psi_0\,d\underline{x}\,ds \\
&-2\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathcal{I}^{-2}\Psi_0\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\Psi_a\nabla^k\Psi_a\cdot\nabla^k\Psi_0\,d\underline{x}\,ds \\
&+2\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathcal{I}^{-1}\Psi_0\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathring{\Psi}_a\nabla^k\Psi_a\cdot\nabla^k\Psi_0\,d\underline{x}\,ds \\
&+2\sum_{k=1}^{4}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\nabla^k\Psi_a\cdot\partial_a\nabla^k\Psi_0\,d\underline{x}\,ds \\
&+2\sum_{k=2}^{5}\int_{s=0}^{t}\int_{\Sigma_s}\nabla^k\Psi_0\cdot F_0^{(k)}\,d\underline{x}\,ds+2\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\nabla^k\Psi_a\cdot F_a^{(k)}\,d\underline{x}\,ds \\
&+2\sum_{k=1}^{4}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\nabla^k\Psi_a\cdot F_a^{(k)}\,d\underline{x}\,ds \\
&+2\mathring{\epsilon}^3\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\nabla\Psi_0\cdot\partial_a\nabla\Psi_a\,d\underline{x}\,ds+2\mathring{\epsilon}^3\int_{s=0}^{t}\int_{\Sigma_s}\nabla\Psi_0\cdot F_0^{(1)}\,d\underline{x}\,ds \\
&+2\mathring{\epsilon}^3\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\Psi_a\partial_a\Psi_0\,d\underline{x}\,ds-2\mathring{\epsilon}^3\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\Psi_0\Psi_a\mathring{\Psi}_a\,d\underline{x}\,ds.
\end{aligned}\tag{4C.2}$$

Proof. We take the time derivative of both sides of (4A.2). The time derivative of the first line of the right-hand side of (4A.2) is given by (4B.2), where the role of (V_0, V_1, V_2, V_3) in (4B.2) is played by $(\nabla^k\Psi_0, \nabla^k\Psi_1, \nabla^k\Psi_2, \nabla^k\Psi_3)$ and the role of the inhomogeneous terms F_α on the right-hand side of (4B.2) is played by the terms $F_\alpha^{(k)}$ from (4C.1a)–(4C.1b). Moreover, with the help of (2B.6b) and (4C.1a)–(4C.1b), we compute that the time derivatives of the terms on the second line of the right-hand side of (4A.2) are equal to

$$2\sum_{k=1}^{4}\sum_{a=1}^{3}\int_{\Sigma_t}\nabla^k\Psi_a\cdot\partial_a\nabla^k\Psi_0\,d\underline{x}+2\sum_{k=1}^{4}\sum_{a=1}^{3}\int_{\Sigma_t}\nabla^k\Psi_a\cdot F_a^{(k)}\,d\underline{x}$$
$$+2\mathring{\epsilon}^3\sum_{a=1}^{3}\int_{\Sigma_t}\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\nabla\Psi_0\cdot\partial_a\nabla\Psi_a+\nabla\Psi_0\cdot F_0^{(1)}+\Psi_a\partial_a\Psi_0-\Psi_0\Psi_a\mathring{\Psi}_a\}\,d\underline{x}. \quad (4C.3)$$

Combining these calculations, we deduce that

$$\begin{aligned}\frac{d}{dt}\mathbb{Q}_{(\mathring{\epsilon})}(t)&=\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{\Sigma_t}(\mathcal{I}^{-1}\Psi_0)^2\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|\nabla^k\Psi_a|^2\,d\underline{x}\\
&+\sum_{k=2}^{5}\sum_{a=1}^{3}\sum_{b=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)(\partial_a\Psi_a)|\nabla^k\Psi_b|^2\,d\underline{x}\\
&+\sum_{k=2}^{5}\sum_{a=1}^{3}\sum_{b=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-2}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)(\Psi_a)^2|\nabla^k\Psi_b|^2\,d\underline{x}\\
&-\sum_{k=2}^{5}\sum_{a=1}^{3}\sum_{b=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\mathring{\Psi}_a\Psi_a|\nabla^k\Psi_b|^2\,d\underline{x}\\
&-2\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-1}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)(\partial_a\Psi_0)\nabla^k\Psi_a\cdot\nabla^k\Psi_0\,d\underline{x}\\
&-2\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-2}\Psi_0\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\Psi_a\nabla^k\Psi_a\cdot\nabla^k\Psi_0\,d\underline{x}\\
&+2\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{\Sigma_t}\mathcal{I}^{-1}\Psi_0\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)\mathring{\Psi}_a\nabla^k\Psi_a\cdot\nabla^k\Psi_0\,d\underline{x}\\
&+2\sum_{k=2}^{5}\int_{\Sigma_t}\nabla^k\Psi_0\cdot F_0^{(k)}\,d\underline{x}+2\sum_{a=1}^{3}\int_{\Sigma_t}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\nabla^k\Psi_a\cdot F_a^{(k)}\,d\underline{x}+(4C.3).\end{aligned} \quad (4C.4)$$

Integrating (4C.4) from time 0 to time t, we arrive at the desired identity (4C.2). □

5. A priori estimates

In this section, we use the data-size and bootstrap assumptions of Section 3 and the energy identities of Section 4 to derive a priori estimates for solutions to (2B.1) and to the renormalized equations of Proposition 2.5.

5A. ***Conventions for constants.*** In our estimates, the explicit constants $C > 0$ and $c > 0$ are free to vary from line to line. *These explicit constants, and implicit ones as well, are allowed to depend on the data-size parameters* $\mathring{A}$ *and* $\mathring{A}_*^{-1}$ *from* Section 3A. However, the constants can be chosen to be independent of the parameters $\mathring{\epsilon}$ and ε whenever $\mathring{\epsilon}$ and ε are sufficiently small relative to $\mathring{A}^{-1}$ and $\mathring{A}_*$ in the sense described in Section 3C. For example, under our conventions, we have $\mathring{A}_*^{-2}\varepsilon = \mathcal{O}(\varepsilon)$.

5B. *Pointwise estimates tied to the integrating factor.* In this subsection, we derive pointwise estimates that are important for analyzing $\mathcal{I}$.

We start by deriving sharp estimates for Ψ_0. The proof is based on separately considering regions where $\mathcal{I}$ is small and $\mathcal{I}$ is large. In Lemma 5.2, we will use these estimates to derive further information about the behavior of Ψ_0 in regions where $\mathcal{I}$ is small (i.e., near the singularity), which is crucial for closing the energy estimates.

Lemma 5.1 (pointwise estimates for Ψ_0). *Under the data-size assumptions of Section 3A, the bootstrap assumptions of Section 3B, and the smallness assumptions of Section 3C, the following pointwise estimates hold for* $(t, \underline{x}) \in [0, T_{(\mathrm{Boot})}) \times \mathbb{R}^3$:

$$\Psi_0(t, \underline{x}) = \mathring{\Psi}_0(\underline{x}) + \mathcal{O}(\varepsilon), \tag{5B.1}$$

where $\mathring{\Psi}_0(\underline{x}) = \Psi_0(0, \underline{x})$.

In addition,

$$-\tfrac{5}{16} \leq \min_{\Sigma_t} \Psi_0. \tag{5B.2}$$

Proof. We first prove (5B.1). We will show that $|\partial_t \Psi_0(t, \underline{x})| \lesssim \varepsilon$. Then from this estimate and the fundamental theorem of calculus, we obtain the desired bound (5B.1). To prove the bound $|\partial_t \Psi_0(t, \underline{x})| \lesssim \varepsilon$, we first consider points $(t, \underline{x})$ such that $\mathcal{I}(t, \underline{x}) > \frac{1}{8}$. Then all factors of $\mathcal{I}^{-1}$ in the evolution equation (2B.6a) can be bounded by $\lesssim 1$. For this reason, the desired bound follows as a straightforward consequence of (2B.6a), the bootstrap assumptions, the data-size assumptions (3A.1a), and the assumptions of Section 2A on $\mathscr{W}$.

To finish the proof of (5B.1), it remains to show that $|\partial_t \Psi_0(t, \underline{x})| \lesssim \varepsilon$ at points $(t, \underline{x})$ such that $0 < \mathcal{I}(t, \underline{x}) \leq \frac{1}{8}$. From the bootstrap assumption (3B.4), we deduce that $1 \lesssim \Psi_0(t, \underline{x})$ at such points. From this bound, the bootstrap assumptions, the data-size assumptions (3A.1a), and the assumptions of Section 2A on $\mathscr{W}$, we deduce the following bound for some factors on the right-hand side of (2B.6a) at the spacetime points under consideration:

$$|\mathcal{I}^{-1}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)| = \Psi_0^{-1}|(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)| \lesssim |(\mathcal{I}^{-1}\Psi_0)\mathscr{W}(\mathcal{I}^{-1}\Psi_0)| \lesssim 1.$$

With the help of this bound, the desired estimate $|\partial_t \Psi_0(t, \underline{x})| \lesssim \varepsilon$ follows as a straightforward consequence of (2B.6a), the bootstrap assumptions, the data-size assumptions (3A.1a), and the assumptions of Section 2A on $\mathscr{W}$. We have therefore proved (5B.1).

The bound (5B.2) then follows from (3A.1c) and (5B.1). □

In the next lemma, we derive sharp estimates for $\mathcal{I}$. The estimates are important for closing the energy estimates up to the singularity and for precisely tying the vanishing of $\mathcal{I}$ to the blowup of $\partial_t \Phi$.

Lemma 5.2 (crucial estimates for the integrating factor). *Under the data-size assumptions of Section 3A, the bootstrap assumptions of Section 3B, and the smallness assumptions of Section 3C, the following* *estimates hold for* $(t, \underline{x}) \in [0, T_{(\text{Boot})}) \times \mathbb{R}^3$:

$$\mathcal{I}(t, \underline{x}) = 1 - t\mathring{\Psi}_0(\underline{x}) + \mathcal{O}(\varepsilon), \tag{5B.3a}$$

$$\mathcal{I}_\star(t) = 1 - t\mathring{A}_* + \mathcal{O}(\varepsilon), \tag{5B.3b}$$

where $\mathring{\Psi}_0(\underline{x}) = \Psi_0(0, \underline{x})$, $\mathcal{I}_\star$ *is defined in* (2B.2), *and* $\mathring{A}_* > 0$ *is the data-size parameter from Definition 3.1.*

Moreover, the following implications hold for $(t, \underline{x}) \in [0, T_{(\text{Boot})}) \times \mathbb{R}^3$:

$$\mathcal{I}(t, \underline{x}) \le \tfrac{1}{4}\min\{1, \mathring{A}_*\} \quad \Longrightarrow \quad \frac{\Psi_0(t, \underline{x})}{\mathcal{I}(t, \underline{x})} \ge 1, \tag{5B.4a}$$

$$\mathcal{I}(t, \underline{x}) \le \tfrac{1}{4} \quad \Longrightarrow \quad \Psi_0(t, \underline{x}) \ge \tfrac{1}{4}\mathring{A}_*. \tag{5B.4b}$$

Finally, the following implications hold for $(t, \underline{x}) \in [0, T_{(\text{Boot})}) \times \mathbb{R}^3$:

$$\Psi(t, \underline{x}) \le 0 \implies \mathcal{I}(t, \underline{x}) \ge 1 - C\varepsilon \qquad \text{and} \qquad \Psi(t, \underline{x}) \le 0 \implies \frac{\Psi_0(t, \underline{x})}{\mathcal{I}(t, \underline{x})} \ge -\tfrac{3}{8}. \tag{5B.5}$$

Remark 5.3 (improvement of a bootstrap assumption). Note in particular that the estimate (5B.4b) provides a strict improvement of the bootstrap assumption (3B.4).

Remark 5.4 (the significance of (5B.5)). Note that (5B.5) is a strict improvement of the bootstrap assumption (3B.3) and that (5B.5) implies that $-\frac{3}{8} \le \partial_t \Phi(t, \underline{x})$ for $(t, \underline{x}) \in [0, T_{(\text{Boot})}) \times \mathbb{R}^3$. In view of the assumption (2A.1) for $\mathscr{W}$, we conclude that on $(t, \underline{x}) \in [0, T_{(\text{Boot})}) \times \mathbb{R}^3$, the wave speed $\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}^{1/2}$ is positive and uniformly bounded from above by a positive constant. In the rest of article, we often silently use this fact. Note, however, that from (2A.4) with $k = 0$, it follows that the wave speed vanishes when $\partial_t \Phi \to \infty$.

Proof. From (2B.1) and the estimate (5B.1), we deduce $\partial_t \mathcal{I}(t, \underline{x}) = -\mathring{\Psi}_0(\underline{x}) + \mathcal{O}(\varepsilon)$. Integrating in time and using the initial condition (2B.1), we find that $\mathcal{I}(t, \underline{x}) = 1 - t\mathring{\Psi}_0(\underline{x}) + \mathcal{O}(\varepsilon)$, which is (5B.3a).

Equation (5B.3b) follows as a simple consequence of (5B.3a) and definitions (2B.2) and (3A.2).

To prove (5B.4a), we first consider the case $\mathring{A}_* \ge 1$. From (5B.3a) and (5B.1), we deduce that $\mathcal{I}(t, \underline{x}) = 1 - t\Psi_0(t, \underline{x}) + \mathcal{O}(\varepsilon)$. It follows that if $\mathcal{I}(t, \underline{x}) \le \frac{1}{4}$, then $t\Psi_0(t, \underline{x}) \ge \frac{1}{2}$. Since $0 \le t \le 2\mathring{A}_*^{-1} \le 2$, we deduce that $\Psi_0(t, \underline{x})/\mathcal{I}(t, \underline{x}) \ge 1$, which is the desired conclusion. Next, we consider the case $\mathring{A}_* < 1$. Using (5B.3a) and (5B.1), we deduce that $\mathcal{I}(t, \underline{x}) = 1 - t\Psi_0(t, \underline{x}) + \mathcal{O}(\varepsilon)$. It follows that if $\mathcal{I}(t, \underline{x}) \le \frac{1}{4}\mathring{A}_*$, then $t\Psi_0(t, \underline{x}) \ge 1 - \frac{1}{2}\mathring{A}_*$. Since $0 \le t \le 2\mathring{A}_*^{-1}$, we deduce that

$$\frac{\Psi_0(t, \underline{x})}{\mathcal{I}(t, \underline{x})} \ge 2\left\{1 - \tfrac{1}{2}\mathring{A}_*\right\} = 2 - \mathring{A}_*,$$

which, in view of our assumption $\mathring{A}_* < 1$, is > 1. This completes our proof of (5B.4a).

The implication (5B.4b) can be proved using arguments similar to the ones that we used to prove (5B.4a), and we therefore omit the details.

Next, we note that when $\Psi_0(t, \underline{x}) \le 0$, the estimate $\mathcal{I}(t, \underline{x}) = 1 - t\Psi_0(t, \underline{x}) + \mathcal{O}(\varepsilon)$ proved above implies that $\mathcal{I}(t, \underline{x}) \ge 1 - C\varepsilon$, which yields the first implication stated in (5B.5). To obtain the second implication stated in (5B.5), we use the first implication and the estimate (5B.2). □

In the next lemma, we derive some simple pointwise estimates showing that the spatial derivatives of $\mathcal{I}$ up to top order can be controlled in terms of the spatial derivatives of $\{\Psi_a\}_{a=1,2,3}$.

Lemma 5.5 (estimates for the derivatives of the integrating factor). *Under the data-size assumptions of Section 3A, the bootstrap assumptions of Section 3B, and the smallness assumptions of Section 3C, the following pointwise estimates hold for* $(t,\underline{x}) \in [0, T_{(\text{Boot})}) \times \mathbb{R}^3$:

$$|\nabla \mathcal{I}| \lesssim \sum_{a=1}^{3} |\Psi_a| + \sum_{a=1}^{3} |\mathring{\Psi}_a|. \tag{5B.6a}$$

Moreover, for $2 \leq k \leq 6$, *the following estimate holds*:

$$|\nabla^k \mathcal{I}| \lesssim \sum_{a=1}^{3} |\nabla^{[1,k-1]}\Psi_a| + \sum_{a=1}^{3} |\nabla^{[1,k-1]}\mathring{\Psi}_a| + \varepsilon \sum_{a=1}^{3} |\mathring{\Psi}_a|. \tag{5B.6b}$$

Finally, the following estimate holds for $t \in [0, T_{(\text{Boot})})$:

$$\|\nabla^{[1,3]}\mathcal{I}\|_{L^\infty(\Sigma_t)} \lesssim \varepsilon. \tag{5B.7}$$

Proof. The estimate (5B.6a) is straightforward consequence of (2B.4) and the bootstrap assumptions. Similarly, the estimate (5B.6b) is straightforward to derive via induction in k with the help of (2B.4), the bootstrap assumptions, the data-size assumptions (3A.1a), and (3C.1). Inequality (5B.7) then follows from (5B.6a)–(5B.6b), the bootstrap assumptions, the data-size assumptions (3A.1a), and (3C.1). □

5C. *Pointwise estimates involving the weight.* In the next lemma, we derive precise pointwise estimates for quantities that involve the weight function $\mathscr{W}$. The detailed information is important for closing the energy estimates and for showing that the spatial derivatives of $\mathscr{W} = \mathscr{W}(\partial_t \Phi) = \mathscr{W}(\mathcal{I}^{-1}\Psi_0)$ are controllable. Some of the analysis is delicate in that $\partial_t \Phi$ and its derivatives are allowed to be arbitrarily large (i.e., the estimates hold uniformly, arbitrarily close to the singularity).

Lemma 5.6 (pointwise estimates involving the weight $\mathscr{W}$). *Let* $\mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}$ *be the characteristic function of the spacetime subset* $\{(t,\underline{x}) \mid 0 < \mathcal{I}(t,\underline{x}) \leq \frac{1}{4}\min\{1,\mathring{A}_*\}\}$. *Under the data-size assumptions of Section 3A, the bootstrap assumptions of Section 3B, and the smallness assumptions of Section 3C, the following pointwise estimates hold for* $(t,\underline{x}) \in [0, T_{(\text{Boot})}) \times \mathbb{R}^3$:

$$\mathscr{W}(\mathcal{I}^{-1}\Psi_0) \lesssim 1, \tag{5C.1a}$$

$$|\nabla\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}| \lesssim \varepsilon \mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}\{\mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|\}^{1/2} + \varepsilon\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}^{1/2} \tag{5C.1b}$$

$$\lesssim \varepsilon. \tag{5C.1c}$$

In addition, for $2 \leq k \leq 5$, *the following estimates hold*:

$$|\nabla^k\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}| \lesssim |\nabla^{[1,k]}\Psi_0| + \sum_{a=1}^{3} |\nabla^{\leq k-1}\Psi_a| + \sum_{a=1}^{3} |\nabla^{\leq k-1}\mathring{\Psi}_a|. \tag{5C.2}$$

Furthermore, the following estimates hold:

$$\mathcal{I}^{-1}\mathscr{W}(\mathcal{I}^{-1}\Psi_0) \lesssim \mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}\{\mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|\}^{1/2} + \{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}^{1/2} \tag{5C.3a}$$

$$\lesssim 1. \tag{5C.3b}$$

Moreover, for $1 \leq k \leq 5$, *the following estimates hold*:

$$|\nabla^k\{\mathcal{I}^{-1}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}| \lesssim |\nabla^{[1,k]}\Psi_0| + \sum_{a=1}^{3}|\nabla^{\leq k-1}\Psi_a| + \sum_{a=1}^{3}|\nabla^{\leq k-1}\mathring{\Psi}_a|. \tag{5C.4}$$

Finally, for $0 \leq P \leq 2$, *the following estimates hold*:

$$\left|\mathcal{I}^{-2}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0) + \mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}\mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|\right| \lesssim \mathscr{W}(\mathcal{I}^{-1}\Psi_0), \tag{5C.5a}$$

$$|\mathcal{I}^{-P}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)| \lesssim 1. \tag{5C.5b}$$

Proof. Throughout this proof, we set

$$y = y(t,\underline{x}) := \frac{\Psi_0(t,\underline{x})}{\mathcal{I}(t,\underline{x})}.$$

Also, we silently use the observations of Remark 5.4.

Proof of (5C.1a): This bound is a trivial consequence of our assumption (2A.4) on $\mathscr{W}$.

Proof of (5C.1b) and (5C.1c): We first prove (5C.1b) at spacetime points $(t,\underline{x})$ such that $\mathcal{I}(t,\underline{x}) > \frac{1}{4}\min\{1,\mathring{A}_*\}$. This is the easy case because $\mathcal{I}^{-1} < 4\max\{1,\mathring{A}_*^{-1}\} \leq C$, and we therefore do not have to concern ourselves with the possibility of small denominators. Specifically, using the identity (2B.4), the bootstrap assumptions, the data-size assumptions (3A.1a), and the assumptions of Section 2A, we deduce that when $\mathcal{I}(t,\underline{x}) > \frac{1}{4}\min\{1,\mathring{A}_*\}$, we have

$$|\nabla\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}| \lesssim |\nabla\Psi_0| + \sum_{a=1}^{3}|\Psi_a| + \sum_{a=1}^{3}|\mathring{\Psi}_a| \lesssim \varepsilon. \tag{5C.6}$$

Next, we use the bootstrap assumptions and the assumptions of Section 2A on $\mathscr{W}$, including the uniform positivity and boundedness of $\mathscr{W}(y)$ on intervals of the form $y \in \left[-\frac{3}{8}, C\right]$, to obtain

$$\mathbf{1}_{\{\mathcal{I}>(1/4)\min\{1,\mathring{A}_*\}\}} \lesssim \mathscr{W}(\mathcal{I}^{-1}\Psi_0) \lesssim \{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}^{1/2}.$$

It follows that the right-hand side of (5C.6) is $\lesssim$ the second term on the right-hand side of (5C.1b) as desired.

We now prove (5C.1b) at points $(t,\underline{x})$ such that $0 < \mathcal{I}(t,\underline{x}) \leq \frac{1}{4}\min\{1,\mathring{A}_*\}$. Along the way, we will prove some additional estimates that we will use later on. We start by defining the following weighted differential operator, which acts on functions $f = f(y)$: $D_Y f := y^2\frac{d}{dy}f$. Note that the chain rule implies

$$\nabla\mathscr{W}(y) = -D_Y\mathscr{W}(y)\nabla(y^{-1}). \tag{5C.7}$$

We therefore inductively deduce that, for $1 \leq k \leq 5$, we have

$$|\nabla^k \mathscr{W}(y)| \lesssim \sum_{n=1}^{k} |D_Y^n \mathscr{W}(y)| \Bigg\{ \sum_{\substack{\sum_{i=1}^n k_i = k \\ k_i \geq 1}} \prod_{i=1}^{n} |\nabla^{k_i}(y^{-1})| \Bigg\}. \tag{5C.8}$$

The case $k = 1$ in (5C.8) yields $|\nabla \mathscr{W}(\mathcal{I}^{-1}\Psi_0)| \lesssim (\mathcal{I}^{-1}\Psi_0)^2 |\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)| |\nabla\{\mathcal{I}\Psi_0^{-1}\}|$. Also using the identity (2B.4), the bootstrap assumptions, the data-size assumptions (3A.1a), (3C.1), the assumptions of Section 2A, and the crucially important estimate (5B.4b) (which implies that $\Psi_0^{-1} \lesssim 1$), we deduce that when $\mathcal{I}(t, \underline{x}) \leq \frac{1}{4}\min\{1, \mathring{A}_*\}$, we have $|\nabla\{\mathcal{I}\Psi_0^{-1}\}| \lesssim \varepsilon$ and thus

$$\begin{aligned} |\nabla \mathscr{W}(\mathcal{I}^{-1}\Psi_0)| &\lesssim \varepsilon (\mathcal{I}^{-1}\Psi_0)^2 |\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)| \lesssim \varepsilon \{(\mathcal{I}^{-1}\Psi_0)^2 |\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|\}^{1/2} \\ &\lesssim \varepsilon \{\mathcal{I}^{-2} |\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|\}^{1/2}, \end{aligned} \tag{5C.9}$$

which is $\lesssim$ the first term on the right-hand side of (5C.1b) as desired. This finishes the proof of (5C.1b). We clarify that to derive the next-to-last inequality in (5C.9), in which we bounded $(\mathcal{I}^{-1}\Psi_0)^2 |\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|$ by its square root, we used (2A.4) to deduce $y^2 |\mathscr{W}'(y)| \lesssim 1$.

We now prove (5C.1c). From Remark 5.4, the assumptions of Section 2A on $\mathscr{W}$, and (5B.4b), we deduce that

$$\begin{aligned} \mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}} \{\mathcal{I}^{-2} |\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|\}^{1/2} &\lesssim \mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}} \{(\mathcal{I}^{-2}\Psi_0^2) |\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|\}^{1/2} \\ &\lesssim 1 \end{aligned} \tag{5C.10}$$

and that $\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}^{1/2} \lesssim 1$. That is, the non-$\varepsilon$ factors on the right-hand side of (5C.1b) are $\lesssim 1$. This yields (5C.1c).

Proof of (5C.2): The proof is similar to that of (5C.1b), but slightly simpler. Note that $k \in [2, 5]$ by assumption in this estimate. We first prove the estimate at points $(t, \underline{x})$ such that $\mathcal{I}(t, \underline{x}) > \frac{1}{4}\min\{1, \mathring{A}_*\}$. This is the easy case because $\mathcal{I}^{-1} < 4\max\{1, \mathring{A}_*^{-1}\} \leq C$, and we therefore do not have to concern ourselves with the possibility of small denominators. Specifically, using the identity (2B.4), the bootstrap assumptions, the data-size assumptions (3A.1a), (3C.1), and the assumptions of Section 2A, we deduce that when $\mathcal{I}(t, \underline{x}) > \frac{1}{4}\min\{1, \mathring{A}_*\}$, we have

$$|\nabla^k\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}| \lesssim |\nabla^{[1,k]}\Psi_0| + \sum_{a=1}^{3} |\nabla^{\leq k-1}\Psi_a| + \sum_{a=1}^{3} |\nabla^{\leq k-1}\mathring{\Psi}_a|, \tag{5C.11}$$

which is $\lesssim$ the right-hand side of (5C.2) as desired.

It remains for us to prove (5C.2) at points $(t, \underline{x})$ such that $0 < \mathcal{I}(t, \underline{x}) \leq \frac{1}{4}\min\{1, \mathring{A}_*\}$. Note that the estimate (5C.8) holds and that by (2A.4) and Remark 5.4, we have the following bound[26] for the factors of $D_Y^n \mathscr{W}(y)$ on the right-hand side of (5C.8): $|D_Y^n \mathscr{W}(y)| \lesssim 1$. From this bound, (5C.8), the bootstrap assumptions, and the data-size assumptions (3A.1a), we see that the desired bound (5C.2) will follow

[26]In obtaining this bound, it is helpful to note that $D_Y f = -\frac{d}{dz} f$, where $z := 1/y$.

once we show that the following bound holds when $2 \leq k \leq 5$ and $\mathcal{I}(t, \underline{x}) \leq \frac{1}{4}\min\{1, \mathring{A}_*\}$:

$$|\nabla^k(y^{-1})| \lesssim |\nabla^{[1,k]}\Psi_0| + \sum_{a=1}^{3} |\nabla^{\leq k-1}\Psi_a| + \sum_{a=1}^{3} |\nabla^{\leq k-1}\mathring{\Psi}_a|. \tag{5C.12}$$

To prove (5C.12), we first note that (5B.4b) implies that $1 \lesssim \Psi_0(t, \underline{x})$ in the present context. Thus, the left-hand side of (5C.11), which is equal to $|\nabla^k(\mathcal{I}/\Psi_0)|$, is the k-th derivative of a ratio with a denominator uniformly bounded from below away from 0, and the desired estimate (5C.2) follows as a straightforward consequence of the identity (2B.4), the data-size assumptions (3A.1a), and the bootstrap assumptions.

Proof of (5C.3a), (5C.3b), and (5C.4): These estimates can be proved using arguments similar to the ones we used to prove (5C.1b) and (5C.2), based on separately considering the cases $\mathcal{I}(t, \underline{x}) > \frac{1}{4}\min\{1, \mathring{A}_*\}$ and $0 < \mathcal{I}(t, \underline{x}) \leq \frac{1}{4}\min\{1, \mathring{A}_*\}$ and using the assumptions of Section 2A. We omit the details, noting only that we can write $\mathcal{I}^{-1}\mathscr{W}(\mathcal{I}^{-1}\Psi_0) = \Psi_0^{-1} y\mathscr{W}(y)$ and that the assumptions of Section 2A (especially (2A.5)), (5B.4a), (5B.4b), and Remark 5.4 imply that we have the following key estimates, relevant for the more difficult case $0 < \mathcal{I}(t, \underline{x}) \leq \frac{1}{4}\min\{1, \mathring{A}_*\}$:

$$\begin{aligned}\mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}\{\Psi_0^{-1} y\mathscr{W}(y)\} &\lesssim \mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}\{y^2|\mathscr{W}'(y)|\}^{1/2} \\ &\lesssim \mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}\{\mathcal{I}^{-2}|\mathscr{W}'(y)|\}^{1/2}\end{aligned} \tag{5C.13}$$

and, for $n \leq 5$, $|D_Y^n(y\mathscr{W}(y))| \lesssim 1$ (footnote 26 is also relevant for obtaining this latter bound).

Proof of (5C.5a): We first note that by (2A.6) and (5B.4a), we have $\mathscr{W}'(\mathcal{I}^{-1}\Psi_0) < 0$ at points $(t, \underline{x})$ such that $\mathcal{I}(t, \underline{x}) \leq \frac{1}{4}\min\{1, \mathring{A}_*\}$. From this fact and the identity $1 = \mathbf{1}_{\{\mathcal{I}>(1/4)\min\{1,\mathring{A}_*\}\}} + \mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}$, it follows that

$$\begin{aligned}\text{LHS (5C.5a)} &= |\mathbf{1}_{\{\mathcal{I}>(1/4)\min\{1,\mathring{A}_*\}\}}\mathcal{I}^{-2}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)| \\ &\lesssim \mathbf{1}_{\{\mathcal{I}>(1/4)\min\{1,\mathring{A}_*\}\}}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|.\end{aligned} \tag{5C.14}$$

Also using the bound $|\mathscr{W}'(y)| \lesssim 1$, which is a simple consequence of (2A.4), we find that the left-hand side of (5C.5a) is $\lesssim \mathbf{1}_{\{\mathcal{I}>(1/4)\min\{1,\mathring{A}_*\}\}}$. Next, we recall the estimate $\mathbf{1}_{\{\mathcal{I}>(1/4)\min\{1,\mathring{A}_*\}\}} \lesssim \mathscr{W}(\mathcal{I}^{-1}\Psi_0)$ that we derived in our proof of (5C.1b). Combining the above estimates, we conclude the desired bound (5C.5a).

Proof of (5C.5b): We first prove (5C.5b) at points $(t, \underline{x})$ such that $\mathcal{I}(t, \underline{x}) > \frac{1}{4}\min\{1, \mathring{A}_*\}$. Using the bootstrap assumptions and the assumptions of Section 2A on $\mathscr{W}$, we deduce, in view of Remark 5.4, that $|\mathcal{I}^{-P}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)| \lesssim |\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)| \lesssim 1$ as desired.

It remains for us to prove (5C.5b) at points $(t, \underline{x})$ such that $0 < \mathcal{I}(t, \underline{x}) \leq \frac{1}{4}\min\{1, \mathring{A}_*\}$. Using (5B.4b), we see that $1 \lesssim \Psi_0(t, \underline{x})$ at such points, and it follows that $|\mathcal{I}^{-P}\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)| \lesssim |\{\mathcal{I}^{-1}\Psi_0\}^P\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|$. Using the assumptions of Section 2A on $\mathscr{W}$ and the assumption $P \in [0, 2]$, we deduce, in view of Remark 5.4, that the right-hand side of the previous expression is $\lesssim 1$ as desired. This finishes the proof of (5C.5b) and completes the proof of the lemma. □

5D. ***Pointwise estimates for the inhomogeneous terms in the commuted evolution equations.*** With the estimates of Lemma 5.6 in hand, we are now ready to derive pointwise estimates for the inhomogeneous terms in the ∇^k-commuted evolution equations.

Lemma 5.7 (pointwise estimates for the inhomogeneous terms in the ∇^k-commuted evolution equations). *Let $\mathcal{I}$ be as in Definition 2.1, and let $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ be a solution to the system (2B.6a)–(2B.6b). Consider the following inhomogeneous PDE system,*[27] *obtained by commuting (2B.6a)–(2B.6b) with ∇^k:*

$$\partial_t \nabla^k \Psi_0 = \mathscr{W}(\mathcal{I}^{-1}\Psi_0)\sum_{a=1}^{3}\partial_a \nabla^k \Psi_a + F_0^{(k)}, \tag{5D.1a}$$

$$\partial_t \nabla^k \Psi_i = \partial_i \nabla^k \Psi_0 + F_i^{(k)}. \tag{5D.1b}$$

Under the data-size assumptions of Section 3A, the bootstrap assumptions of Section 3B, and the smallness assumptions of Section 3C, for $k = 2, 3, 4, 5$ and $(t, \underline{x}) \in [0, T_{(\mathrm{Boot})}) \times \mathbb{R}^3$, the following estimate holds:

$$\begin{aligned}|F_0^{(k)}| \lesssim{}& \varepsilon|\nabla^{[2,k]}\Psi_0| + \varepsilon \mathbf{1}_{\{0<\mathcal{I}\le(1/4)\min\{1,\mathring{A}_*\}\}}\sum_{a=1}^{3}\{\mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|\}^{1/2}|\nabla^k\Psi_a| \\ &+ \varepsilon\sum_{a=1}^{3}\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}^{1/2}|\nabla^k\Psi_a| + \sum_{a=1}^{3}|\nabla^{[1,k-1]}\Psi_a| + \varepsilon^2\sum_{a=1}^{3}|\Psi_a| + \sum_{a=1}^{3}|\nabla^{\le k}\mathring{\Psi}_a|. \end{aligned}\tag{5D.2}$$

Moreover, for $k = 0, 1, 2, 3, 4$, the following estimate holds:

$$|F_0^{(k)}| \lesssim \underbrace{\varepsilon|\nabla^{[2,k]}\Psi_0|}_{\text{absent if } k\le 1} + \underbrace{\sum_{a=1}^{3}|\nabla^{[1,k]}\Psi_a|}_{\text{absent if } k=0} + \underbrace{\varepsilon^2\sum_{a=1}^{3}|\Psi_a|}_{\text{absent if } k=0} + \underbrace{\varepsilon\sum_{a=1}^{3}|\Psi_a|}_{\text{absent if } k\ge 1} + \sum_{a=1}^{3}|\nabla^{\le k}\mathring{\Psi}_a|. \tag{5D.3}$$

Finally, for $k = 0, 1, 2, 3, 4, 5$, the following estimate holds:

$$\sum_{a=1}^{3}|F_a^{(k)}| \lesssim \underbrace{\varepsilon|\nabla^{[2,k]}\Psi_0|}_{\text{absent if } k=0,1} + \sum_{a=1}^{3}|\nabla^{\le k}\mathring{\Psi}_a|. \tag{5D.4}$$

Proof. The estimate (5D.4) follows in a straightforward fashion from commuting (2B.6b) with ∇^k and using the bootstrap assumptions, the data-size assumptions (3A.1a), and (3C.1).

To prove (5D.2), we first commute (2B.6a) with ∇^k to obtain (5D.1a). The only products in $F_0^{(k)}$ that are difficult to bound are those that feature a factor in which k total derivatives fall on Ψ_a, specifically the products $\sum_{a=1}^{3}\{\nabla[\mathscr{W}(\mathcal{I}^{-1}\Psi_0)]\}\partial_a\nabla^{k-1}\Psi_a$, $\sum_{a=1}^{3}\mathcal{I}^{-1}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\Psi_a\nabla^k\Psi_a$, and $\sum_{a=1}^{3}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\mathring{\Psi}_a\nabla^k\Psi_a$. To bound the first of these, we use the estimate (5C.1b), which implies that the product is bounded by the sum of the second and third terms on the right-hand side of (5D.2) as desired. To handle the second and third products, we use (5C.1a), (5C.3a), the bootstrap assumptions, the data-size assumptions (3A.1a), and (3C.1) to bound them in magnitude by

$$\lesssim \varepsilon \mathbf{1}_{\{0<\mathcal{I}\le(1/4)\min\{1,\mathring{A}_*\}\}}\sum_{a=1}^{3}\{\mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)|\}^{1/2}|\nabla^k\Psi_a| + \varepsilon\sum_{a=1}^{3}\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}^{1/2}|\nabla^k\Psi_a|,$$

which is in turn bounded by the sum of the second and third terms on the right-hand side of (5D.2) as desired. The remaining terms in $F_0^{(k)}$ feature $\le k-1$ derivatives of Ψ_a. These terms can easily seen to

[27] We do not bother to state the precise form of the inhomogeneous terms $F_\alpha^{(k)}$ here.

be $\lesssim$ the right-hand side of (5D.2) with the help of the estimates (5C.1a), (5C.1c), (5C.2), (5C.3b), and (5C.4), the bootstrap assumptions, the data-size assumptions (3A.1a), and (3C.1).

The estimate (5D.3) is easier to prove and can be obtained in a similar fashion with the help of the estimates (5C.1a), (5C.1c), (5C.2), (5C.3b), (5C.4), the bootstrap assumptions, the data-size assumptions (3A.1a), and (3C.1). □

5E. *The main a priori estimates.* We now derive the main result of this section: a priori estimates that hold up to top order and that in particular yield a strict improvement of the bootstrap assumptions. These are the main ingredients in the proof of our main theorem.

Proposition 5.8 (the main a priori estimates). *Let $\mathcal{I}$ be the integrating factor from Definition 2.1, and let $\mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}$ be the characteristic function of the spacetime subset*

$$\left\{(t,\underline{x})\;\middle|\;0<\mathcal{I}(t,\underline{x})\leq\tfrac{1}{4}\min\{1,\mathring{A}_*\}\right\}.$$

There exists a constant $C>0$ such that under the data-size assumptions of Section 3A, the bootstrap assumptions of Section 3B, and the smallness assumptions of Section 3C, for solutions $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ to the system (2B.6a)–(2B.6b), the L^2-controlling quantity $\mathbb{Q}_{(\mathring{\epsilon})}$ of Definition 4.2 satisfies the following estimate for $t\in[0,T_{(\text{Boot})})$:

$$\mathbb{Q}_{(\mathring{\epsilon})}(t)+\frac{1}{20}\mathring{A}_*^2\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}\mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)||\nabla^k\Psi_a|^2\,d\underline{x}\,ds\leq C\mathring{\epsilon}^2. \tag{5E.1}$$

In addition the following estimates hold for $t\in[0,T_{(\text{Boot})})$ and $i=1,2,3$:

$$\mathring{\epsilon}\|\partial_t\Psi_0\|_{L^2(\Sigma_t)}^2+\|\nabla\partial_t\Psi_0\|_{H^3(\Sigma_t)}^2\leq C\mathring{\epsilon}^2, \tag{5E.2a}$$

$$\mathring{\epsilon}^3\|\partial_t\Psi_i\|_{L^2(\Sigma_t)}^2+\|\nabla\partial_t\Psi_i\|_{H^3(\Sigma_t)}^2\leq C\mathring{\epsilon}^2. \tag{5E.2b}$$

Moreover, $\mathcal{I}$ satisfies the following estimate for $t\in[0,T_{(\text{Boot})})$:

$$\mathring{\epsilon}^3\|\nabla\mathcal{I}\|_{L^2(\Sigma_t)}^2+\|\nabla^{[2,5]}\mathcal{I}\|_{L^2(\Sigma_t)}^2+\int_{s=0}^{t}\int_{\Sigma_s}\mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}\mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)||\nabla^6\mathcal{I}|^2\,d\underline{x}\,ds\leq C\mathring{\epsilon}^2. \tag{5E.3}$$

Finally, we have the following estimates for $t\in[0,T_{(\text{Boot})})$, which in particular yield strict improvements of the bootstrap assumptions (3B.5a)–(3B.5d) whenever $C\mathring{\epsilon}<\varepsilon$:

$$\|\Psi_0\|_{L^\infty(\Sigma_t)}\leq\mathring{A}+C\mathring{\epsilon}, \tag{5E.4a}$$

$$\|\nabla^{[1,3]}\Psi_0\|_{L^\infty(\Sigma_t)}\leq C\mathring{\epsilon}, \tag{5E.4b}$$

$$\|\nabla^{\leq 2}\Psi_i\|_{L^\infty(\Sigma_t)}\leq C\mathring{\epsilon}, \tag{5E.4c}$$

$$\|\mathcal{I}\|_{L^\infty(\Sigma_t)}\leq 1+2\mathring{A}_*^{-1}\mathring{A}+C\mathring{\epsilon}, \tag{5E.4d}$$

$$\|\nabla^{[1,3]}\mathcal{I}\|_{L^\infty(\Sigma_t)}\leq C\mathring{\epsilon}. \tag{5E.4e}$$

Proof. Proof of (5E.1): The main step is to derive the following estimate:

$$\mathbb{Q}_{(\mathring{\epsilon})}(t) + \tfrac{1}{16}\mathring{A}_*^2 \sum_{k=2}^{5}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s} \mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}\mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)||\nabla^k\Psi_a|^2\,d\underline{x}\,ds$$
$$\leq C\mathring{\epsilon}^2 + C\varepsilon\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}\mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)||\nabla^k\Psi_a|^2\,d\underline{x}\,ds$$
$$+ C\int_{s=0}^{t}\mathbb{Q}_{(\mathring{\epsilon})}(s)\,ds. \quad (5E.5)$$

Once we have shown (5E.5), we can absorb the second term on the right-hand side of (5E.5) into the second term on the left-hand side of (5E.5), which, for ε sufficiently small, at most reduces the coefficient of $\frac{1}{16}\mathring{A}_*^2$ in front of the second term on the left to the value of $\frac{1}{20}\mathring{A}_*^2$, as is stated on the left-hand side of (5E.1). We then use Gronwall's inequality and the assumption $0 < t < T_{(\mathrm{Boot})} \leq 2\mathring{A}_*^{-1}$ to conclude that the left-hand side of (5E.1) is $\leq C\exp(Ct)\mathring{\epsilon}^2 \leq C\exp(C\mathring{A}_*^{-1})\mathring{\epsilon}^2 \leq C\mathring{\epsilon}^2$ as desired.

To prove (5E.5), we must bound the terms on the right-hand side of (4C.2). As a first step, we note the following bound for the first term on the right-hand side of (4C.2): $\mathbb{Q}_{(\mathring{\epsilon})}(0) \leq C\mathring{\epsilon}^2$, an estimate that follows as a straightforward consequence of definition (4A.2), the data-size assumptions (3A.1a)–(3A.1c), the initial condition $\mathcal{I}|_{\Sigma_0} = 1$ stated in (2B.1), and the assumptions of Section 2A on $\mathscr{W}$.

Next, we treat the spacetime integral on the first line of the right-hand side of (4C.2). Using (5B.4b), (5C.5a), and the bootstrap assumption (3B.5a) for $\|\Psi_0\|_{L^\infty(\Sigma_t)}$, we can express the integral as the negative integral

$$-\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}(\mathcal{I}^{-1}\Psi_0)^2|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)||\nabla^k\Psi_a|^2\,d\underline{x}\,ds,$$

which is bounded from above by the negative "favorable integral"

$$-\tfrac{1}{16}\mathring{A}_*^2\sum_{k=2}^{5}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}}\mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)||\nabla^k\Psi_a|^2\,d\underline{x}\,ds,$$

plus an error integral that is bounded in magnitude by

$$\lesssim \sum_{k=2}^{5}\sum_{a=1}^{3}\int_{s=0}^{t}\int_{\Sigma_s}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)|\nabla^k\Psi_a|^2\,d\underline{x}\,ds.$$

We can therefore bring the favorable integral over to the left-hand side of (5E.5), where it appears with a "+" sign. Moreover, from Definition 4.2, we deduce that the error integral is bounded by the last term on the right-hand side of (5E.5) as desired.

We now bound the spacetime integrals on lines two to four of the right-hand side of (4C.2). Using the estimate (5C.5b), the bootstrap assumptions, the data-size assumptions (3A.1a), and (3C.1), we deduce that all three integrands are bounded in magnitude by $\lesssim \sum_{k=2}^{5}\sum_{a=1}^{3}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)|\nabla^k\Psi_a|^2$. From Definition 4.2, we conclude that the corresponding error integrals are bounded by the last term on the right-hand side of (5E.5) as desired. Using similar reasoning and Young's inequality, we bound the last two spacetime integrals on the right-hand side of (4C.2) by $\leq$ the right-hand side of (5E.5) as desired.

We now bound the spacetime integrals on lines five to seven of the right-hand side of (4C.2). Using the estimates (5C.5a) and (5C.5b), the bootstrap assumptions, the data-size assumptions (3A.1a), (3C.1), and Young's inequality, we deduce that all three integrands are bounded in magnitude by

$$\lesssim \varepsilon \mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}} \sum_{k=2}^{5}\sum_{a=1}^{3} \mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)||\nabla^k\Psi_a|^2 + \varepsilon \sum_{k=2}^{5}\sum_{a=1}^{3} \mathscr{W}(\mathcal{I}^{-1}\Psi_0)|\nabla^k\Psi_a|^2 + \sum_{k=2}^{5}|\nabla^k\Psi_0|^2.$$

Appealing to Definition 4.2, we conclude that the corresponding error integrals are bounded in magnitude by $\leq$ the right-hand side of (5E.5) as desired.

We now bound the spacetime integral on line eight of the right-hand side of (4C.2), in which the integrand is $2\sum_{k=1}^{4}\sum_{a=1}^{3}\nabla^k\Psi_a\cdot\partial_a\nabla^k\Psi_0$. Using Young's inequality, we bound this integrand by $\lesssim |\nabla^{[2,5]}\Psi_0|^2+\sum_{a=1}^{3}|\nabla^{[1,4]}\Psi_a|^2$. From Definition 4.2, we conclude that the integral of the right-hand side of this expression over the spacetime domain $(s,\underline{x})\in[0,t]\times\mathbb{R}^3$ is bounded by the last term on the right-hand side of (5E.5) as desired.

We now bound the first spacetime integral on line nine of the right-hand side of (4C.2), in which the integrand is $2\sum_{k=2}^{5}\nabla^k\Psi_0\cdot F_0^{(k)}$. Using Young's inequality, (5D.2), and (3C.1), we pointwise bound this integrand in magnitude by

$$\lesssim |\nabla^{[2,5]}\Psi_0|^2 + \varepsilon \mathbf{1}_{\{0<\mathcal{I}\leq(1/4)\min\{1,\mathring{A}_*\}\}} \sum_{a=1}^{3} \mathcal{I}^{-2}|\mathscr{W}'(\mathcal{I}^{-1}\Psi_0)||\nabla^{[2,5]}\Psi_a|^2 \\ + \sum_{a=1}^{3}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)|\nabla^{[2,5]}\Psi_a|^2 + \sum_{a=1}^{3}|\nabla^{[1,4]}\Psi_a|^2 + \mathring{\epsilon}^3\sum_{a=1}^{3}|\Psi_a|^2 + \sum_{a=1}^{3}|\nabla^{\leq 5}\mathring{\Psi}_a|^2. \quad (5E.6)$$

From Definition 4.2 and the data-size assumptions (3A.1a), we conclude that the integral of the right-hand side of (5E.6) over the spacetime domain $(s,\underline{x})\in[0,t]\times\mathbb{R}^3$ is $\leq$ the right-hand side of (5E.5) as desired.

We now bound the second spacetime integral on line nine of the right-hand side of (4C.2), in which the integrand is $2\sum_{k=2}^{5}\sum_{a=1}^{3}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\nabla^k\Psi_a\cdot F_a^{(k)}$. Using Young's inequality, (5C.1a), and (5D.4), we pointwise bound this integrand in magnitude by $\lesssim |\nabla^{[2,5]}\Psi_0|^2+\sum_{a=1}^{3}\mathscr{W}(\mathcal{I}^{-1}\Psi_0)|\nabla^{[2,5]}\Psi_a|^2+\sum_{a=1}^{3}|\nabla^{\leq 5}\mathring{\Psi}_a|^2$. From Definition 4.2 and the data-size assumptions (3A.1a), we conclude that the integral of the right-hand side of this expression over the spacetime domain $(s,\underline{x})\in[0,t]\times\mathbb{R}^3$ is $\leq$ the right-hand side of (5E.5) as desired.

We now bound the spacetime integral on line ten of the right-hand side of (4C.2), in which the integrand is $2\sum_{k=1}^{4}\sum_{a=1}^{3}\nabla^k\Psi_a\cdot F_a^{(k)}$. Using Young's inequality and (5D.4), we pointwise bound this integrand in magnitude by $\lesssim |\nabla^{[2,4]}\Psi_0|^2+\sum_{a=1}^{3}|\nabla^{[1,4]}\Psi_a|^2+\sum_{a=1}^{3}|\nabla^{\leq 4}\mathring{\Psi}_a|^2$. From Definition 4.2 and the data-size assumptions (3A.1a), we conclude that the integral of the right-hand side of this expression over the spacetime domain $(s,\underline{x})\in[0,t]\times\mathbb{R}^3$ is $\leq$ the right-hand side of (5E.5) as desired.

We now bound the first spacetime integral on line eleven of the right-hand side of (4C.2), in which the integrand is $2\mathring{\epsilon}^3\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\nabla\Psi_0\cdot\sum_{a=1}^{3}\partial_a\nabla\Psi_a$. Using the estimate (5C.1a) and Young's inequality, we bound this integrand by $\lesssim \mathring{\epsilon}^3|\nabla\Psi_0|^2+\sum_{a=1}^{3}|\nabla^2\Psi_a|^2$. From Definition 4.2, we conclude that the integral of the right-hand side of this expression over the spacetime domain $(s,\underline{x})\in[0,t]\times\mathbb{R}^3$ is bounded by the last term on the right-hand side of (5E.5) as desired.

Finally, we bound the second spacetime integral on line eleven of the right-hand side of (4C.2), in which the integrand is $2\mathring{\epsilon}^3 \nabla\Psi_0 \cdot F_0^{(1)}$. Using Young's inequality and (5D.3), we pointwise bound this integrand in magnitude by

$$\lesssim \mathring{\epsilon}^3 |\nabla\Psi_0|^2 + \sum_{a=1}^{3} |\nabla\Psi_a|^2 + \mathring{\epsilon}^3 \sum_{a=1}^{3} |\Psi_a|^2 + \sum_{a=1}^{3} |\nabla^{\leq 1}\mathring{\Psi}_a|^2. \tag{5E.7}$$

From Definition 4.2 and the data-size assumptions (3A.1a), we conclude that the integral of the right-hand side of (5E.7) over the spacetime domain $(s, \underline{x}) \in [0, t] \times \mathbb{R}^3$ is $\leq$ the right-hand side of (5E.5) as desired. This completes our proof of (5E.5) and therefore finishes the proof of (5E.1).

Proof of (5E.4b) and (5E.4c): In view of Definition 4.2, we see that the estimates $\|\nabla^{[2,3]}\Psi_0\|_{L^\infty(\Sigma_t)} \lesssim \mathring{\epsilon}$ and $\|\nabla^{[1,2]}\Psi_i\|_{L^\infty(\Sigma_t)} \lesssim \mathring{\epsilon}$ follow from (5E.1) and Sobolev embedding $H^2(\mathbb{R}^3) \hookrightarrow L^\infty(\mathbb{R}^3)$. To bound $\|\nabla\Psi_0\|_{L^\infty(\Sigma_t)}$, we first use (5D.1a), the bootstrap assumptions, the data-size assumptions (3A.1a), the estimates (5C.1a) and (5D.3), inequality (3C.1), and the already-proven bound $\|\nabla^{[1,2]}\Psi_i\|_{L^\infty(\Sigma_t)} \lesssim \mathring{\epsilon}$ to obtain $|\partial_t \nabla\Psi_0| \lesssim \varepsilon^2 + \mathring{\epsilon} + \sum_{a=1}^{3} |\nabla^{[1,2]}\Psi_a| \lesssim \mathring{\epsilon}$. From this bound, the fundamental theorem of calculus, and the data-size assumptions (3A.1a), we find that $|\nabla\Psi_0| \lesssim \mathring{\epsilon} + \int_{s=0}^{t} \mathring{\epsilon}\, ds \lesssim \mathring{\epsilon}$. This implies $\|\nabla\Psi_0\|_{L^\infty(\Sigma_t)} \lesssim \mathring{\epsilon}$, which completes the proof of (5E.4b). Similarly, from (2B.6b), the bootstrap assumptions, the data-size assumptions (3A.1a), and the already-proven bound $\|\nabla\Psi_0\|_{L^\infty(\Sigma_t)} \lesssim \mathring{\epsilon}$, we deduce $\sum_{a=1}^{3} |\partial_t \Psi_a| \lesssim \mathring{\epsilon}$. From this bound, the fundamental theorem of calculus, and the data-size assumption (3A.1a), we find that $\sum_{a=1}^{3} |\Psi_a| \lesssim \mathring{\epsilon}$, which implies $\sum_{a=1}^{3} \|\Psi_a\|_{L^\infty(\Sigma_t)} \lesssim \mathring{\epsilon}$, thereby completing the proof of (5E.4c).

Proof of (5E.4a): We first use (2B.6a), the estimates (5C.1a) and (5C.3b), the bootstrap assumptions, the data-size assumptions (3A.1a), and the already-proven bound $\|\nabla^{\leq 1}\Psi_i\|_{L^\infty(\Sigma_t)} \lesssim \mathring{\epsilon}$ to obtain $|\partial_t \Psi_0| \lesssim \mathring{\epsilon}$. From this bound, the fundamental theorem of calculus, the data-size assumption (3A.1b), and the fact that $0 < t \leq 2\mathring{A}_*^{-1}$, we find that $\|\Psi_0\|_{L^\infty(\Sigma_t)} \leq \|\mathring{\Psi}_0\|_{L^\infty(\Sigma_0)} + C\mathring{\epsilon} \leq \mathring{A} + C\mathring{\epsilon}$, which is the desired bound (5E.4a).

Proof of (5E.4d) and (5E.4e): We repeat the proof of (5B.3a), but using the bootstrap assumption (3B.5d) and the estimates (5E.4a)–(5E.4c) instead of using the full set of bootstrap assumptions. We find that $\mathcal{I}(t, \underline{x}) = 1 - t\mathring{\Psi}_0(\underline{x}) + \mathcal{O}(\mathring{\epsilon})$. From this estimate, the fact that $0 < t < 2\mathring{A}_*^{-1}$, and the data-size assumption (3A.1b), we conclude the desired bound (5E.4d). Similarly, to prove (5E.4e), we repeat the proof of (5B.7), but using the estimates (5E.4a)–(5E.4d) instead of the bootstrap assumptions.

Proof of (5E.3): The estimate (5E.3) follows as a straightforward consequence of the pointwise estimates (5B.6a)–(5B.6b), the weight estimate (5C.5b), the energy estimate (5E.1), and the data-size assumptions (3A.1a).

Proof of (5E.2a) and (5E.2b): To prove (5E.2a), we first use (5D.1a) and the estimate (5C.1a) to deduce that

$$\begin{aligned}
&\mathring{\epsilon}\|\partial_t\Psi_0\|_{L^2(\Sigma_t)}^2 + \|\nabla\partial_t\Psi_0\|_{H^3(\Sigma_t)}^2 \\
&\lesssim \sum_{k=2}^{5}\sum_{a=1}^{3} \|\{\mathscr{W}(\mathcal{I}^{-1}\Psi_0)\}^{1/2}\nabla^k\Psi_a\|_{L^2(\Sigma_t)}^2 + \sum_{a=1}^{3}\|\nabla\Psi_a\|_{L^2(\Sigma_t)}^2 + \mathring{\epsilon}\|F_0^{(0)}\|_{L^2(\Sigma_t)}^2 + \sum_{k=1}^{4}\|F_0^{(k)}\|_{L^2(\Sigma_t)}^2.
\end{aligned} \tag{5E.8}$$

Next, we recall that the already-proven estimates (5E.4a)–(5E.4d) imply that the bootstrap assumptions (3B.5a)–(3B.5d) hold with $C\mathring{\epsilon}$ in place of ε. It follows that the pointwise estimate (5D.3) holds with $C\mathring{\epsilon}$ in place of ε. From this fact, Definition 4.2, the energy estimate (5E.1), and the data-size assumptions (3A.1a), we deduce that the right-hand side of (5E.8) is $\lesssim \mathring{\epsilon}^2$, which is the desired bound (5E.2a).

The estimate (5E.2b) can be proved using similar arguments based on the evolution equation (5D.1b) and the pointwise estimate (5D.4), and we omit the details. □

6. Local well-posedness and continuation criteria

In this section, we provide a proposition that yields standard well-posedness results and continuation criteria pertaining to the quantities $\{\partial_\alpha \Phi\}_{\alpha=0,1,2,3}$, $\mathcal{I}$, and $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$.

Proposition 6.1. *Let $N \geq 3$ be an integer and let $(\partial_t \Phi|_{\Sigma_0}, \partial_1 \Phi|_{\Sigma_0}, \partial_2 \Phi|_{\Sigma_0}, \partial_3 \Phi|_{\Sigma_0}) = (\mathring{\Psi}_0, \mathring{\Psi}_1, \mathring{\Psi}_2, \mathring{\Psi}_3)$ be initial data for the wave equation (1B.1a) satisfying $\mathring{\Psi}_\alpha \in H^N(\mathbb{R}^3)$, $\alpha = 0, 1, 2, 3$, and with $\partial_i \mathring{\Psi}_j = \partial_j \mathring{\Psi}_i$ for $i, j = 1, 2, 3$ (see Remark 1.1). Let $\mathcal{H} := \left(-\frac{1}{2}, \infty\right)$ denote the regime of hyperbolicity, and note that the following holds: (1B.1a) is a nondegenerate[28] wave equation at points $(t, \underline{x})$ such that $\partial_t \Phi(t, \underline{x}) \in \mathcal{H}$ (see (2A.1) for justification of this assertion). Assume that $\mathring{\Psi}_0(\mathbb{R}^3)$ is contained in a compact subset $\mathfrak{K}$ of $\mathcal{H}$. Let $\mathcal{I}$, $\mathcal{I}_\star$, and $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ be the quantities defined in Definitions 2.1 and 2.3. Then there exist a compact set $\mathfrak{K}'$ of $\mathcal{H}$ containing $\mathfrak{K}$ in its interior and a time $T > 0$, depending on $\mathfrak{K}$ and $\sum_{\alpha=0}^{3} \|\mathring{\Psi}_\alpha\|_{H^N}$, such that a unique[29] classical solution to (1B.1a) exists on $[0, T) \times \mathbb{R}^3$, such that $\partial_t \Phi([0, T) \times \mathbb{R}^3) \subset \mathfrak{K}'$, and such that the following regularity properties hold for $\alpha = 0, 1, 2, 3$:*

$$\partial_\alpha \Phi \in C([0, T), H^N). \tag{6.1}$$

In addition, the solution depends continuously on the data.

Let $T_{(\text{Lifespan})}$ be the supremum of all times $T > 0$ such that the classical solution to (1B.1a) exists on $[0, T) \times \mathbb{R}^3$ and satisfies the above properties. Then either $T_{(\text{Lifespan})} = \infty$, or $T_{(\text{Lifespan})} < \infty$ and one of the following two breakdown scenarios must occur:

1. *There exists a sequence of points $\{(t_n, \underline{x}_n)\}_{n=1}^{\infty} \subset [0, T_{(\text{Lifespan})}) \times \mathbb{R}^3$ such that $\partial_t \Phi(t_n, \underline{x}_n)$ escapes every compact subset of $\mathcal{H}$ as $n \to \infty$.*
2. $\lim_{t \uparrow T_{(\text{Lifespan})}} \sup_{s \in [0,t)} \sum_{\alpha=0}^{3} \|\partial_\alpha \partial_t \Phi\|_{L^\infty(\Sigma_s)} = \infty$.

Moreover, on $[0, T_{(\text{Lifespan})}) \times \mathbb{R}^3$, $\mathcal{I}$ and $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ are classical solutions to (2B.1) and (2B.6a)–(2B.6b) such that

$$\mathcal{I} - 1 \in C([0, T_{(\text{Lifespan})}), H^{N+1}(\mathbb{R}^3)), \quad \Psi_\alpha \in C([0, T_{(\text{Lifespan})}), H^N(\mathbb{R}^3)). \tag{6.2}$$

Finally, the quantity $\mathcal{I}_\star$ defined in (2B.2) satisfies the following estimates:

$$0 < \mathcal{I}_\star(t) < \infty \quad \text{for } t \in [0, T_{(\text{Lifespan})}). \tag{6.3}$$

[28] By nondegenerate, we mean that relative to the Cartesian coordinates, the 4×4 matrix of components $g_{\alpha\beta}$ has signature $(-, +, +, +)$, where $g := -dt^2 + 1/(\mathscr{W}(\partial_t \Phi)) \sum_{a=1}^{3} (dx^a)^2$ is the metric corresponding to (1B.1a).

[29] More precisely, Φ is uniquely determined only up to a constant when only its first derivatives along Σ_0 are prescribed; see Remark 1.1.

Proof. The statements concerning Φ are standard and can be proved using the ideas found, for example, in [Speck 2008].

Next, we note that the evolution equation plus the initial condition for $\mathcal{I}$ stated in (2B.1), the fact that $\mathcal{I}(t,\cdot)-1$ is compactly supported in space (see Remark 3.3), and the fact that

$$\partial_t\Phi \in C([0,T_{(\text{Lifespan})}), H^N(\mathbb{R}^3)) \subset C([0,T_{(\text{Lifespan})}), C^1(\mathbb{R}^3))$$

(i.e., (6.1)) can be used to deduce (6.3). Similarly, from (2B.1), the identity (2B.4), the definition $\Psi_\alpha := \mathcal{I}\partial_\alpha\Phi$ (see Definition 2.3), (6.1), and the standard Sobolev–Moser calculus, it is straightforward to deduce (6.2). □

7. The main theorem

In this section, we state and prove our main stable blowup-result.

Theorem 7.1 (stable ODE-type blowup). *Assume that the weight function $\mathscr{W}$ appearing in the wave equation (1B.1a) satisfies the assumptions stated in Section 2A. Consider compactly supported initial data $(\partial_t\Phi|_{\Sigma_0}, \partial_1\Phi|_{\Sigma_0}, \partial_2\Phi|_{\Sigma_0}, \partial_3\Phi|_{\Sigma_0}) = (\mathring{\Psi}_0, \mathring{\Psi}_1, \mathring{\Psi}_2, \mathring{\Psi}_3)$ for the wave equation (1B.1a) with $\partial_i\mathring{\Psi}_j = \partial_j\mathring{\Psi}_i$ for $i, j = 1, 2, 3$ (see Remark 1.1) that satisfy the data-size assumptions (3A.1a)–(3A.1c) involving the parameters $\mathring{\epsilon}$ and $\mathring{A}$, and let $\mathring{A}_*$ be the data-size parameter defined in (3A.2). Let $\mathcal{I}$, $\mathcal{I}_\star$, and $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ be the quantities defined in Definitions 2.1 and 2.3. We define*

$$T_{(\text{Lifespan})} := \sup\{t>0 \mid \{\partial_\alpha\Phi\}_{\alpha=0,1,2,3} \textit{ exist classically on } [0,t)\times\mathbb{R}^3\}. \tag{7.1}$$

If $\mathring{\epsilon}>0$, $\mathring{A}>0$, and $\mathring{A}_*>0$, and if $\mathring{\epsilon}$ is small relative to $\mathring{A}^{-1}$ and $\mathring{A}_*$ in the sense explained in Section 3C, then the following conclusions hold.

- *Characterization of the solution's classical lifespan: The solution's classical lifespan is characterized by $\mathcal{I}_\star$ as follows:*

$$T_{(\text{Lifespan})} = \sup\Big\{t>0 \,\Big|\, \inf_{s\in[0,t)} \mathcal{I}_\star(s) > 0\Big\}. \tag{7.2}$$

Moreover,

$$\mathcal{I}(t,\underline{x}) > 0 \quad \textit{for } (t,\underline{x}) \in [0,T_{(\text{Lifespan})})\times\mathbb{R}^3, \tag{7.3a}$$

$$\lim_{t\uparrow T_{(\text{Lifespan})}} \mathcal{I}_\star(t) = 0. \tag{7.3b}$$

In addition, the following estimate holds:

$$T_{(\text{Lifespan})} = \mathring{A}_*^{-1}\{1+\mathcal{O}(\mathring{\epsilon})\}. \tag{7.4}$$

- *Regularity properties of Ψ_α and $\mathcal{I}$ on $[0,T_{(\text{Lifespan})})\times\mathbb{R}^3$: On the slab $[0,T_{(\text{Lifespan})})\times\mathbb{R}^3$, the solution satisfies the energy bounds (5E.1)–(5E.3), the L^∞ estimates (5E.4a)–(5E.4e), (5B.1)–(5B.2), and (5B.3a)–(5B.5) (with $C\mathring{\epsilon}$ on the right-hand side in place of ε in these equations). Moreover, $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ and $\mathcal{I}$ enjoy the following regularity:*

$$\Psi_0 \in C([0,T_{(\text{Lifespan})}), H^5(\mathbb{R}^3)) \cap L^\infty([0,T_{(\text{Lifespan})}), H^5(\mathbb{R}^3)), \tag{7.5a}$$

$$\Psi_i \in C([0,T_{(\text{Lifespan})}), H^5(\mathbb{R}^3)) \cap L^\infty([0,T_{(\text{Lifespan})}), H^4(\mathbb{R}^3)), \tag{7.5b}$$

$$\mathcal{I}-1 \in C([0,T_{(\text{Lifespan})}), H^6(\mathbb{R}^3)) \cap L^\infty([0,T_{(\text{Lifespan})}), H^5(\mathbb{R}^3)). \tag{7.5c}$$

- *Regularity properties of* Ψ_α *and* $\mathcal{I}$ *on* $[0, T_{(\text{Lifespan})}] \times \mathbb{R}^3$: Ψ_α *and* $\mathcal{I}$ *do not blow up at time* $T_{(\text{Lifespan})}$, *but rather continuously extend to* $[0, T_{(\text{Lifespan})}] \times \mathbb{R}^3$ *as functions that enjoy the following regularity, where* N *is any real number with* $N < 5$:

$$\Psi_0 \in L^\infty([0, T_{(\text{Lifespan})}], H^5(\mathbb{R}^3)) \cap C([0, T_{(\text{Lifespan})}], H^N(\mathbb{R}^3)), \tag{7.6a}$$

$$\Psi_i \in C([0, T_{(\text{Lifespan})}], H^4(\mathbb{R}^3)), \tag{7.6b}$$

$$\mathcal{I} - 1 \in C([0, T_{(\text{Lifespan})}], H^5(\mathbb{R}^3)). \tag{7.6c}$$

- *Description of the vanishing of* $\mathcal{I}$ *and the blowup of* $\partial_t \Phi$: *For* $(t, \underline{x}) \in [0, T_{(\text{Lifespan})}) \times \mathbb{R}^3$, *we have*

$$\mathcal{I}(t, \underline{x}) \leq \tfrac{1}{4} \quad \Longrightarrow \quad \partial_t \Phi(t, \underline{x}) \geq \frac{\mathring{A}_*}{4\mathcal{I}(t, \underline{x})}. \tag{7.7}$$

  Let

$$\Sigma^{\text{Blowup}}_{T_{(\text{Lifespan})}} := \{(T_{(\text{Lifespan})}, \underline{x}) \mid \mathcal{I}(T_{(\text{Lifespan})}, \underline{x}) = 0\}. \tag{7.8}$$

  Then if $(T_{(\text{Lifespan})}, \underline{x}) \in \Sigma^{\text{Blowup}}_{T_{(\text{Lifespan})}}$, *we have*[30]

$$\lim_{t \uparrow T_{(\text{Lifespan})}} \partial_t \Phi(t, \underline{x}) = \infty. \tag{7.9}$$

  Finally, if $(T_{(\text{Lifespan})}, \underline{x}) \notin \Sigma^{\text{Blowup}}_{T_{(\text{Lifespan})}}$, *then there exists an open ball* $B_{\underline{x}} \subset \mathbb{R}^3$ *centered at* $\underline{x}$ *such that, for* $\alpha = 0, 1, 2, 3$, *we have* $\partial_\alpha \Phi \in C([0, T_{(\text{Lifespan})}], H^5(B_{\underline{x}}))$.

Proof. Let $C_* > 1$ be a constant; we will enlarge C_* as needed throughout the proof. Let $T_{(\text{Max})}$ be the supremum of the set of real numbers T with $0 \leq T \leq 2\mathring{A}_*^{-1}$ such that the following properties hold:

- $\{\partial_\alpha \Phi\}_{\alpha=0,1,2,3}$ is a classical solution to (1B.1a) on $[0, T) \times \mathbb{R}^3$ (see Remark 1.1) satisfying the properties stated in Proposition 6.1 (with $N = 5$ in the proposition).
- $\mathcal{I}$ is a classical solution to (2B.1) on $[0, T) \times \mathbb{R}^3$ satisfying the properties stated in Proposition 6.1 (with $N = 5$ in the proposition).
- $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ are classical solutions to (2B.6a)–(2B.6b) for $(t, \underline{x}) \in [0, T) \times \mathbb{R}^3$ such that the properties stated in Proposition 6.1 hold (with $N = 5$ in the proposition).
- The bootstrap assumptions (3B.3) and (3B.4) hold for $(t, \underline{x}) \in [0, T) \times \mathbb{R}^3$.
- The L^∞ bootstrap assumptions (3B.5a)–(3B.5d) hold for $t \in [0, T)$ with $\varepsilon := C_* \mathring{\epsilon}$.
- $\inf\{\mathcal{I}_\star(t) \mid t \in [0, T)\} > 0$, where $\mathcal{I}_\star$ is defined in (2B.2). Note that this implies that the bootstrap assumption (3B.2) holds for $t \in [0, T)$.

Throughout the rest of the proof, we will assume that $\mathring{\epsilon}$ is sufficiently small and that C_* is sufficiently large without explicitly mentioning it every time. Next, we note that the hypotheses of Proposition 6.1 hold with $N = 5$. Hence, by Proposition 6.1 and Sobolev embedding, we have $T_{(\text{Max})} > 0$.

[30] See also Remark 1.5 concerning the blowup of Φ itself, if initial data for Φ itself are prescribed.

We will now show that $T_{(\text{Max})} = T_{(\text{Lifespan})}$, where $T_{(\text{Lifespan})}$ is defined by (7.1). We first note that clearly, $T_{(\text{Max})} \le T_{(\text{Lifespan})}$. To proceed, we assume for the sake of deriving a contradiction that

$$\inf_{s\in[0,T_{(\text{Max})})} \mathcal{I}_\star(s) > 0. \tag{7.10}$$

Then, in view of Definitions 2.1 and 2.3, (2B.6a)–(2B.6b), the bootstrap assumptions, the assumptions of Section 2A on $\mathscr{W}$, the data-size assumptions (3A.1a), and (3C.1), we see that (7.10) implies

$$\lim_{t\uparrow T_{(\text{Max})}} \sup_{s\in[0,t)} \left\{ \|\partial_t\Phi\|_{L^\infty(\Sigma_s)} + \sum_{\alpha=0}^{3} \|\partial_\alpha\partial_t\Phi\|_{L^\infty(\Sigma_s)} \right\} < \infty.$$

Also taking into account Remark 5.4, we see that neither of the two breakdown scenarios of Proposition 6.1 occur on $[0, T_{(\text{Max})}) \times \mathbb{R}^3$. Moreover, by Proposition 5.8, if C_* is large enough, then the bootstrap assumption inequalities (3B.5a)–(3B.5d) hold in a strict sense (that is, with "$\le$" replaced by "$<$") on $[0, T_{(\text{Max})}) \times \mathbb{R}^3$. Moreover, all estimates proved prior to Proposition 5.8 hold with ε replaced by $C\mathring{\epsilon}$, and we will use this fact in the rest of the proof without mentioning it again. Furthermore, (5B.5) and (5B.4b) respectively yield strict improvements of the bootstrap assumptions (3B.3) and (3B.4) for $(t, \underline{x}) \in [0, T_{(\text{Max})}) \times \mathbb{R}^3$. Next, we note that the estimate (5B.3b) implies that $\mathcal{I}_\star(t)$ cannot remain positive for t larger than $\mathring{A}_*^{-1}\{1+\mathcal{O}(\mathring{\epsilon})\}$. From this fact, it follows that $T_{(\text{Max})} < 2\mathring{A}_*^{-1}$. Combining these facts and appealing to Proposition 6.1, we deduce that $\{\partial_\alpha\Phi\}_{\alpha=0,1,2,3}$, $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$, and $\mathcal{I}$ extend as classical solutions to a region of the form $[0, T_{(\text{Max})} + \Delta) \times \mathbb{R}^3$ for some $\Delta > 0$ with $T_{(\text{Max})} + \Delta < 2\mathring{A}_*^{-1}$ (on which these variables enjoy the regularity properties guaranteed by Proposition 6.1), such that $\inf_{s\in[0,T_{(\text{Max})}+\Delta)} \mathcal{I}_\star(s) > 0$, and such that the bootstrap assumptions (3B.3)–(3B.5d) hold for $(t, \underline{x}) \in [0, T_{(\text{Max})} + \Delta) \times \mathbb{R}^3$. In total, this contradicts the definition of $T_{(\text{Max})}$. Therefore, (7.10) is impossible, and it follows that

$$T_{(\text{Max})} = \sup\left\{t > 0 \;\middle|\; \inf_{s\in[0,t)} \mathcal{I}_\star(s) > 0\right\}, \tag{7.11}$$

that

$$\inf_{s\in[0,T_{(\text{Max})})} \mathcal{I}_\star(s) = 0, \tag{7.12}$$

and that the estimates (5E.1)–(5E.3) and (5E.4a)–(5E.4e) hold for $t \in [0, T_{(\text{Max})})$.

Next, we note that the estimate (7.7) follows from (5B.4b). Then, using (7.7) and (7.12), we see that $\lim_{t\uparrow T_{(\text{Max})}} \sup_{s\in[0,t)} \|\partial_t\Phi\|_{L^\infty(\Sigma_s)} = \infty$, that is, that $\partial_t\Phi$ blows up at time $T_{(\text{Max})}$. We have therefore shown that $T_{(\text{Max})} = T_{(\text{Lifespan})}$ and that $T_{(\text{Lifespan})}$ is characterized by (7.2). Moreover, from the estimate (5B.3b), we find that $\mathcal{I}_\star$ vanishes for the first time at $T_{(\text{Lifespan})} = \mathring{A}_*^{-1}\{1+\mathcal{O}(\mathring{\epsilon})\}$, which in total yields (7.3a) and (7.4).

In the rest of this proof, we sometimes silently use that $\Psi_0 \in L^\infty([0, T_{(\text{Lifespan})}), L^2(\mathbb{R}^3))$ and $\mathcal{I} - 1 \in L^\infty([0, T_{(\text{Lifespan})}), L^2(\mathbb{R}^3))$. These facts do not follow from the energy estimates (5E.1) and (5E.3), but instead follow from (5E.4a), (5E.4d), and the compactly supported (in space) nature of Ψ_0 and $\mathcal{I} - 1$. Next, we easily conclude from the definition (4A.2) of $\mathbb{Q}_{(\mathring{\epsilon})}(t)$ and the fact that the estimate (5E.1) holds on $[0, T_{(\text{Lifespan})})$ that $\Psi_0 \in L^\infty([0, T_{(\text{Lifespan})}), H^5(\mathbb{R}^3))$ (as is stated in (7.5a)) and that $\Psi_i \in L^\infty([0, T_{(\text{Lifespan})}), H^4(\mathbb{R}^3))$ (as is stated in (7.5b)). The same reasoning yields that $\Psi_0 \in L^\infty([0, T_{(\text{Lifespan})}], H^5(\mathbb{R}^3))$ (as is stated in (7.6a)), where the open time interval is replaced with

$[0, T_{(\text{Lifespan})}]$. The fact that $\Psi_\alpha \in C([0, T_{(\text{Lifespan})}), H^5(\mathbb{R}^3))$ (as is stated in (7.5a)–(7.5b)) is a standard result that can be proved using energy-based arguments (similar to the ones we used to prove (5E.1)) and standard facts from functional analysis. We omit the details and instead refer the reader to [Speck 2008, Section 2.7.5]. We clarify that in proving this "soft result", it is important that, for fixed $t \in [0, T_{(\text{Lifespan})})$, we have $\min_{[0,t]\times\mathbb{R}^3} \mathcal{I} > 0$, which, in view of Remark 5.4, implies in particular that the weight $\mathscr{W}(\mathcal{I}^{-1}\Psi_0)$ on the right-hand side of (4A.2) is bounded from above and from below by t-dependent positive constants on $[0, t] \times \mathbb{R}^3$ (and thus the energy estimates are nondegenerate away from $\Sigma_{T_{(\text{Lifespan})}}$). Through similar reasoning based on (2B.1) (which states that $\partial_t \mathcal{I} = -\Psi_0$), the identity (2B.4), and the estimate (5E.3), we deduce that $\mathcal{I} - 1 \in C([0, T_{(\text{Lifespan})}), H^6(\mathbb{R}^3)) \cap L^\infty([0, T_{(\text{Lifespan})}), H^5(\mathbb{R}^3))$. We have therefore proved (7.5a)–(7.5c).

We will now prove (7.6a)–(7.6c). We first note that the estimates (5E.1) and (5E.2a)–(5E.2b) and (2B.1) imply that $\partial_t \Psi_\alpha \in L^\infty([0, T_{(\text{Lifespan})}), H^4(\mathbb{R}^3))$ and $\partial_t \mathcal{I} \in L^\infty([0, T_{(\text{Lifespan})}), H^5(\mathbb{R}^3))$. Hence, from the fundamental theorem of calculus, the initial conditions (2B.1) and (3A.1a)–(3A.1b), and the completeness of the Sobolev spaces $H^M(\mathbb{R}^3)$, we obtain $\Psi_\alpha \in C([0, T_{(\text{Lifespan})}], H^4(\mathbb{R}^3))$ and $\mathcal{I} - 1 \in C([0, T_{(\text{Lifespan})}], H^5(\mathbb{R}^3))$. In particular, we have shown (7.6b)–(7.6c). Moreover, (7.6c) and Sobolev embedding together yield that $\mathcal{I} \in C([0, T_{(\text{Lifespan})}], C(\mathbb{R}^3))$ and thus $\mathcal{I}_\star \in C([0, T_{(\text{Lifespan})}], [0, \infty))$. Since we have already shown that $T_{(\text{Lifespan})}$ is equal to the right-hand side of (7.11) and shown (7.12), it follows that $\mathcal{I}_\star(T_{(\text{Lifespan})}) = 0$ and that $\lim_{t \uparrow T_{(\text{Lifespan})}} \mathcal{I}_\star(t) = 0$, that is, that (7.3b) holds. To obtain that, for any real number $N < 5$, we have $\Psi_0 \in C([0, T_{(\text{Lifespan})}], H^N(\mathbb{R}^3))$ (as is stated in (7.6a)), we interpolate between[31] L^2 and H^5 and use the already-shown facts $\Psi_0 \in L^\infty([0, T_{(\text{Lifespan})}], H^5(\mathbb{R}^3)) \cap C([0, T_{(\text{Lifespan})}], H^4(\mathbb{R}^3))$. We have therefore proved (7.6a).

The desired localized blowup-result (7.9) for points in $\Sigma^{\text{Blowup}}_{T_{(\text{Lifespan})}}$ (where $\Sigma^{\text{Blowup}}_{T_{(\text{Lifespan})}}$ is defined in (7.8)) now follows from (7.7) and the continuous extension property $\mathcal{I} \in C([0, T_{(\text{Lifespan})}], C(\mathbb{R}^3))$ mentioned in the previous paragraph.

Finally, we will show that if $(T_{(\text{Lifespan})}, \underline{x}) \notin \Sigma^{\text{Blowup}}_{T_{(\text{Lifespan})}}$, then there exists an open ball $B_{\underline{x}} \subset \mathbb{R}^3$ centered at $\underline{x}$ such that, for $\alpha = 0, 1, 2, 3$, we have $\partial_\alpha \Phi \in C([0, T_{(\text{Lifespan})}], H^5(B_{\underline{x}}))$. To proceed, we first note that if $(T_{(\text{Lifespan})}, \underline{x}) \notin \Sigma^{\text{Blowup}}_{T_{(\text{Lifespan})}}$, then the results proved above imply that there exist a $\delta_{\underline{x}} > 0$ and a radius $r_{\underline{x}} > 0$ such that, with $B_{\underline{x};r_{\underline{x}}} \subset \mathbb{R}^3$ denoting the open ball of radius $r_{\underline{x}}$ centered at $\underline{x}$, we have $\mathcal{I}(t, \underline{y}) > 0$ for $(t, \underline{y}) \in [T_{(\text{Lifespan})} - \delta_{\underline{x}}, T_{(\text{Lifespan})}] \times \bar{B}_{\underline{x};r_{\underline{x}}}$ (where $\bar{B}_{\underline{x};r_{\underline{x}}}$ denotes the closure of $B_{\underline{x};r_{\underline{x}}}$) and such that, for $t \in [T_{(\text{Lifespan})} - \delta_{\underline{x}}, T_{(\text{Lifespan})})$, we have $\|\Psi_\alpha\|_{H^5(\{t\}\times B_{\underline{x};r_{\underline{x}}})} < \infty$ and $\|\mathcal{I} - 1\|_{H^6(\{t\}\times B_{\underline{x};r_{\underline{x}}})} < \infty$. Hence, since the wave speed of the system is uniformly bounded from above on $[T_{(\text{Lifespan})} - \delta_{\underline{x}}, T_{(\text{Lifespan})}] \times \bar{B}_{\underline{x};r_{\underline{x}}}$ (see Remark 3.3), since $\mathcal{I}$ is uniformly bounded from above and from below by positive ($\underline{x}$-dependent) constants on $[T_{(\text{Lifespan})} - \delta_{\underline{x}}, T_{(\text{Lifespan})}] \times \bar{B}_{\underline{x};r_{\underline{x}}}$, and since the estimates (7.5a)–(7.5c) and (7.6a)–(7.6c) hold, we can derive Sobolev estimates (using energy-based arguments) similar to the ones that we derived three paragraphs above, but localized in space,[32] for

[31] Here, we mean the following standard inequality: if $f \in H^5(\Sigma_t)$ and $0 \le N \le 5$, then there exists a constant $C_N > 0$ such that $\|f\|_{H^N(\Sigma_t)} \le C_N \|f\|^{1-N/5}_{L^2(\Sigma_t)} \|f\|^{N/5}_{H^5(\Sigma_t)}$.

[32] For example, for $\sigma > 0$ chosen sufficiently large, for t near $T_{(\text{Lifespan})}$, and for $s \in [t, T_{(\text{Lifespan})}]$, one can view the state of the solution on $\{t\} \times B_{\underline{x};r_{\underline{x}}}$ as an "initial" condition and use energy identities to obtain Sobolev estimates on $\{s\} \times B_{\underline{x};r_{\underline{x}}-\sigma s} \subset \Sigma_s$.

(2B.1) and (2B.6a)–(2B.6b), starting from initial conditions on $\{t\} \times B_{\underline{x};r_{\underline{x}}}$ for some t sufficiently close to $T_{(\text{Lifespan})}$. This yields the existence of an open ball $B_{\underline{x}} \subset B_{\underline{x};r_{\underline{x}}}$ centered at $\underline{x}$ such that the following regularity properties hold: $\Psi_\alpha \in C([0, T_{(\text{Lifespan})}], H^5(B_{\underline{x}}))$ and $\mathcal{I} - 1 \in C([0, T_{(\text{Lifespan})}], H^6(B_{\underline{x}}))$. We clarify that in deriving these spatially localized energy estimates on the *closed* time interval $[0, T_{(\text{Lifespan})}]$, we have exploited the following crucially important consequence of the bounds noted above and the assumptions of Section 2A on $\mathscr{W}$: the weight $\mathscr{W}(\mathcal{I}^{-1}\Psi_0)$ (which appears, for example, on the right-hand side of (4A.2)) is strictly positive on the domain $[T_{(\text{Lifespan})} - \delta_{\underline{x}}, T_{(\text{Lifespan})}] \times \bar{B}_{\underline{x}}$. Finally, from the regularity properties of $\{\Psi_\alpha\}_{\alpha=0,1,2,3}$ and $\mathcal{I}$ mentioned above, the positivity of $\mathcal{I}$ on $[T_{(\text{Lifespan})} - \delta_{\underline{x}}, T_{(\text{Lifespan})}] \times \bar{B}_{\underline{x}}$, and the standard Sobolev–Moser calculus, we conclude, in view of Definition 2.3, the desired result $\partial_\alpha \Phi \in C([0, T_{(\text{Lifespan})}], H^5(B_{\underline{x}}))$. We have therefore proved the theorem. □

References

[Alinhac 1999a] S. Alinhac, "Blowup of small data solutions for a class of quasilinear wave equations in two space dimensions, II", *Acta Math.* **182**:1 (1999), 1–23. MR Zbl

[Alinhac 1999b] S. Alinhac, "Blowup of small data solutions for a quasilinear wave equation in two space dimensions", *Ann. of Math.* (2) **149**:1 (1999), 97–127. MR Zbl

[Alinhac 2001] S. Alinhac, "The null condition for quasilinear wave equations in two space dimensions, II", *Amer. J. Math.* **123**:6 (2001), 1071–1101. MR Zbl

[Beyer and LeFloch 2010] F. Beyer and P. G. LeFloch, "Second-order hyperbolic Fuchsian systems and applications", *Classical Quantum Gravity* **27**:24 (2010), art. id. 245012. MR Zbl

[Bricmont and Kupiainen 1994] J. Bricmont and A. Kupiainen, "Universality in blow-up for nonlinear heat equations", *Nonlinearity* **7**:2 (1994), 539–575. MR Zbl

[Chatzikaleas and Donninger 2019] A. Chatzikaleas and R. Donninger, "Stable blowup for the cubic wave equation in higher dimensions", *J. Differential Equations* **266**:10 (2019), 6809–6865. MR Zbl

[Christodoulou 2007] D. Christodoulou, *The formation of shocks in 3-dimensional fluids*, Eur. Math. Soc., Zürich, 2007. MR Zbl

[Christodoulou 2009] D. Christodoulou, *The formation of black holes in general relativity*, Eur. Math. Soc., Zürich, 2009. MR Zbl

[Christodoulou 2019] D. Christodoulou, *The shock development problem*, EMS Monographs in Math. **8**, Eur. Math. Soc., Zürich, 2019. Zbl

[Christodoulou and Klainerman 1993] D. Christodoulou and S. Klainerman, *The global nonlinear stability of the Minkowski space*, Princeton Math. Series **41**, Princeton Univ. Press, 1993. MR Zbl

[Christodoulou and Lisibach 2016] D. Christodoulou and A. Lisibach, "Shock development in spherical symmetry", *Ann. PDE* **2**:1 (2016), art. id. 3. MR Zbl

[Christodoulou and Miao 2014] D. Christodoulou and S. Miao, *Compressible flow and Euler's equations*, Surveys of Modern Math. **9**, Int. Press, Somerville, MA, 2014. MR Zbl

[Christodoulou and Perez 2016] D. Christodoulou and D. R. Perez, "On the formation of shocks of electromagnetic plane waves in non-linear crystals", *J. Math. Phys.* **57**:8 (2016), art. id. 081506. MR Zbl

[Collot 2018] C. Collot, *Type II blow up manifolds for the energy supercritical semilinear wave equation*, Mem. Amer. Math. Soc. **1205**, Amer. Math. Soc., Providence, RI, 2018. MR Zbl

[Costin et al. 2016] O. Costin, R. Donninger, and X. Xia, "A proof for the mode stability of a self-similar wave map", *Nonlinearity* **29**:8 (2016), 2451–2473. MR Zbl

[Costin et al. 2017] O. Costin, R. Donninger, and I. Glogić, "Mode stability of self-similar wave maps in higher dimensions", *Comm. Math. Phys.* **351**:3 (2017), 959–972. MR Zbl

[Côte et al. 2015] R. Côte, C. E. Kenig, A. Lawrie, and W. Schlag, "Characterization of large energy solutions of the equivariant wave map problem, I", *Amer. J. Math.* **137**:1 (2015), 139–207. MR Zbl

[Côte et al. 2018] R. Côte, C. E. Kenig, A. Lawrie, and W. Schlag, "Profiles for the radial focusing 4*d* energy-critical wave equation", *Comm. Math. Phys.* **357**:3 (2018), 943–1008. MR Zbl

[Dafermos 2000] C. M. Dafermos, *Hyperbolic conservation laws in continuum physics*, Grundlehren der Math. Wissenschaften **325**, Springer, 2000. MR Zbl

[Donninger 2010] R. Donninger, "Nonlinear stability of self-similar solutions for semilinear wave equations", *Comm. Partial Differential Equations* **35**:4 (2010), 669–684. MR Zbl

[Donninger 2011] R. Donninger, "On stable self-similar blowup for equivariant wave maps", *Comm. Pure Appl. Math.* **64**:8 (2011), 1095–1147. MR Zbl

[Donninger 2017] R. Donninger, "Strichartz estimates in similarity coordinates and stable blowup for the critical wave equation", *Duke Math. J.* **166**:9 (2017), 1627–1683. MR Zbl

[Donninger and Schörkhuber 2012] R. Donninger and B. Schörkhuber, "Stable self-similar blow up for energy subcritical wave equations", *Dyn. Partial Differ. Equ.* **9**:1 (2012), 63–87. MR Zbl

[Donninger and Schörkhuber 2014] R. Donninger and B. Schörkhuber, "Stable blow up dynamics for energy supercritical wave equations", *Trans. Amer. Math. Soc.* **366**:4 (2014), 2167–2189. MR Zbl

[Donninger and Schörkhuber 2016] R. Donninger and B. Schörkhuber, "On blowup in supercritical wave equations", *Comm. Math. Phys.* **346**:3 (2016), 907–943. MR Zbl

[Donninger et al. 2012] R. Donninger, B. Schörkhuber, and P. C. Aichelburg, "On stable self-similar blow up for equivariant wave maps: the linearized problem", *Ann. Henri Poincaré* **13**:1 (2012), 103–144. MR Zbl

[Donninger et al. 2014] R. Donninger, M. Huang, J. Krieger, and W. Schlag, "Exotic blowup solutions for the u^5 focusing wave equation in $\mathbb{R}^3$", *Michigan Math. J.* **63**:3 (2014), 451–501. MR Zbl

[Duyckaerts et al. 2011] T. Duyckaerts, C. Kenig, and F. Merle, "Universality of blow-up profile for small radial type II blow-up solutions of the energy-critical wave equation", *J. Eur. Math. Soc.* **13**:3 (2011), 533–599. MR Zbl

[Duyckaerts et al. 2012a] T. Duyckaerts, C. Kenig, and F. Merle, "Profiles of bounded radial solutions of the focusing, energy-critical wave equation", *Geom. Funct. Anal.* **22**:3 (2012), 639–698. MR Zbl

[Duyckaerts et al. 2012b] T. Duyckaerts, C. Kenig, and F. Merle, "Universality of the blow-up profile for small type II blow-up solutions of the energy-critical wave equation: the nonradial case", *J. Eur. Math. Soc.* **14**:5 (2012), 1389–1454. MR Zbl

[Duyckaerts et al. 2013] T. Duyckaerts, C. Kenig, and F. Merle, "Classification of radial solutions of the focusing, energy-critical wave equation", *Camb. J. Math.* **1**:1 (2013), 75–144. MR Zbl

[Gao and Krieger 2015] C. Gao and J. Krieger, "Optimal polynomial blow up range for critical wave maps", *Comm. Pure Appl. Anal.* **14**:5 (2015), 1705–1741. MR Zbl

[Guo and Tahvildar-Zadeh 1999] Y. Guo and A. S. Tahvildar-Zadeh, "Formation of singularities in relativistic fluid dynamics and in spherically symmetric plasma dynamics", pp. 151–161 in *Nonlinear partial differential equations* (Evanston, IL, 1998), edited by G.-Q. Chen and E. DiBenedetto, Contemp. Math. **238**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[Hillairet and Raphaël 2012] M. Hillairet and P. Raphaël, "Smooth type II blow-up solutions to the four-dimensional energy-critical wave equation", *Anal. PDE* **5**:4 (2012), 777–829. MR Zbl

[Holzegel et al. 2016] G. Holzegel, S. Klainerman, J. Speck, and W. W.-Y. Wong, "Small-data shock formation in solutions to 3D quasilinear wave equations: an overview", *J. Hyperbolic Differ. Equ.* **13**:1 (2016), 1–105. MR Zbl

[Jendrej 2016] J. Jendrej, "Bounds on the speed of type II blow-up for the energy critical wave equation in the radial case", *Int. Math. Res. Not.* **2016**:21 (2016), 6656–6688. MR Zbl

[Jendrej 2017] J. Jendrej, "Construction of type II blow-up solutions for the energy-critical wave equation in dimension 5", *J. Funct. Anal.* **272**:3 (2017), 866–917. MR Zbl

[John 1974] F. John, "Formation of singularities in one-dimensional nonlinear wave propagation", *Comm. Pure Appl. Math.* **27** (1974), 377–405. MR Zbl

[John 1979] F. John, "Blow-up of solutions of nonlinear wave equations in three space dimensions", *Manuscripta Math.* **28**:1-3 (1979), 235–268. MR Zbl

[John 1981] F. John, "Blow-up for quasilinear wave equations in three space dimensions", *Comm. Pure Appl. Math.* **34**:1 (1981), 29–51. MR

[Kato 1980] T. Kato, "Blow-up of solutions of some nonlinear hyperbolic equations", *Comm. Pure Appl. Math.* **33**:4 (1980), 501–505. 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

[Kichenassamy 1996] S. Kichenassamy, "The blow-up problem for exponential nonlinearities", *Comm. Partial Differential Equations* **21**:1-2 (1996), 125–162. MR Zbl

[Klainerman and Rodnianski 2003] S. Klainerman and I. Rodnianski, "Improved local well-posedness for quasilinear wave equations in dimension three", *Duke Math. J.* **117**:1 (2003), 1–124. MR Zbl

[Klainerman and Rodnianski 2005] S. Klainerman and I. Rodnianski, "Rough solutions of the Einstein-vacuum equations", *Ann. of Math.* (2) **161**:3 (2005), 1143–1193. MR Zbl

[Klainerman et al. 2015] S. Klainerman, I. Rodnianski, and J. Szeftel, "The bounded L^2 curvature conjecture", *Invent. Math.* **202**:1 (2015), 91–216. MR Zbl

[Krieger and Schlag 2014] J. Krieger and W. Schlag, "Full range of blow up exponents for the quintic wave equation in three dimensions", *J. Math. Pures Appl.* (9) **101**:6 (2014), 873–900. MR Zbl

[Krieger et al. 2008] J. Krieger, W. Schlag, and D. Tataru, "Renormalization and blow up for charge one equivariant critical wave maps", *Invent. Math.* **171**:3 (2008), 543–615. MR Zbl

[Krieger et al. 2009] J. Krieger, W. Schlag, and D. Tataru, "Slow blow-up solutions for the $H^1(\mathbb{R}^3)$ critical focusing semilinear wave equation", *Duke Math. J.* **147**:1 (2009), 1–53. MR Zbl

[Lax 1964] P. D. Lax, "Development of singularities of solutions of nonlinear hyperbolic partial differential equations", *J. Math. Phys.* **5** (1964), 611–613. MR Zbl

[Lax 1972] P. D. Lax, "The formation and decay of shock waves", *Amer. Math. Monthly* **79**:3 (1972), 227–241. MR Zbl

[Lax 1973] P. D. Lax, *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*, CBMS Regional Conf. Series Appl. Math. **11**, Soc. Indust. Appl. Math., Philadelphia, 1973. MR Zbl

[Luk 2018] J. Luk, "Weak null singularities in general relativity", *J. Amer. Math. Soc.* **31**:1 (2018), 1–63. MR Zbl

[Luk and Speck 2016] J. Luk and J. Speck, "The hidden null structure of the compressible Euler equations and a prelude to applications", 2016. To appear in *J. Hyperbolic Differ. Equ.* arXiv

[Luk and Speck 2018] J. Luk and J. Speck, "Shock formation in solutions to the 2D compressible Euler equations in the presence of non-zero vorticity", *Invent. Math.* **214**:1 (2018), 1–169. MR Zbl

[Majda 1981] A. Majda, "The existence and stability of multidimensional shock fronts", *Bull. Amer. Math. Soc.* (*N.S.*) **4**:3 (1981), 342–344. MR Zbl

[Majda 1983a] A. Majda, *The existence of multidimensional shock fronts*, Mem. Amer. Math. Soc. **281**, Amer. Math. Soc., Providence, RI, 1983. MR Zbl

[Majda 1983b] A. Majda, *The stability of multidimensional shock fronts*, Mem. Amer. Math. Soc. **275**, Amer. Math. Soc., Providence, RI, 1983. MR Zbl

[Merle and Zaag 1997] F. Merle and H. Zaag, "Stability of the blow-up profile for equations of the type $u_t = \Delta u + |u|^{p-1}u$", *Duke Math. J.* **86**:1 (1997), 143–195. MR Zbl

[Merle and Zaag 2003] F. Merle and H. Zaag, "Determination of the blow-up rate for the semilinear wave equation", *Amer. J. Math.* **125**:5 (2003), 1147–1164. MR Zbl

[Merle and Zaag 2005] F. Merle and H. Zaag, "Determination of the blow-up rate for a critical semilinear wave equation", *Math. Ann.* **331**:2 (2005), 395–416. MR Zbl

[Merle and Zaag 2007] F. Merle and H. Zaag, "Existence and universality of the blow-up profile for the semilinear wave equation in one space dimension", *J. Funct. Anal.* **253**:1 (2007), 43–121. MR Zbl

[Merle and Zaag 2012] F. Merle and H. Zaag, "Isolatedness of characteristic points at blowup for a 1-dimensional semilinear wave equation", *Duke Math. J.* **161**:15 (2012), 2837–2908. MR Zbl

[Merle and Zaag 2015] F. Merle and H. Zaag, "On the stability of the notion of non-characteristic point and blow-up profile for semilinear wave equations", *Comm. Math. Phys.* **333**:3 (2015), 1529–1562. MR Zbl

[Merle and Zaag 2016] F. Merle and H. Zaag, "Dynamics near explicit stationary solutions in similarity variables for solutions of a semilinear wave equation in higher dimensions", *Trans. Amer. Math. Soc.* **368**:1 (2016), 27–87. MR Zbl

[Merle et al. 2010] F. Merle, P. Raphaël, and J. Szeftel, "Stable self-similar blow-up dynamics for slightly L^2 super-critical NLS equations", *Geom. Funct. Anal.* **20**:4 (2010), 1028–1071. MR Zbl

[Merle et al. 2014] F. Merle, P. Raphaël, and J. Szeftel, "On collapsing ring blow-up solutions to the mass supercritical nonlinear Schrödinger equation", *Duke Math. J.* **163**:2 (2014), 369–431. MR Zbl

[Miao 2018] S. Miao, "On the formation of shock for quasilinear wave equations with weak intensity pulse", *Ann. PDE* **4**:1 (2018), art. id. 10. MR Zbl

[Miao and Yu 2017] S. Miao and P. Yu, "On the formation of shocks for quasilinear wave equations", *Invent. Math.* **207**:2 (2017), 697–831. MR Zbl

[Raphaël and Rodnianski 2012] P. Raphaël and I. Rodnianski, "Stable blow up dynamics for the critical co-rotational wave maps and equivariant Yang–Mills problems", *Publ. Math. Inst. Hautes Études Sci.* **115** (2012), 1–122. MR Zbl

[Rauch 1986] J. Rauch, "BV estimates fail for most quasilinear hyperbolic systems in dimensions greater than one", *Comm. Math. Phys.* **106**:3 (1986), 481–484. MR Zbl

[Riemann 1860] B. Riemann, "Über die Fortpflanzung ebener Luftwellen von endlicher Schwingungsweite", *Abh. König. Ges. Wiss. Göttengen* **8** (1860), 43–66.

[Rodnianski and Speck 2018a] I. Rodnianski and J. Speck, "A regime of linear stability for the Einstein-scalar field system with applications to nonlinear big bang formation", *Ann. of Math.* (2) **187**:1 (2018), 65–156. MR Zbl

[Rodnianski and Speck 2018b] I. Rodnianski and J. Speck, "Stable big bang formation in near-FLRW solutions to the Einstein-scalar field and Einstein-stiff fluid systems", *Selecta Math.* (*N.S.*) **24**:5 (2018), 4293–4459. MR Zbl

[Rodnianski and Sterbenz 2010] I. Rodnianski and J. Sterbenz, "On the formation of singularities in the critical O(3) σ-model", *Ann. of Math.* (2) **172**:1 (2010), 187–242. MR Zbl

[Sbierski 2016] J. Sbierski, "On the existence of a maximal Cauchy development for the Einstein equations: a dezornification", *Ann. Henri Poincaré* **17**:2 (2016), 301–329. MR Zbl

[Shahshahani 2016] S. Shahshahani, "Renormalization and blow-up for wave maps from $S^2 \times \mathbb{R}$ to S^2", *Trans. Amer. Math. Soc.* **368**:8 (2016), 5621–5654. MR Zbl

[Sideris 1984a] T. C. Sideris, "Formation of singularities in solutions to nonlinear hyperbolic equations", *Arch. Rational Mech. Anal.* **86**:4 (1984), 369–381. MR Zbl

[Sideris 1984b] T. C. Sideris, "Nonexistence of global solutions to semilinear wave equations in high dimensions", *J. Differential Equations* **52**:3 (1984), 378–406. MR Zbl

[Sideris 1985] T. C. Sideris, "Formation of singularities in three-dimensional compressible fluids", *Comm. Math. Phys.* **101**:4 (1985), 475–485. MR Zbl

[Smith and Tataru 2005] H. F. Smith and D. Tataru, "Sharp local well-posedness results for the nonlinear wave equation", *Ann. of Math.* (2) **162**:1 (2005), 291–366. MR Zbl

[Speck 2008] J. R. Speck, *On the questions of local and global well-posedness for the hyperbolic PDEs occurring in some relativistic theories of gravity and electromagnetism*, Ph.D. thesis, Rutgers University, 2008, available at https://tinyurl.com/speckphd.

[Speck 2016] J. Speck, *Shock formation in small-data solutions to 3D quasilinear wave equations*, Math. Surveys and Monographs **214**, Amer. Math. Soc., Providence, RI, 2016. MR Zbl

[Speck 2017] J. Speck, "Finite-time degeneration of hyperbolicity without blowup for quasilinear wave equations", *Anal. PDE* **10**:8 (2017), 2001–2030. MR Zbl

[Speck 2018] J. Speck, "Shock formation for $2D$ quasilinear wave systems featuring multiple speeds: blowup for the fastest wave, with non-trivial interactions up to the singularity", *Ann. PDE* **4**:1 (2018), art. id. 6. MR Zbl

[Speck 2019a] J. Speck, "Multidimensional nonlinear geometric optics for transport operators with applications to stable shock formation", *Pure Appl. Anal.* **1**:3 (2019), 447–514. MR Zbl

[Speck 2019b] J. Speck, "A new formulation of the 3D compressible Euler equations with dynamic entropy: remarkable null structures and regularity properties", *Arch. Ration. Mech. Anal.* **234**:3 (2019), 1223–1279. MR Zbl

[Speck et al. 2016] J. Speck, G. Holzegel, J. Luk, and W. Wong, "Stable shock formation for nearly simple outgoing plane symmetric waves", *Ann. PDE* **2**:2 (2016), art. id. 10. MR Zbl

[Wong 2013] W. W.-Y. Wong, "A comment on the construction of the maximal globally hyperbolic Cauchy development", *J. Math. Phys.* **54**:11 (2013), art. id. 113511. MR Zbl

Received 16 Nov 2017. Revised 9 Dec 2018. Accepted 12 Feb 2019.

JARED SPECK: jared.speck@vanderbilt.edu
Vanderbilt University, Nashville, TN, United States

ANALYSIS AND PDE
Vol. 13, No. 1, 2020
dx.doi.org/10.2140/apde.2020.13.147

ASYMPTOTIC EXPANSIONS OF FUNDAMENTAL SOLUTIONS IN PARABOLIC HOMOGENIZATION

JUN GENG AND ZHONGWEI SHEN

For a family of second-order parabolic systems with rapidly oscillating and time-dependent periodic coefficients, we investigate the asymptotic behavior of fundamental solutions and establish sharp estimates for the remainders.

1. Introduction

In this paper we study the asymptotic behavior of fundamental solutions $\Gamma_\varepsilon(x, t; y, s)$ for a family of second-order parabolic operators $\partial_t + \mathcal{L}_\varepsilon$ with rapidly oscillating and time-dependent periodic coefficients. Specifically, we consider

$$\mathcal{L}_\varepsilon = -\operatorname{div}(A(x/\varepsilon, t/\varepsilon^2)\nabla) \tag{1-1}$$

in $\mathbb{R}^d \times \mathbb{R}$, where $\varepsilon > 0$ and $A(y, s) = (a_{ij}^{\alpha\beta}(y, s))$ with $1 \le i, j \le d$ and $1 \le \alpha, \beta \le m$. Throughout the paper we will assume that the coefficient matrix $A = A(y, s)$ is real, bounded measurable and satisfies the ellipticity condition

$$\|A\|_\infty \le \mu^{-1} \quad \text{and} \quad \mu|\xi|^2 \le a_{ij}^{\alpha\beta}(y, s)\xi_i^\alpha \xi_j^\beta \tag{1-2}$$

for any $\xi = (\xi_i^\alpha) \in \mathbb{R}^{m\times d}$ and a.e. $(y, s) \in \mathbb{R}^{d+1}$, where $\mu > 0$. We also assume that A is 1-periodic; i.e.,

$$A(y+z, s+t) = A(y, s) \quad \text{for } (z, t) \in \mathbb{Z}^{d+1} \text{ and a.e. } (y, s) \in \mathbb{R}^{d+1}. \tag{1-3}$$

Under these assumptions it is known that as $\varepsilon \to 0$, the operator $\partial_t + \mathcal{L}_\varepsilon$ G-converges to a parabolic operator $\partial_t + \mathcal{L}_0$ with constant coefficients [Bensoussan et al. 1978].

In the scalar case $m = 1$, it follows from a celebrated theorem of John Nash [1958] that local solutions of $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = 0$ are Hölder continuous. More precisely, if $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = 0$ in $Q_{2r} = Q_{2r}(x_0, t_0)$ for some $(x_0, t_0) \in \mathbb{R}^{d+1}$ and $0 < r < \infty$, where

$$Q_r(x_0, t_0) = B(x_0, r) \times (t_0 - r^2, t_0), \tag{1-4}$$

then there exists some $\sigma \in (0, 1)$, depending only on d and μ, such that

$$\|u_\varepsilon\|_{C^{\sigma,\sigma/2}(Q_r)} \le C r^{-\sigma} \left(\frac{1}{|Q_{2r}|} \int_{Q_{2r}} |u_\varepsilon|^2 \right)^{1/2}, \tag{1-5}$$

Geng was supported in part by the NNSF of China (no. 11971212 and no. 11571152). Shen was supported in part by NSF grant DMS-1600520.
MSC2010: 35B27, 35E05, 35K40.
Keywords: homogenization, parabolic system, fundamental solution.

where $C > 0$ depends only on d and μ. In particular, C and σ are independent of $\varepsilon > 0$. The periodicity assumption (1-3) is not needed here. It follows that the fundamental solution $\Gamma_\varepsilon(x, t; y, s)$ for $\partial_t + \mathcal{L}_\varepsilon$ exists and satisfies the Gaussian estimate

$$|\Gamma_\varepsilon(x, t; y, s)| \le \frac{C}{(t-s)^{d/2}} \exp\left\{-\frac{\kappa|x-y|^2}{t-s}\right\} \tag{1-6}$$

for any $x, y \in \mathbb{R}^d$ and $-\infty < s < t < \infty$, where $\kappa > 0$ depends only on μ and $C > 0$ depends on d and μ (also see [Aronson 1967; Fabes and Stroock 1986] for lower bounds).

If $m \ge 2$, the global Hölder estimate (1-5) for $1 < r < \infty$ was established recently in [Geng and Shen 2015] for any $\sigma \in (0, 1)$ under the assumptions that A is elliptic, periodic, and $A \in \mathrm{VMO}_x$ (see (2-4) for the definition of VMO_x). We mention that the local Hölder estimate for $0 < r < \varepsilon$ without the periodicity condition was obtained earlier in [Byun 2007; Krylov 2007]. Consequently, by [Hofmann and Kim 2004; Cho et al. 2008], the matrix of fundamental solutions $\Gamma_\varepsilon(x, t; y, s) = (\Gamma_\varepsilon^{\alpha\beta}(x, t; y, s))$, with $1 \le \alpha, \beta \le m$, exists and satisfies the estimate (1-6), where $\kappa > 0$ depends only on μ. The constant $C > 0$ in (1-6) depends on d, m, μ and the function $A^\#(r)$ in (2-5), but not on $\varepsilon > 0$.

The primary purpose of this paper is to study the asymptotic behavior, as $\varepsilon \to 0$, of $\Gamma_\varepsilon(x, t; y, s)$, $\nabla_x \Gamma_\varepsilon(x, t; y, s)$, $\nabla_y \Gamma_\varepsilon(x, t; y, s)$, and $\nabla_x \nabla_y \Gamma_\varepsilon(x, t; y, s)$. Our main results extend the analogous estimates for elliptic operators $-\operatorname{div}(A(x/\varepsilon)\nabla)$ in [Avellaneda and Lin 1991; Kenig et al. 2014] to the parabolic setting. As demonstrated in the elliptic case [Kenig and Shen 2011], the estimates in this paper open the doors for the use of layer potentials in solving initial-boundary value problems for the parabolic operators $\partial_t + \mathcal{L}_\varepsilon$ with sharp estimates that are uniform in $\varepsilon > 0$.

Let $\Gamma_0(x, t; y, s)$ denote the matrix of fundamental solutions for the homogenized operator $\partial_t + \mathcal{L}_0$, where $\mathcal{L}_0 = -\operatorname{div}(\hat{A}\nabla)$ and $\hat{A} = (\hat{a}_{ij}^{\alpha\beta})$ is given by (2-7). Since $\hat{A}$ is constant and satisfies the ellipticity condition (2-8), it is well known that $\Gamma_0(x, t; y, s) = \Gamma_0(x-y, t-s; 0, 0)$ and for any $x, y \in \mathbb{R}^d$ and $-\infty < s < t < \infty$,

$$|\nabla_x^M \partial_t^N \Gamma_0(x, t; y, s)| \le \frac{C}{(t-s)^{(d+M+2N)/2}} \exp\left\{-\frac{\kappa|x-y|^2}{t-s}\right\} \tag{1-7}$$

for any $M, N \ge 0$, where $\kappa > 0$ depends only on μ, and C depends on d, m, M, N, and μ.

Our first result provides the sharp estimate for $\Gamma_\varepsilon - \Gamma_0$.

Theorem 1.1. *Suppose that the coefficient matrix A satisfies conditions (1-2) and (1-3). If $m \ge 2$, we also assume that $A \in \mathrm{VMO}_x$. Then*

$$|\Gamma_\varepsilon(x, t; y, s) - \Gamma_0(x, t; y, s)| \le \frac{C\varepsilon}{(t-s)^{(d+1)/2}} \exp\left\{-\frac{\kappa|x-y|^2}{t-s}\right\} \tag{1-8}$$

for any $x, y \in \mathbb{R}^d$ and $-\infty < s < t < \infty$, where $\kappa > 0$ depends only on μ. The constant C depends on d, m, μ, and $A^\#$ (if $m \ge 2$).

Let $\chi(y, s) = (\chi_j^{\alpha\beta}(y, s))$, where $1 \le j \le d$ and $1 \le \alpha, \beta \le m$, denote the matrix of correctors for $\partial_t + \mathcal{L}_\varepsilon$ (see Section 2 for its definition). The next theorem gives an asymptotic expansion for $\nabla_x \Gamma_\varepsilon(x, t; y, s)$.

Theorem 1.2. *Suppose that the coefficient matrix A satisfies conditions (1-2) and (1-3). Also assume that A is Hölder continuous,*

$$|A(x,t)-A(y,s)|\le \tau(|x-y|+|t-s|^{1/2})^{\lambda} \tag{1-9}$$

for any $(x,t),(y,s)\in\mathbb{R}^{d+1}$, where $\tau\ge 0$ and $\lambda\in(0,1)$. Then

$$\begin{aligned}|\nabla_x\Gamma_\varepsilon(x,t;y,s)-(I+\nabla\chi(x/\varepsilon,t/\varepsilon^2))\nabla_x\Gamma_0(x,t;y,s)|\\ \le \frac{C\varepsilon}{(t-s)^{(d+2)/2}}\log(2+\varepsilon^{-1}|t-s|^{1/2})\exp\Big\{-\frac{\kappa|x-y|^2}{t-s}\Big\}\end{aligned} \tag{1-10}$$

for any $x,y\in\mathbb{R}^d$ and $-\infty<s<t<\infty$, where $\kappa>0$ depends only on μ. The constant C depends on d, m, μ, and (λ,τ) in (1-9).

With the summation convention this means that for $1\le i\le d$ and $1\le\alpha,\beta\le m$

$$\left|\frac{\partial\Gamma_\varepsilon^{\alpha\beta}}{\partial x_i}(x,t;y,s)-\frac{\partial\Gamma_0^{\alpha\beta}}{\partial x_i}(x,t;y,s)-\frac{\partial\chi_j^{\alpha\gamma}}{\partial x_i}(x/\varepsilon,t/\varepsilon^2)\frac{\partial\Gamma_0^{\gamma\beta}}{\partial x_j}(x,t;y,s)\right| \tag{1-11}$$

is bounded by the right-hand side of (1-10). Let $\tilde A(y,s)=(\tilde a_{ij}^{\alpha\beta}(y,s))$, where $\tilde a_{ij}^{\alpha\beta}(y,s)=a_{ji}^{\beta\alpha}(y,-s)$. Let $\widetilde\Gamma_\varepsilon(x,t;y,s)=(\widetilde\Gamma_\varepsilon^{\alpha\beta}(x,t;y,s))$ denote the matrix of fundamental solutions for the operator $\partial_t+\widetilde{\mathcal{L}}_\varepsilon$, where $\widetilde{\mathcal{L}}_\varepsilon=-\operatorname{div}(\tilde A(x/\varepsilon,t/\varepsilon^2)\nabla)$. Then

$$\Gamma_\varepsilon^{\alpha\beta}(x,t;y,s)=\widetilde\Gamma_\varepsilon^{\beta\alpha}(y,-s;x,-t). \tag{1-12}$$

Since $\tilde A$ satisfies the same conditions as A, it follows from (1-10), (1-11) and (1-12) that

$$\left|\frac{\partial\Gamma_\varepsilon^{\beta\alpha}}{\partial y_i}(x,t;y,s)-\frac{\partial\Gamma_0^{\beta\alpha}}{\partial y_i}(x,t;y,s)-\frac{\partial\tilde\chi_j^{\alpha\gamma}}{\partial y_i}(y/\varepsilon,-s/\varepsilon^2)\frac{\partial\Gamma_0^{\beta\gamma}}{\partial y_j}(x,t;y,s)\right| \tag{1-13}$$

is bounded by the right-hand side of (1-10), where $\tilde\chi(y,s)=(\tilde\chi_j^{\alpha\beta}(y,s))$ denotes the correctors for $\partial_t+\widetilde{\mathcal{L}}_\varepsilon$. That is,

$$\begin{aligned}|\nabla_y\Gamma_\varepsilon^T(x,t;y,s)-(I+\nabla\tilde\chi(y/\varepsilon,-s/\varepsilon^2))\nabla_y\Gamma_0^T(x,t;y,s)|\\ \le \frac{C\varepsilon}{(t-s)^{(d+2)/2}}\log(2+\varepsilon^{-1}|t-s|^{1/2})\exp\Big\{-\frac{\kappa|x-y|^2}{t-s}\Big\},\end{aligned} \tag{1-14}$$

where Γ_ε^T denotes the transpose of the matrix Γ_ε.

We also obtain an asymptotic expansion for $\nabla_x\nabla_y\Gamma_\varepsilon(x,t;y,s)$.

Theorem 1.3. *Under the same assumptions on A as in Theorem 1.2, the estimate*

$$\begin{aligned}&\Big|\frac{\partial}{\partial x_i\partial y_j}\{\Gamma_\varepsilon^{\alpha\beta}(x,t;y,s)\}\\ &-\frac{\partial}{\partial x_i}\{\delta^{\alpha\gamma}x_k+\varepsilon\chi_k^{\alpha\gamma}(x/\varepsilon,t/\varepsilon^2)\}\frac{\partial^2}{\partial x_k\partial y_\ell}\{\Gamma_0^{\gamma\sigma}(x,t;y,s)\}\frac{\partial}{\partial y_j}\{\delta^{\beta\sigma}y_\ell+\varepsilon\tilde\chi_\ell^{\beta\sigma}(y/\varepsilon,-s/\varepsilon^2)\}\Big|\\ &\le \frac{C\varepsilon}{(t-s)^{(d+3)/2}}\log(2+\varepsilon^{-1}|t-s|^{1/2})\exp\Big\{-\frac{\kappa|x-y|^2}{t-s}\Big\}\end{aligned} \tag{1-15}$$

holds for $x, y \in \mathbb{R}^d$ and $-\infty < s < t < \infty$, where κ depends only on μ. The constant C depends on d, m, μ, and (λ, τ) in (1-9).

Remark 1.4. The estimates (1-10), (1-14) and (1-15) are sharp, up to the logarithmic factor $\log(2+\varepsilon^{-1}|t-s|^{1/2})$, which is probably not necessary. It may be possible to remove the logarithmic factor by using higher-order correctors in the proof. However, we will not pursue this idea in the present paper.

In the scale case $m=1$, the estimate (1-8), *without* the exponential factor, is known under the conditions that A is elliptic, periodic, symmetric, and time-independent; see [Jikov et al. 1994, p. 77]. This was proved by using the Floquet–Bloch decomposition of the fundamental solutions and by studying the spectral properties of elliptic operators

$$-(\nabla+ik)\cdot A(\nabla+ik)$$

in a periodic cell, where $i=\sqrt{-1}$ and $k \in \mathbb{R}^d$. Such an approach is not available when the coefficient matrix A is time-dependent. To the best of authors' knowledge, the Gaussian bound in Theorem 1.1 as well as our estimates in Theorems 1.2 and 1.3 are new even in the case that $m=1$ and A is time-independent.

As a corollary of Theorems 1.1 and 1.2, we establish an interesting result on equistabilization for time-dependent coefficients; cf. [Jikov et al. 1994, p. 77].

Corollary 1.5. *Assume that A satisfies the same conditions as in Theorem 1.1. Let $f \in L^\infty(\mathbb{R}^d)$ and u_ε be the bounded solution of the Cauchy problem,*

$$\begin{cases} (\partial_t+\mathcal{L}_\varepsilon)u_\varepsilon = 0 & \text{in } \mathbb{R}^d \times (0,\infty), \\ u_\varepsilon = f & \text{on } \mathbb{R}^d \times \{t=0\}, \end{cases} \tag{1-16}$$

with $\varepsilon = 1$ or 0. Then for any $x \in \mathbb{R}^d$ and $t \geq 1$,

$$|u_1(x,t)-u_0(x,t)| \leq Ct^{-1/2}\|f\|_\infty. \tag{1-17}$$

Furthermore, if A is Hölder continuous,

$$\left|\nabla u_1^\alpha(x,t)-\nabla u_0^\alpha(x,t)-\nabla \chi_j^{\alpha\beta}(x,t)\frac{\partial u_0^\beta}{\partial x_j}(x,t)\right| \leq Ct^{-1}\log(2+t)\|f\|_\infty \tag{1-18}$$

for any $x \in \mathbb{R}^d$ and $t \geq 1$.

We now describe some of the key ideas in the proof of Theorems 1.1, 1.2, and 1.3. As indicated earlier, our main results extend the analogous results in [Avellaneda and Lin 1991; Kenig et al. 2014] for the elliptic operators $-\operatorname{div}(A(x/\varepsilon)\nabla)$, where $A = A(y)$ is elliptic and periodic. Our general approach is inspired by [Kenig et al. 2014], which uses a two-scale expansion and relies on regularity estimates that are uniform in $\varepsilon > 0$. Following [Geng and Shen 2017], we consider the two-scale expansion

$$w_\varepsilon = u_\varepsilon(x,t)-u_0(x,t)-\varepsilon\chi(x/\varepsilon,t/\varepsilon^2)S_\varepsilon(\nabla u_0)-\varepsilon^2\phi(x/\varepsilon,t/\varepsilon^2)\nabla S_\varepsilon(\nabla u_0), \tag{1-19}$$

where $\chi(y,s)$ and $\phi(y,s)$ are correctors and dual correctors respectively for $\partial_t + \mathcal{L}_\varepsilon$ (see Section 2 for their definitions). In (1-19) the operator S_ε is a parabolic smoothing operator at scale ε. In comparison with the elliptic case, an extra term is added in the right-hand side of (1-19). This modification allows us to show that if $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = (\partial_t + \mathcal{L}_0)u_0$, then

$$(\partial_t + \mathcal{L}_\varepsilon)w_\varepsilon = \varepsilon \operatorname{div}(F_\varepsilon) \tag{1-20}$$

for some function F_ε, which depends only on u_0. As a consequence, we may apply the uniform interior L^∞ estimates established in [Geng and Shen 2015] to the function w_ε. To fully utilize the ideas above, we will consider the functions

$$\begin{aligned} u_\varepsilon(x,t) &= \int_{-\infty}^{t}\int_{\mathbb{R}^d} \Gamma_\varepsilon(x,t;y,s) f(y,s) e^{-\psi(y)}\,dy\,ds, \\ u_0(x,t) &= \int_{-\infty}^{t}\int_{\mathbb{R}^d} \Gamma_0(x,t;y,s) f(y,s) e^{-\psi(y)}\,dy\,ds, \end{aligned} \tag{1-21}$$

where ψ is a Lipschitz function in $\mathbb{R}^d$ and $f \in C_0^\infty(Q_r(y_0,s_0);\mathbb{R}^m)$. The main technical step in proving Theorem 1.1 involves bounding the L^∞ norm

$$\|e^{\psi}(u_\varepsilon - u_0)\|_{L^\infty(Q_r(x_0,t_0))} \tag{1-22}$$

by $\|f\|_{L^2(Q_r(y_0,s_0))}$, where $0 < \varepsilon < r = c\sqrt{t_0 - s_0}$. We remark that the use of weighted inequalities with weight e^{ψ} to generate the exponential factor in the Gaussian bound is more or less well known. Our approach may be regarded as a variation of the standard one found in [Hofmann and Kim 2004; Cho et al. 2008]; also see earlier work in [Fabes and Stroock 1986; Davies 1987a; Davies 1987b].

The proof of Theorem 1.2 uses the estimate in Theorem 1.1. The stronger assumption that A is Hölder continuous allows us to apply the uniform interior Lipschitz estimate obtained in [Geng and Shen 2015] to the function w_ε in (1-19). To see Theorem 1.3, one uses the fact that as a function of (x,t), $\nabla_y \Gamma_\varepsilon(x,t;y,s)$ is a solution of $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = 0$, away from the pole (y,s).

We end this section with some notation that will be used throughout the paper. A function $h = h(y,s)$ in $\mathbb{R}^{d+1}$ is said to be 1-periodic if h is periodic with respect to $\mathbb{Z}^{d+1}$. We will use the notation

$$\fint_E f = \frac{1}{|E|}\int_E f \quad \text{and} \quad h^\varepsilon(x,t) = h(x/\varepsilon, t/\varepsilon^2)$$

for $\varepsilon > 0$, as well as the summation convention that the repeated indices are summed. Finally, we shall use κ to denote positive constants that depend only on μ, and C constants that depend at most on d, m, μ and the smoothness of A, but never on ε.

2. Preliminaries

Let $\mathcal{L}_\varepsilon = -\operatorname{div}(A^\varepsilon(x,t)\nabla)$, where $A^\varepsilon(x,t) = A(x/\varepsilon, t/\varepsilon^2)$. Assume that $A(y,s)$ is 1-periodic in (y,s) and satisfies the ellipticity condition (1-2). For $1 \le j \le d$ and $1 \le \beta \le m$, the corrector $\chi_j^\beta = \chi_j^\beta(y,s) =$

$(\chi_j^{\alpha\beta}(y,s))$ is defined as the weak solution of the cell problem

$$\begin{cases} (\partial_s + \mathcal{L}_1)(\chi_j^\beta) = -\mathcal{L}_1(P_j^\beta) \quad \text{in } Y, \\ \chi_j^\beta = \chi_j^\beta(y,s) \text{ is 1-periodic in } (y,s), \\ \int_Y \chi_j^\beta = 0, \end{cases} \tag{2-1}$$

where $Y = [0,1)^{d+1}$, $P_j^\beta(y) = y_j e^\beta$, and $e^\beta = (0, \dots, 1, \dots, 0)$ with 1 in the β-th position. Note that

$$(\partial_s + \mathcal{L}_1)(\chi_j^\beta + P_j^\beta) = 0 \quad \text{in } \mathbb{R}^{d+1}. \tag{2-2}$$

By the rescaling property of $\partial_t + \mathcal{L}_\varepsilon$, one obtains

$$(\partial_t + \mathcal{L}_\varepsilon)\{\varepsilon\chi_j^\beta(x/\varepsilon, t/\varepsilon^2) + P_j^\beta(x)\} = 0 \quad \text{in } \mathbb{R}^{d+1}. \tag{2-3}$$

We say $A \in \mathrm{VMO}_x$ if

$$\lim_{r\to 0} A^\#(r) = 0, \tag{2-4}$$

where

$$A^\#(r) = \sup_{\substack{0<\rho<r \\ (x,t)\in\mathbb{R}^{d+1}}} \fint_{t-\rho^2}^t \fint_{y\in B(x,\rho)} \fint_{z\in B(x,\rho)} |A(y,s) - A(z,s)|\, dz\, dy\, ds. \tag{2-5}$$

Observe that if $A(y,s)$ is continuous in the variable y, uniformly in (y,s), then $A \in \mathrm{VMO}_x$.

Lemma 2.1. *Assume that $A(y,s)$ is 1-periodic in (y,s) and satisfies (1-2). If $m \geq 2$, we also assume $A \in \mathrm{VMO}_x$. Then $\chi_j^\beta \in L^\infty(Y; \mathbb{R}^m)$.*

Proof. In the scalar case $m = 1$, this follows from (2-2) by Nash's classical estimate. Moreover, the estimate

$$\left(\fint_{Q_r(x,t)} |\nabla \chi_j^\beta|^2 \right)^{1/2} \leq C r^{\sigma-1} \tag{2-6}$$

holds for any $0 < r < 1$ and $(x,t) \in \mathbb{R}^{d+1}$, where $Q_r(x,t) = B(x,r) \times (t - r^2, t)$, and $C > 0$ and $\sigma \in (0,1)$ depend on d and μ. If $m \geq 2$ and $A \in \mathrm{VMO}_x$, the boundedness of χ_j^β follows from the interior $W^{1,p}$ estimates for local solutions of $(\partial_t + \mathcal{L}_1)(u) = \mathrm{div}(f)$ [Byun 2007; Krylov 2007]. In this case the estimate (2-6) holds for any $\sigma \in (0,1)$. □

Let $\hat{A} = (\hat{a}_{ij}^{\alpha\beta})$, where $1 \leq i, j \leq d$, $1 \leq \alpha, \beta \leq m$, and

$$\hat{a}_{ij}^{\alpha\beta} = \fint_Y \left[a_{ij}^{\alpha\beta} + a_{ik}^{\alpha\gamma} \frac{\partial}{\partial y_k} \chi_j^{\gamma\beta} \right]; \tag{2-7}$$

that is

$$\hat{A} = \fint_Y \{A + A\nabla\chi\}.$$

It is known that the constant matrix $\hat{A}$ satisfies the ellipticity condition

$$\mu|\xi|^2 \leq \hat{a}_{ij}^{\alpha\beta} \xi_i^\alpha \xi_j^\beta \quad \text{for any } \xi = (\xi_j^\beta) \in \mathbb{R}^{m\times d}, \tag{2-8}$$

and $|\hat{a}_{ij}^{\alpha\beta}| \le \mu_1$, where $\mu_1 > 0$ depends only on d, m and μ [Bensoussan et al. 1978]. Define

$$\mathcal{L}_0 = -\operatorname{div}(\hat{A}\nabla).$$

Then $\partial_t + \mathcal{L}_0$ is the homogenized operator for the family of parabolic operators $\partial_t + \mathcal{L}_\varepsilon$, $\varepsilon > 0$.

To introduce the dual correctors, we consider the 1-periodic matrix-valued function

$$B = A + A\nabla\chi - \hat{A}. \tag{2-9}$$

More precisely, $B = B(y,s) = (b_{ij}^{\alpha\beta})$, where $1 \le i, j \le d$, $1 \le \alpha, \beta \le m$, and

$$b_{ij}^{\alpha\beta} = a_{ij}^{\alpha\beta} + a_{ik}^{\alpha\gamma}\frac{\partial \chi_j^{\gamma\beta}}{\partial y_k} - \hat{a}_{ij}^{\alpha\beta}. \tag{2-10}$$

Lemma 2.2. *Let $1 \le j \le d$ and $1 \le \alpha, \beta \le m$. Then there exist 1-periodic functions $\phi_{kij}^{\alpha\beta}(y,s)$ in $\mathbb{R}^{d+1}$ such that $\phi_{kij}^{\alpha\beta} \in H^1(Y)$,*

$$b_{ij}^{\alpha\beta} = \frac{\partial}{\partial y_k}(\phi_{kij}^{\alpha\beta}) \quad \text{and} \quad \phi_{kij}^{\alpha\beta} = -\phi_{ikj}^{\alpha\beta}, \tag{2-11}$$

where $1 \le k, i \le d+1$, $b_{ij}^{\alpha\beta}$ is defined by (2-10) for $1 \le i \le d$, $b_{(d+1)j}^{\alpha\beta} = -\chi_j^{\alpha\beta}$, and we have used the notation $y_{d+1} = s$.

Proof. This lemma was proved in [Geng and Shen 2015]. We give a proof here for reader's convenience. By (2-1) and (2-7), $b_{ij}^{\alpha\beta} \in L^2(Y)$ and

$$\int_Y b_{ij}^{\alpha\beta} = 0 \tag{2-12}$$

for $1 \le i \le d+1$. It follows that there exist $f_{ij}^{\alpha\beta} \in H^2(Y)$ such that

$$\begin{aligned} &\Delta_{d+1} f_{ij}^{\alpha\beta} = b_{ij}^{\alpha\beta} \quad \text{in } \mathbb{R}^{d+1}, \\ &f_{ij}^{\alpha\beta} \text{ is 1-periodic in } \mathbb{R}^{d+1}, \end{aligned} \tag{2-13}$$

where Δ_{d+1} denotes the Laplacian in $\mathbb{R}^{d+1}$. Write

$$b_{ij}^{\alpha\beta} = \frac{\partial}{\partial y_k}\left\{\frac{\partial}{\partial y_k} f_{ij}^{\alpha\beta} - \frac{\partial}{\partial y_i} f_{kj}^{\alpha\beta}\right\} + \frac{\partial}{\partial y_i}\left\{\frac{\partial}{\partial y_k} f_{kj}^{\alpha\beta}\right\}, \tag{2-14}$$

where the index k is summed from 1 to $d+1$. Note that by (2-1),

$$\sum_{i=1}^{d+1} \frac{\partial b_{ij}^{\alpha\beta}}{\partial y_i} = \sum_{i=1}^{d} \frac{\partial}{\partial y_i} b_{ij}^{\alpha\beta} - \frac{\partial}{\partial s}\chi_j^{\alpha\beta} = 0. \tag{2-15}$$

In view of (2-13) this implies

$$\sum_{i=1}^{d+1} \frac{\partial}{\partial y_i} f_{ij}^{\alpha\beta}$$

is harmonic in $\mathbb{R}^{d+1}$. Since it is 1-periodic, it must be constant. Consequently, by (2-14), we obtain

$$b_{ij}^{\alpha\beta} = \frac{\partial}{\partial y_k}(\phi_{kij}^{\alpha\beta}), \tag{2-16}$$

where

$$\phi_{kij}^{\alpha\beta} = \frac{\partial}{\partial y_k} f_{ij}^{\alpha\beta} - \frac{\partial}{\partial y_i} f_{kj}^{\alpha\beta} \tag{2-17}$$

is 1-periodic and belongs to $H^1(Y)$. It is easy to see that $\phi_{kij}^{\alpha\beta} = -\phi_{ikj}^{\alpha\beta}$. □

The 1-periodic functions $(\phi_{kij}^{\alpha\beta})$, given by Lemma 2.2, are called dual correctors for the family of parabolic operators $\partial_t + \mathcal{L}_\varepsilon$, $\varepsilon > 0$.

Lemma 2.3. *Let $\phi = (\phi_{kij}^{\alpha\beta})$ be the dual correctors, given by Lemma 2.2. Under the same assumptions as in Lemma 2.1, one has $\phi_{kij}^{\alpha\beta} \in L^\infty(Y)$.*

Proof. It follows from (2-6) that if $(x, t) \in \mathbb{R}^{d+1}$ and $0 < r < 1$,

$$\int_{Q_r(x,t)} |b_{ij}^{\alpha\beta}|^2 \le C r^{d+2\sigma} \tag{2-18}$$

for some $\sigma \in (0, 1)$. By covering the interval $(t - r, t)$ with intervals of length r^2, we obtain

$$\int_{B_r(x,t)} |b_{ij}^{\alpha\beta}|^2 \le C r^{d-1+2\sigma},$$

where $B_r(x, t) = B(x, r) \times (t - r, t)$. Hence, by Hölder's inequality,

$$\int_{B_r(x,t)} |b_{ij}^{\alpha\beta}| \le C r^{d+\sigma}.$$

Thus, for any $(x, t) \in Y$,

$$\int_Y \frac{|b_{ij}^{\alpha\beta}(y, s)|}{(|x - y| + |t - s|)^d}\, dy\, ds \le C \sum_{j=1}^{\infty} 2^{jd} \int_{|y-x|+|t-s|\sim 2^{-j}} |b_{ij}^{\alpha\beta}(y, s)|\, dy\, ds \le C. \tag{2-19}$$

In view of (2-13), by using the fundamental solution for Δ_{d+1} in $\mathbb{R}^{d+1}$, we may show that

$$\|\nabla_{y,s} f_{ij}^{\alpha\beta}\|_{L^\infty(Y)} \le C\|\nabla_{y,s} f_{ij}^{\alpha\beta}\|_{L^2(Y)} + \sup_{(x,t)\in Y} \int_Y \frac{|b_{ij}^{\alpha\beta}(y, s)|}{(|x - y| + |t - s|)^d}\, dy\, ds,$$

where $\nabla_{y,s}$ denotes the gradient in $\mathbb{R}^{d+1}$. This, together with (2-19), shows that $|\nabla_{y,s} f_{ij}^{\alpha\beta}| \in L^\infty(Y)$. By (2-17) we obtain $\phi_{kij}^{\alpha\beta} \in L^\infty(Y)$. □

Remark 2.4. Suppose $A = A(y, s)$ is Hölder continuous in (y, s). By (2-2) and the standard regularity theory for $\partial_s + \mathcal{L}_1$, we have $\nabla\chi(y, s)$ is Hölder continuous in (y, s). It follows that $b_{ij}^{\alpha\beta}(y, s)$ is Hölder continuous in (y, s). In view of (2-13) and (2-17) one may deduce that $\nabla_{y,s}\phi_{kij}^{\alpha\beta}$ is Hölder continuous in (y, s). This will be used in the proof of Theorems 1.2 and 1.3.

Theorem 2.5. *Suppose that A satisfies the conditions (1-2) and (1-3). If $m \ge 2$, we also assume $A \in \mathrm{VMO}_x$. Let u_ε be a weak solution of $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = \mathrm{div}(f)$ in $Q_{2r} = Q_{2r}(x_0, t_0)$ for some $0 < r < \infty$, where $f = (f_i^\alpha) \in L^p(Q_{2r}; \mathbb{R}^{m\times d})$ for some $p > d + 2$. Then*

$$\|u_\varepsilon\|_{L^\infty(Q_r)} \le C\left\{\left(\fint_{Q_{2r}} |u_\varepsilon|^2\right)^{1/2} + r\left(\fint_{Q_{2r}} |f|^p\right)^{1/p}\right\}, \tag{2-20}$$

where C depends only on d, m, p, μ, and $A^\#$ in (2-5) (if $m \ge 2$).

Proof. If $m=1$, this follows from the well-known Nash's estimate. The periodicity is not needed. If $m \geq 2$, (2-20) follows from the uniform interior Hölder estimate in [Geng and Shen 2015, Theorem 1.1]. □

Under the assumptions on A in Theorem 2.5, the matrix of fundamental solutions for $\partial_t + \mathcal{L}_\varepsilon$ in $\mathbb{R}^{d+1}$ exists and satisfies the Gaussian estimate (1-6). This follows from the L^∞ estimate (2-20) by a general result in [Hofmann and Kim 2004]; also see [Auscher 1996; Cho et al. 2008].

Theorem 2.6. *Suppose that A satisfies conditions (1-2) and (1-3). Also assume that A satisfies the Hölder condition (1-9). Let u_ε be a weak solution of $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = F$ in $Q_{2r} = Q_{2r}(x_0, t_0)$ for some $0 < r < \infty$, where $F \in L^p(Q_{2r}; \mathbb{R}^m)$ for some $p > d+2$. Then*

$$\|\nabla u_\varepsilon\|_{L^\infty(Q_r)} \leq C\left\{\frac{1}{r}\left(\fint_{Q_{2r}} |u_\varepsilon|^2\right)^{1/2} + r\left(\fint_{Q_{2r}} |F|^p\right)^{1/p}\right\}, \tag{2-21}$$

where C depends only on d, m, p, μ, and (λ, τ) in (1-9).

Proof. This was proved in [Geng and Shen 2015, Theorem 1.2]. □

The Lipschitz estimate (2-21) allows us to bound $\nabla_x \Gamma_\varepsilon(x,t;y,s)$, $\nabla_y \Gamma_\varepsilon(x,t;y,s)$ and $\nabla_x\nabla_y \Gamma_\varepsilon(x,t;y,s)$.

Theorem 2.7. *Assume that A satisfies the same conditions as in Theorem 2.6. Then*

$$|\nabla_x \Gamma_\varepsilon(x,t;y,s)| + |\nabla_y \Gamma_\varepsilon(x,t;y,s)| \leq \frac{C}{(t-s)^{(d+1)/2}} \exp\left\{-\frac{\kappa|x-y|^2}{t-s}\right\}, \tag{2-22}$$

$$|\nabla_x \nabla_y \Gamma_\varepsilon(x,t;y,s)| \leq \frac{C}{(t-s)^{(d+2)/2}} \exp\left\{-\frac{\kappa|x-y|^2}{t-s}\right\} \tag{2-23}$$

for any $x, y \in \mathbb{R}^d$ and $-\infty < s < t < \infty$, where $\kappa > 0$ depends only on μ. The constant C depends on d, m, μ, and (λ, τ) in (1-9).

Proof. Fix $x_0, y_0 \in \mathbb{R}^d$ and $s_0 < t_0$. Let $u_\varepsilon(x,t) = \Gamma_\varepsilon(x,t;y_0,s_0)$. Then $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = 0$ in $Q_{2r}(x_0,t_0)$, where $r = \sqrt{t_0 - s_0}/8$. The estimate for $|\nabla_x \Gamma(x_0,t_0;y_0,s_0)|$ now follows from (2-21) and (1-6) (with a different κ). In view of (1-12) this also gives the estimate for $|\nabla_y \Gamma_\varepsilon(x_0,t_0;y_0,s_0)|$. Finally, to see (2-23), we let $v_\varepsilon(x,t) = \nabla_y \Gamma_\varepsilon(x,t;y_0,s_0)$. Then $(\partial_t + \mathcal{L}_\varepsilon)v_\varepsilon = 0$ in $Q_{2r}(x_0,t_0)$. By applying (2-21) to v_ε and using the estimate in (2-22) for $\nabla_y \Gamma_\varepsilon(x,t;y,s)$, we obtain the desired estimate for $|\nabla_x \nabla_y \Gamma_\varepsilon(x_0,t_0;y_0,s_0)|$. □

3. A two-scale expansion

Suppose that

$$(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = (\partial_t + \mathcal{L}_0)u_0 \tag{3-1}$$

in $\Omega \times (T_0, T_1)$, where $\Omega \subset \mathbb{R}^d$. Let S_ε be a linear operator to be chosen later. Following [Geng and Shen 2017], we consider the two-scale expansion $w_\varepsilon = (w_\varepsilon^\alpha)$, where

$$w_\varepsilon^\alpha(x,t) = u_\varepsilon^\alpha(x,t) - u_0^\alpha(x,t) - \varepsilon\chi_j^{\alpha\beta}(x/\varepsilon, t/\varepsilon^2) S_\varepsilon\left(\frac{\partial u_0^\beta}{\partial x_j}\right) - \varepsilon^2 \phi_{(d+1)ij}^{\alpha\beta}(x/\varepsilon, t/\varepsilon^2) \frac{\partial}{\partial x_i} S_\varepsilon\left(\frac{\partial u_0^\beta}{\partial x_j}\right), \tag{3-2}$$

and $\chi_j^{\alpha\beta}$, $\phi_{(d+1)ij}^{\alpha\beta}$ are the correctors and dual correctors introduced in the last section. The repeated indices i, j in (3-2) are summed from 1 to d.

Proposition 3.1. *Let $u_\varepsilon \in L^2(T_0, T_1; H^1(\Omega))$ and $u_0 \in L^2(T_0, T_1; H^2(\Omega))$. Let w_ε be defined by (3-2). Assume (3-1) holds in $\Omega \times (T_0, T_1)$. Then*

$$(\partial_t + \mathcal{L}_\varepsilon) w_\varepsilon = \varepsilon \operatorname{div}(F_\varepsilon) \tag{3-3}$$

in $\Omega \times (T_0, T_1)$, where $F_\varepsilon = (F^\alpha_{\varepsilon,i})$ and

$$\begin{aligned} F^\alpha_{\varepsilon,i}(x,t) &= \varepsilon^{-1}(a^{\alpha\beta}_{ij}(x/\varepsilon, t/\varepsilon^2) - \hat{a}^{\alpha\beta}_{ij})\left(\frac{\partial u^\beta_0}{\partial x_j} - S_\varepsilon\left(\frac{\partial u^\beta_0}{\partial x_j}\right)\right) \\ &\quad + a^{\alpha\beta}_{ij}(x/\varepsilon, t/\varepsilon^2)\chi^{\beta\gamma}_k(x/\varepsilon, t/\varepsilon^2)\frac{\partial}{\partial x_j} S_\varepsilon\left(\frac{\partial u^\gamma_0}{\partial x_k}\right) \\ &\quad + \phi^{\alpha\beta}_{ikj}(x/\varepsilon, t/\varepsilon^2)\frac{\partial}{\partial x_k} S_\varepsilon\left(\frac{\partial u^\beta_0}{\partial x_j}\right) + \varepsilon\phi^{\alpha\beta}_{i(d+1)j}(x/\varepsilon, t/\varepsilon^2)\partial_t S_\varepsilon\left(\frac{\partial u^\beta_0}{\partial x_j}\right) \\ &\quad - a^{\alpha\beta}_{ij}(x/\varepsilon, t/\varepsilon^2)\left(\frac{\partial}{\partial x_j}(\phi^{\beta\gamma}_{(d+1)\ell k})\right)(x/\varepsilon, t/\varepsilon^2)\frac{\partial}{\partial x_\ell} S_\varepsilon\left(\frac{\partial u^\gamma_0}{\partial x_k}\right) \\ &\quad - \varepsilon a^{\alpha\beta}_{ij}(x/\varepsilon, t/\varepsilon^2)\phi^{\beta\gamma}_{(d+1)\ell k}(x/\varepsilon, t/\varepsilon^2)\frac{\partial^2}{\partial x_j \partial x_\ell} S_\varepsilon\left(\frac{\partial u^\gamma_0}{\partial x_k}\right). \end{aligned} \tag{3-4}$$

The repeated indices i, j, k, ℓ are summed from 1 to d.

Proof. This proposition was proved in [Geng and Shen 2017, Theorem 2.2]. □

We now introduce a parabolic smoothing operator. Let

$$\mathcal{O} = \{(x,t) \in \mathbb{R}^{d+1} : |x|^2 + |t| \le 1\}.$$

Fix a nonnegative function $\theta = \theta(x,t) \in C^\infty_0(\mathcal{O})$ such that $\int_{\mathbb{R}^{d+1}} \theta = 1$. Let $\theta_\varepsilon(x,t) = \varepsilon^{-d-2}\theta(x/\varepsilon, t/\varepsilon^2)$. Define

$$S_\varepsilon(f)(x,t) = f * \theta_\varepsilon(x,t) = \int_{\mathbb{R}^{d+1}} f(x-y, t-s)\theta_\varepsilon(y,s)\,dy\,ds. \tag{3-5}$$

Lemma 3.2. *Let $g = g(x,t)$ be a 1-periodic function in (x,t) and $\psi = \psi(x)$ a bounded Lipschitz function in $\mathbb{R}^d$. Then*

$$\|e^\psi g^\varepsilon S_\varepsilon(f)\|_{L^p(\mathbb{R}^{d+1})} \le C\, e^{\varepsilon\|\nabla\psi\|_\infty}\|g\|_{L^p(Y)}\|e^\psi f\|_{L^p(\mathbb{R}^{d+1})} \tag{3-6}$$

for any $1 \le p < \infty$, where $g^\varepsilon(x,t) = g(x/\varepsilon, t/\varepsilon^2)$ and C depends only on d and p.

Proof. Using $\int_{\mathbb{R}^{d+1}} \theta_\varepsilon = 1$ and Hölder's inequality, we obtain

$$|S_\varepsilon(e^{-\psi} f)(x,t)|^p \le \int_{\mathbb{R}^{d+1}} |e^{-\psi(y)} f(y,s)|^p\, \theta_\varepsilon(x-y, t-s)\,dy\,ds.$$

It follows that

$$\begin{aligned} |e^{\psi(x)} S_\varepsilon(e^{-\psi} f)(x,t)|^p &\le \int_{\mathbb{R}^{d+1}} |e^{\psi(x)-\psi(y)} f(y,s)|^p\, \theta_\varepsilon(x-y, t-s)\,dy\,ds \\ &\le e^{\varepsilon p\|\nabla\psi\|_\infty} \int_{\mathbb{R}^{d+1}} |f(y,s)|^p\, \theta_\varepsilon(x-y, t-s)\,dy\,ds, \end{aligned}$$

where we have used the facts that $|\psi(x)-\psi(y)| \le \|\nabla\psi\|_\infty |x-y|$ and $\theta_\varepsilon(x-y,t-s)=0$ if $|x-y|>\varepsilon$, for the last step. Hence, by Fubini's theorem,

$$\begin{aligned}\int_{\mathbb{R}^{d+1}} &|g^\varepsilon(x,t)|^p |e^\psi S_\varepsilon(e^{-\psi} f)(x,t)|^p\,dx\,dt \\ &\le e^{\varepsilon p\|\nabla\psi\|_\infty} \sup_{(y,s)\in\mathbb{R}^{d+1}} \int_{\mathbb{R}^{d+1}} |g^\varepsilon(x,t)|^p \theta_\varepsilon(x-y,t-s)\,dx\,dt \int_{\mathbb{R}^{d+1}} |f(y,s)|^p\,dy\,ds \\ &\le C\, e^{\varepsilon p\|\nabla\psi\|_\infty} \|g\|^p_{L^p(Y)} \|f\|^p_{L^p(\mathbb{R}^{d+1})},\end{aligned}$$

where C depends only on d. This gives (3-6). □

Remark 3.3. Let $\Omega \subset \mathbb{R}^d$ and $(T_0,T_1)\subset\mathbb{R}$. Define

$$\Omega_\varepsilon = \{x\in\mathbb{R}^d : \operatorname{dist}(x,\Omega) < \varepsilon\}. \tag{3-7}$$

Observe that for $(x,t)\in\Omega\times(T_0,T_1)$, we have $S_\varepsilon(f)(x,t) = S_\varepsilon(f\eta_\varepsilon)(x,t)$, where $\eta_\varepsilon=\eta_\varepsilon(x,t)$ is the characteristic function of $\Omega_\varepsilon\times(T_0-\varepsilon^2, T_1+\varepsilon^2)$. By applying (3-6) to the function $f\eta_\varepsilon$, one may deduce that

$$\int_{T_0}^{T_1}\int_\Omega |e^\psi g^\varepsilon S_\varepsilon(f)|^p\,dx\,dt \le C e^{\varepsilon p\|\nabla\psi\|_\infty} \|g\|^p_{L^p(Y)} \int_{T_0-\varepsilon^2}^{T_1+\varepsilon^2}\int_{\Omega_\varepsilon} |e^\psi f|^p\,dx\,dt. \tag{3-8}$$

Using $\int_{\mathbb{R}^{d+1}} |\nabla\theta_\varepsilon|\,dx\,dt \le C\varepsilon^{-1}$, the same argument as in the proof of Lemma 3.2 also shows that

$$\int_{T_0}^{T_1}\int_\Omega |e^\psi g^\varepsilon \nabla S_\varepsilon(f)|^p\,dx\,dt \le C\varepsilon^{-p} e^{\varepsilon p\|\nabla\psi\|_\infty} \|g\|^p_{L^p(Y)} \int_{T_0-\varepsilon^2}^{T_1+\varepsilon^2}\int_{\Omega_\varepsilon} |e^\psi f|^p\,dx\,dt \tag{3-9}$$

for $1\le p<\infty$, where C depends only on d and p.

Lemma 3.4. *Let S_ε be defined as in (3-5). Let $1\le p<\infty$ and ψ be a bounded Lipschitz function in $\mathbb{R}^d$. Then for $\Omega\subset\mathbb{R}^d$ and $(T_0,T_1)\subset\mathbb{R}$,*

$$\int_{T_0}^{T_1}\int_\Omega |e^\psi (S_\varepsilon(\nabla f) - \nabla f)|^p\,dx\,dt \le C\varepsilon^p e^{\varepsilon p\|\nabla\psi\|_\infty} \int_{T_0-\varepsilon^2}^{T_1+\varepsilon^2}\int_{\Omega_\varepsilon} |e^\psi (|\nabla^2 f| + |\partial_t f|)|^p\,dx\,dt, \tag{3-10}$$

where Ω_ε is given by (3-7) and C depends only on d and p.

Proof. Write

$$S_\varepsilon(\nabla f)(x,t) - \nabla f(x,t) = J_1(x,t) + J_2(x,t),$$

where

$$\begin{aligned} J_1(x,t) &= \int_{\mathbb{R}^{d+1}} \theta_\varepsilon(y,s)(\nabla f(x-y,t-s) - \nabla f(x-y,t))\,dy\,ds, \\ J_2(x,t) &= \int_{\mathbb{R}^{d+1}} \theta_\varepsilon(y,s)(\nabla f(x-y,t) - \nabla f(x,t))\,dy,\,ds.\end{aligned}$$

To estimate J_2, we observe that by Hölder's inequality and the fact $\int_{\mathbb{R}^{d+1}} \theta_\varepsilon\,dy\,ds = 1$,

$$|J_2(x,t)|^p \le \int_{\mathbb{R}^{d+1}} \theta_\varepsilon(y,s) |\nabla f(x-y,t) - \nabla f(x,t)|^p\,dy\,ds,$$

and

$$|\nabla f(x-y,t)-\nabla f(x,t)| = \left|\int_0^1 \frac{\partial}{\partial\tau}\nabla f(x-\tau y,t)\,d\tau\right| \\ \le |y|\int_0^1 |\nabla^2 f(x-\tau y,t)|\,d\tau \le |y|\left(\int_0^1 |\nabla^2 f(x-\tau y,t)|^p\,d\tau\right)^{1/p}.$$

It follows by Fubini's theorem that

$$\begin{aligned}\int_{T_0}^{T_1}\int_\Omega |e^{\psi(x)}J_2(x,t)|^p\,dx\,dt \\ &\le \int_{T_0}^{T_1}\int_\Omega\int_{\mathbb{R}^{d+1}}\int_0^1 e^{p\psi(x)}\theta_\varepsilon(y,s)|y|^p|\nabla^2 f(x-\tau y,t)|^p\,d\tau\,dy\,ds\,dx\,dt \\ &\le \varepsilon^p e^{\varepsilon p\|\nabla\psi\|_\infty}\int_{T_0}^{T_1}\int_\Omega\int_{\mathbb{R}^{d+1}}\int_0^1 e^{p\psi(x-\tau y)}\theta_\varepsilon(y,s)|\nabla^2 f(x-\tau y,t)|^p\;d\tau\,dy\,ds\,dx\,dt \\ &\le \varepsilon^p e^{\varepsilon p\|\nabla\psi\|_\infty}\int_{T_0}^{T_1}\int_{\Omega_\varepsilon}|e^\psi\nabla^2 f|^p\,dx\,dt,\end{aligned}$$

where we have used the facts that $|\psi(x)-\psi(x-\tau y)|\le|\tau|\,|y|\,\|\nabla\psi\|_\infty$ and $\theta_\varepsilon(y,s)=0$ if $|y|>\varepsilon$.

Finally, to estimate J_1, we first use integration by parts to obtain

$$|J_1(x,t)| \le \int_{\mathbb{R}^{d+1}}|\nabla\theta_\varepsilon(y,s)|\,|f(x-y,t-s)-f(x-y,t)|\,dy\,ds.$$

By Hölder's inequality,

$$|J_1(x,t)|^p \le C\varepsilon^{1-p}\int_{\mathbb{R}^{d+1}}|\nabla\theta_\varepsilon(y,s)|\,|f(x-y,t-s)-f(x-y,t)|^p\,dy\,ds,$$

where we have also used the fact $\int_{\mathbb{R}^{d+1}}|\nabla\theta_\varepsilon|\,dy\,ds\le C\varepsilon^{-1}$. Using

$$|f(x-y,t-s)-f(x-y,t)| \le \left|\int_0^1 \frac{\partial}{\partial\tau}f(x-y,t-\tau s)\,d\tau\right| \le |s|\left(\int_0^1|\partial_t f(x-y,t-\tau s)|^p\,d\tau\right)^{1/p},$$

we see that by Fubini's theorem,

$$\begin{aligned}\int_{T_0}^{T_1}\int_\Omega |e^{\psi(x)}J_1(x,t)|^p\,dx\,dt \\ &\le C\varepsilon^{1-p}\int_{T_0}^{T_1}\int_\Omega\int_{\mathbb{R}^{d+1}}\int_0^1 e^{p\psi(x)}|\nabla\theta_\varepsilon(y,s)|\,|s|^p|\partial_t f(x-y,t-\tau s)|^p\,d\tau\,dy\,ds\,dx\,dt \\ &\le C\varepsilon^{1+p}e^{\varepsilon p\|\nabla\psi\|_\infty}\int_{T_0}^{T_1}\int_\Omega\int_{\mathbb{R}^{d+1}}\int_0^1 e^{p\psi(x-y)}|\nabla\theta_\varepsilon(y,s)|\,|\partial_t f(x-y,t-\tau s)|^p\,d\tau\,dy\,ds\,dx\,dt \\ &\le C\varepsilon^p e^{\varepsilon p\|\nabla\psi\|_\infty}\int_{T_0-\varepsilon^2}^{T_1+\varepsilon^2}\int_{\Omega_\varepsilon}|e^\psi\partial_t f|^p\,dx\,dt,\end{aligned}$$

where we have used the facts that $|\psi(x)-\psi(x-y)|\le\|\nabla\psi\|_\infty|y|$ and $\theta_\varepsilon(y,s)=0$ if $|y|>\varepsilon$ or $|s|>\varepsilon^2$. This, together with the estimate for J_2, completes the proof. □

Theorem 3.5. *Let $F_\varepsilon = (F^\alpha_{\varepsilon,i})$ be given by (3-4) and $1 \le p < \infty$. Then for any $\Omega \subset \mathbb{R}^d$ and $(T_0, T_1) \subset \mathbb{R}$,*

$$\int_{T_0}^{T_1}\int_\Omega |e^\psi F_\varepsilon|^p\,dx\,dt \le Ce^{\varepsilon p\|\nabla\psi\|_\infty}\int_{T_0-\varepsilon^2}^{T_1+\varepsilon^2}\int_{\Omega_\varepsilon}\{|e^\psi\nabla^2 u_0|^p + |e^\psi\partial_t u_0|^p\}\,dx\,dt, \tag{3-11}$$

where Ω_ε is given by (3-7) and C depends only on d, m, p and μ.

Proof. Observe that

$$\begin{aligned}
&\int_{T_0}^{T_1}\int_\Omega |e^\psi F_\varepsilon|^p\,dx\,dt\\
&\le C\varepsilon^{-p}\int_{T_0}^{T_1}\int_\Omega |\nabla u_0 - S_\varepsilon(\nabla u_0)|^p e^{p\psi}\,dx\,dt + C\int_{T_0}^{T_1}\int_\Omega |\chi^\varepsilon|^p|S_\varepsilon(\nabla^2 u_0)|^p e^{p\psi}\,dx\,dt\\
&\quad + C\int_{T_0}^{T_1}\int_\Omega |\phi^\varepsilon|^p|S_\varepsilon(\nabla^2 u_0)|^p e^{p\psi}\,dx\,dt + C\varepsilon^p\int_{T_0}^{T_1}\int_\Omega |\phi^\varepsilon|^p|\nabla S_\varepsilon(\partial_t u_0)|^p e^{p\psi}\,dx\,dt\\
&\quad + C\int_{T_0}^{T_1}\int_\Omega |(\nabla\phi)^\varepsilon|^p|S_\varepsilon(\nabla^2 u_0)|^p e^{p\psi}\,dx\,dt + C\varepsilon^p\int_{T_0}^{T_1}\int_\Omega |\phi^\varepsilon|^p|\nabla S_\varepsilon(\nabla^2 u_0)|^p e^{p\psi}\,dx\,dt,
\end{aligned} \tag{3-12}$$

where C depends only on d and μ. In (3-12) we have also used the observation that $\partial_t S_\varepsilon(\nabla u_0) = \nabla S_\varepsilon(\partial_t u_0)$ and $\nabla S_\varepsilon(\nabla u_0) = S_\varepsilon(\nabla^2 u_0)$.

We now proceed to bound each term in the right-hand side of (3-12), using Lemma 3.4 and Remark 3.3. By Lemma 3.4, the first term in the right-hand side of (3-12) is bounded by

$$Ce^{p\varepsilon\|\nabla\psi\|_\infty}\int_{T_0-\varepsilon^2}^{T_1+\varepsilon^2}\int_{\Omega_\varepsilon} |e^\psi(|\nabla^2 u_0| + |\partial_t u_0|)|^p\,dx\,dt. \tag{3-13}$$

Using (3-8) we may bound the second, third, fifth terms in the right-hand side of (3-12) by

$$Ce^{p\varepsilon\|\nabla\psi\|_\infty}\int_{T_0-\varepsilon^2}^{T_1+\varepsilon^2}\int_{\Omega_\varepsilon} |e^\psi\nabla^2 u_0|^p\,dx\,dt. \tag{3-14}$$

Finally, by (3-9), the fourth and sixth terms in the right-hand side of (3-12) are bounded by (3-13). □

4. Weighted estimates for $\partial_t + \mathcal{L}_0$

Recall that $\Gamma_0(x, t; y, s)$ denotes the matrix of fundamental solutions for the homogenized operator $\partial_t + \mathcal{L}_0$ in $\mathbb{R}^{d+1}$. Let $\psi : \mathbb{R}^d \to \mathbb{R}$ be a bounded Lipschitz function and

$$u_0(x, t) = \int_{-\infty}^t\int_{\mathbb{R}^d} \Gamma_0(x, t; y, s)f(y, s)e^{-\psi(y)}\,dy\,ds, \tag{4-1}$$

where $f \in C_0^\infty(\mathbb{R}^{d+1}; \mathbb{R}^m)$. Then

$$(\partial_t + \mathcal{L}_0)u_0 = e^{-\psi} f \quad \text{in } \mathbb{R}^{d+1}. \tag{4-2}$$

The goal of this section is to prove the following.

Theorem 4.1. *Let u_0 be defined by (4-1). Suppose that $f(x,t)=0$ for $t \le s_0$. Then*

$$\int_{s_0}^t \int_{\mathbb{R}^d} |e^{\psi}(|\nabla^2 u_0| + |\partial_t u_0|)|^2 \, dx \, dt \le C e^{\kappa(t-s_0)\|\nabla\psi\|_\infty^2} \int_{s_0}^t \int_{\mathbb{R}^d} |f|^2 \, dx \, dt \tag{4-3}$$

for any $s_0 < t < \infty$, where $\kappa > 0$ depends only on μ and C depends only on d and μ.

We start with an estimate on a lower-order term.

Lemma 4.2. *Let u_0 be defined by (4-1). Suppose that $f(x,t)=0$ for $t < s_0$. Then*

$$\int_{s_0}^t \int_{\mathbb{R}^d} |e^{\psi} \nabla u_0|^2 \, dx \, dt \le C(t-s_0) e^{\kappa_1(t-s_0)\|\nabla\psi\|_\infty^2} \int_{s_0}^t \int_{\mathbb{R}^d} |f|^2 \, dx \, dt \tag{4-4}$$

for any $s_0 < t < \infty$, where $\kappa_1 > 0$ depends only on μ and C depends only on d and μ.

Proof. It follows from (1-7) that for $x, y \in \mathbb{R}^d$ and $t > s$,

$$\begin{aligned} e^{\psi(x)-\psi(y)} |\nabla_x \Gamma_0(x,t;y,s)| &\le \frac{C}{(t-s)^{(d+1)/2}} \exp\left\{\psi(x)-\psi(y) - \frac{\kappa|x-y|^2}{t-s}\right\} \\ &\le \frac{C}{(t-s)^{(d+1)/2}} \exp\left\{\|\nabla\psi\|_\infty |x-y| - \frac{\kappa|x-y|^2}{t-s}\right\}. \end{aligned}$$

This, together with the inequality

$$\|\nabla\psi\|_\infty |x-y| \le \frac{(t-s)\|\nabla\psi\|_\infty^2}{2\kappa} + \frac{\kappa|x-y|^2}{2(t-s)}, \tag{4-5}$$

yields

$$e^{\psi(x)-\psi(y)} |\nabla_x \Gamma_0(x,t;y,s)| \le C e^{(t-s)\|\nabla\psi\|_\infty^2/(2\kappa)} \cdot \frac{1}{(t-s)^{(d+1)/2}} e^{-\kappa|x-y|^2/(2(t-s))}. \tag{4-6}$$

It follows that

$$\begin{aligned} |e^{\psi(x)} \nabla u_0(x,t)| &\le \int_{s_0}^t \int_{\mathbb{R}^d} e^{\psi(x)-\psi(y)} |\nabla_x \Gamma_0(x,t;y,s)| |f(y,s)| \, dy \, ds \\ &\le C e^{(t-s_0)\|\nabla\psi\|_\infty^2/(2\kappa)} \int_{s_0}^t \int_{\mathbb{R}^d} \frac{1}{(t-s)^{(d+1)/2}} e^{-\kappa|x-y|^2/(2(t-s))} |f(y,s)| \, dy \, ds \\ &\le C e^{(t-s_0)\|\nabla\psi\|_\infty^2/(2\kappa)} (t-s_0)^{1/4} \left(\int_{s_0}^t \int_{\mathbb{R}^d} \frac{1}{(t-s)^{(d+1)/2}} e^{-\kappa|x-y|^2/(2(t-s))} |f(y,s)|^2 \, dy \, ds \right)^{1/2}, \end{aligned}$$

where we have used Hölder's inequality for the last step. The estimate (4-4) now follows by Fubini's theorem. □

Proof of Theorem 4.1. In view of (4-2) we have

$$(\partial_t + \mathcal{L}_0) \frac{\partial u_0}{\partial x_k} = \frac{\partial}{\partial x_k} (e^{-\psi} f)$$

in $\mathbb{R}^{d+1}$. It follows that

$$\int_{\mathbb{R}^d} \partial_t \nabla u_0 \cdot (\nabla u_0) e^{2\psi}\, dx - \int_{\mathbb{R}^d} \hat{a}_{ij}^{\alpha\beta} \frac{\partial^3 u_0^\beta}{\partial x_i \partial x_j \partial x_k} \cdot \frac{\partial u_0^\alpha}{\partial x_k} e^{2\psi}\, dx = \int_{\mathbb{R}^d} \frac{\partial}{\partial x_k}(e^{-\psi} f^\alpha) \frac{\partial u_0^\alpha}{\partial x_k} e^{2\psi}\, dx.$$

Using integration by parts, we obtain

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}^d} |\nabla u_0|^2 e^{2\psi}\, dx + \int_{\mathbb{R}^d} \hat{a}_{ij}^{\alpha\beta} \frac{\partial^2 u_0^\beta}{\partial x_j \partial x_k} \cdot \frac{\partial^2 u_0^\alpha}{\partial x_i \partial x_k} e^{2\psi}\, dx$$
$$= -\int_{\mathbb{R}^d} f \cdot (\Delta u_0) e^{\psi}\, dx - \int_{\mathbb{R}^d} e^{-\psi} f^\alpha \frac{\partial u_0^\alpha}{\partial x_k} \frac{\partial e^{2\psi}}{\partial x_k}\, dx - \int_{\mathbb{R}^d} \hat{a}_{ij}^{\alpha\beta} \frac{\partial^2 u_0^\beta}{\partial x_j \partial x_k} \cdot \frac{\partial u_0^\alpha}{\partial x_k} \frac{\partial e^{2\psi}}{\partial x_i}\, dx.$$

By the ellipticity of $\mathcal{L}_0$, this yields

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}^d} |\nabla u_0|^2 e^{2\psi}\, dx + \mu \int_{\mathbb{R}^d} |\nabla^2 u_0|^2 e^{2\psi}\, dx$$
$$\le C\int_{\mathbb{R}^d} |f||\nabla^2 u_0| e^{\psi}\, dx + C\int_{\mathbb{R}^d} |f|^2\, dx + C\|\nabla\psi\|_\infty^2 \int_{\mathbb{R}^d} |\nabla u_0|^2 e^{2\psi}\, dx + C\|\nabla\psi\|_\infty \int_{\mathbb{R}^d} |\nabla^2 u_0||\nabla u_0| e^{2\psi}\, dx,$$

where C depends only on d and μ. Using the Cauchy inequality, we may further deduce that

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}^d} |\nabla u_0|^2 e^{2\psi}\, dx + \frac{\mu}{2}\int_{\mathbb{R}^d} |\nabla^2 u_0|^2 e^{2\psi}\, dx \le C\int_{\mathbb{R}^d} |f|^2\, dx + C\|\nabla\psi\|_\infty^2 \int_{\mathbb{R}^d} |\nabla u_0|^2 e^{2\psi}\, dx.$$

We now integrate the inequality above in t over the interval (s_0, s_1). This leads to

$$\frac{1}{2}\int_{\mathbb{R}^d} |\nabla u_0(x, s_1)|^2 e^{2\psi}\, dx + \frac{\mu}{2}\int_{s_0}^{s_1}\int_{\mathbb{R}^d} |\nabla^2 u_0|^2 e^{2\psi}\, dx\, dt$$
$$\le C\int_{s_0}^{s_1}\int_{\mathbb{R}^d} |f|^2\, dx\, dt + C\|\nabla\psi\|_\infty^2 \int_{s_0}^{s_1}\int_{\mathbb{R}^d} |\nabla u_0|^2 e^{2\psi}\, dx\, dt$$
$$\le C e^{\kappa(s_1 - s_0)\|\nabla\psi\|_\infty^2} \int_{s_0}^{s_1}\int_{\mathbb{R}^d} |f|^2\, dx\, dt, \tag{4-7}$$

where we have used (4-4) for the last inequality. Estimate (4-3) follows readily from (4-7). □

5. Proof of Theorem 1.1

We start with some weighted estimates.

Lemma 5.1. *Suppose that*

$$\begin{cases} (\partial_t + \mathcal{L}_\varepsilon) w_\varepsilon = \varepsilon \operatorname{div}(F_\varepsilon) & \text{in } \mathbb{R}^d \times (s_0, \infty), \\ w_\varepsilon = 0 & \text{on } \mathbb{R}^d \times \{t = s_0\}. \end{cases} \tag{5-1}$$

Let $\psi : \mathbb{R}^d \to \mathbb{R}$ be a bounded Lipschitz function. Then for any $t > s_0$

$$\int_{\mathbb{R}^d} |w_\varepsilon(x, t)|^2 e^{2\psi(x)}\, dx \le C\varepsilon^2 e^{\kappa(t - s_0)\|\nabla\psi\|_\infty^2} \int_{s_0}^{t}\int_{\mathbb{R}^d} |F_\varepsilon(x, s)|^2 e^{2\psi(x)}\, dx\, ds, \tag{5-2}$$

where $\kappa > 0$ and $C > 0$ depends only on μ.

Proof. Let

$$I(t)=\int_{\mathbb{R}^d}|w_\varepsilon(x,t)|^2 e^{2\psi(x)}\,dx. \tag{5-3}$$

Note that

$$\begin{aligned}
I'(t)&=2\int_{\mathbb{R}^d}\langle \partial_t w_\varepsilon, e^{2\psi}w_\varepsilon\rangle\,dx\\
&=-2\int_{\mathbb{R}^d}\langle \mathcal{L}_\varepsilon(w_\varepsilon), e^{2\psi}w_\varepsilon\rangle\,dx+2\varepsilon\int_{\mathbb{R}^d}\langle \operatorname{div}(F_\varepsilon), e^{2\psi}w_\varepsilon\rangle\,dx\\
&=-2\int_{\mathbb{R}^d}A^\varepsilon\nabla w_\varepsilon\cdot\nabla(e^{2\psi}w_\varepsilon)\,dx-2\varepsilon\int_{\mathbb{R}^d}F_\varepsilon\cdot\nabla(e^{2\psi}w_\varepsilon)\,dx\\
&=-2\int_{\mathbb{R}^d}A^\varepsilon\nabla w_\varepsilon\cdot(\nabla w_\varepsilon)e^{2\psi}\,dx-2\int_{\mathbb{R}^d}A^\varepsilon\nabla w_\varepsilon\cdot\nabla(e^{2\psi})w_\varepsilon\,dx\\
&\qquad -2\varepsilon\int_{\mathbb{R}^d}F_\varepsilon\cdot(\nabla w_\varepsilon)e^{2\psi}\,dx-2\varepsilon\int_{\mathbb{R}^d}F_\varepsilon\cdot\nabla(e^{2\psi})w_\varepsilon\,dx,
\end{aligned}$$

where $\langle\ ,\ \rangle$ denotes the pairing in $H^{-1}(\mathbb{R}^d;\mathbb{R}^m)\times H^1(\mathbb{R}^d;\mathbb{R}^m)$. It follows that

$$\begin{aligned}
I'(t)&\le -2\mu\int_{\mathbb{R}^d}|\nabla w_\varepsilon|^2 e^{2\psi}\,dx+\kappa\|\nabla\psi\|_\infty\int_{\mathbb{R}^d}|\nabla w_\varepsilon|\,|w_\varepsilon|e^{2\psi}\,dx\\
&\qquad +2\varepsilon\int_{\mathbb{R}^d}|\nabla w_\varepsilon|\,|F_\varepsilon|e^{2\psi}\,dx+4\varepsilon\|\nabla\psi\|_\infty\int_{\mathbb{R}^d}|w_\varepsilon|\,|F_\varepsilon|e^{2\psi}\,dx,
\end{aligned}$$

where $\kappa>0$ depends only on μ. By the Cauchy inequality this implies

$$I'(t)\le \kappa\|\nabla\psi\|_\infty^2 I(t)+\kappa\varepsilon^2\int_{\mathbb{R}^d}|F_\varepsilon(x,t)|^2e^{2\psi}\,dx, \tag{5-4}$$

where $\kappa>0$ depends only on μ. Hence,

$$\frac{d}{dt}\{I(t)e^{-\kappa(t-s_0)\|\nabla\psi\|_\infty^2}\}\le C\varepsilon^2 e^{-\kappa(t-s_0)\|\nabla\psi\|_\infty^2}\int_{\mathbb{R}^d}|F_\varepsilon(x,t)|^2e^{2\psi}\,dx.$$

Since $I(s_0)=0$, it follows that

$$\begin{aligned}
I(t)&\le C\varepsilon^2\int_{s_0}^t\int_{\mathbb{R}^d}e^{\kappa(t-s)\|\nabla\psi\|_\infty^2}|F_\varepsilon(x,s)|^2e^{2\psi}\,dx\,ds\\
&\le C\varepsilon^2e^{\kappa(t-s_0)\|\nabla\psi\|_\infty^2}\int_{s_0}^t\int_{\mathbb{R}^d}|F_\varepsilon(x,s)|^2e^{2\psi}\,dx\,ds. \qquad \square
\end{aligned}$$

Lemma 5.2. *Suppose that* $u_\varepsilon\in L^2((-\infty,T);H^1(\mathbb{R}^d))$ *and* $u_0\in L^2((-\infty,T);H^2(\mathbb{R}^d))$ *for any* $T\in\mathbb{R}$, *and that*

$$\begin{cases}(\partial_t+\mathcal{L}_\varepsilon)u_\varepsilon=(\partial_t+\mathcal{L}_0)u_0 & \text{in } \mathbb{R}^{d+1},\\ u_\varepsilon(x,t)=u_0(x,t)=0 & \text{for } t\le s_0.\end{cases}$$

Let w_ε be defined by (3-2), where the operator S_ε is given by (3-5). Then for any $t > s_0$,

$$\int_{\mathbb{R}^d} |w_\varepsilon(x,t)|^2 e^{2\psi(x)}\,dx \\ \leq C\varepsilon^2 e^{2\varepsilon\|\nabla\psi\|_\infty+\kappa(t-s_0)\|\nabla\psi\|_\infty^2} \int_{s_0}^{t+\varepsilon^2}\int_{\mathbb{R}^d} \{|\nabla^2 u_0(x,s)|^2 + |\partial_s u_0(x,s)|^2\} e^{2\psi(x)}\,dx\,ds, \quad (5\text{-}5)$$

where ψ is a bounded Lipschitz function in $\mathbb{R}^d$, κ depends only on μ, and C depends only on d, m and μ.

Proof. This follows readily from Lemma 5.1 and Theorem 3.5 with $p = 2$. □

The next theorem gives a weighted L^∞ estimate.

Theorem 5.3. *Assume that A is 1-periodic and satisfies (1-2). If $m \geq 2$, we also assume that $A \in \mathrm{VMO}_x$. Suppose that $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = (\partial_t + \mathcal{L}_0)u_0$ in $B(x_0, 3r) \times (t_0 - 5r^2, t_0 + r^2)$ for some $(x_0, t_0) \in \mathbb{R}^{d+1}$ and $\varepsilon \leq r < \infty$. Then*

$$\|e^\psi(u_\varepsilon - u_0)\|_{L^\infty(Q_r(x_0,t_0))} \leq Ce^{3r\|\nabla\psi\|_\infty}\left(\fint_{Q_{2r}(x_0,t_0)} |e^\psi(u_\varepsilon - u_0)|^2\right)^{1/2} \\ + C\varepsilon r e^{3r\|\nabla\psi\|_\infty}\|e^\psi(|\nabla^2 u_0| + |\partial_t u_0|)\|_{L^\infty(B(x_0,3r)\times(t_0-5r^2,t_0+r^2))} \\ + C\varepsilon e^{3r\|\nabla\psi\|_\infty}\|e^\psi \nabla u_0\|_{L^\infty(B(x_0,3r)\times(t_0-5r^2,t_0+r^2))}, \quad (5\text{-}6)$$

where ψ is a Lipschitz function in $\mathbb{R}^d$ and C depends only on d, m, μ and $A^\#$ (if $m \geq 2$).

Proof. Let w_ε be defined by (3-2). Then $(\partial_t + \mathcal{L}_\varepsilon)w_\varepsilon = \varepsilon\,\mathrm{div}(F_\varepsilon)$ in $Q_{2r}(x_0, t_0)$, where F_ε is given by (3-4). It follows by Theorem 2.5 that

$$\|w_\varepsilon\|_{L^\infty(Q_r(x_0,t_0))} \leq C\left\{\left(\fint_{Q_{2r}(x_0,t_0)} |w_\varepsilon|^2\right)^{1/2} + \varepsilon r\left(\fint_{Q_{2r}(x_0,t_0)} |F_\varepsilon|^p\right)^{1/p}\right\}, \quad (5\text{-}7)$$

where $p > d + 2$. This leads to

$$\|u_\varepsilon - u_0\|_{L^\infty(Q_r(x_0,t_0))} \leq C\left(\fint_{Q_{2r}(x_0,t_0)} |u_\varepsilon - u_0|^2\right)^{1/2} + C\varepsilon r\left(\fint_{Q_{2r}(x_0,t_0)} |F_\varepsilon|^p\right)^{1/p} \\ + C\varepsilon\|S_\varepsilon(\nabla u_0)\|_{L^\infty(Q_{2r}(x_0,t_0))} + C\varepsilon^2\|S_\varepsilon(\nabla^2 u_0)\|_{L^\infty(Q_{2r}(x_0,t_0))},$$

where we have used the boundedness of χ and ϕ in Lemmas 2.1 and 2.3. Hence, using $|\psi(x) - \psi(y)| \leq 2r\|\nabla\psi\|_\infty$ for $x, y \in B(x_0, 2r)$, we obtain

$$\|e^\psi(u_\varepsilon - u_0)\|_{L^\infty(Q_r(x_0,t_0))} \\ \leq Ce^{2r\|\nabla\psi\|_\infty}\left(\fint_{Q_{2r}(x_0,t_0)} |e^\psi(u_\varepsilon - u_0)|^2\right)^{1/2} + C\varepsilon r e^{2r\|\nabla\psi\|_\infty}\left(\fint_{Q_{2r}(x_0,t_0)} |e^\psi F_\varepsilon|^p\right)^{1/p} \\ + C\varepsilon e^{2r\|\nabla\psi\|_\infty}\|e^\psi S_\varepsilon(\nabla u_0)\|_{L^\infty(Q_{2r}(x_0,t_0))} + C\varepsilon^2 e^{2r\|\nabla\psi\|_\infty}\|e^\psi S_\varepsilon(\nabla^2 u_0)\|_{L^\infty(Q_{2r}(x_0,t_0))}. \quad (5\text{-}8)$$

Finally, we use Theorem 3.5 to bound the second term in the right-hand side of (5-8). This yields

$$\begin{aligned}&\|e^{\psi}(u_\varepsilon-u_0)\|_{L^\infty(Q_r(x_0,t_0))}\\ &\le Ce^{2r\|\nabla\psi\|_\infty}\biggl(\fint_{Q_{2r}(x_0,t_0)}|e^{\psi}(u_\varepsilon-u_0)|^2\biggr)^{1/2}+C\varepsilon re^{3r\|\nabla\psi\|_\infty}\biggl(\fint_{t_0-5r^2}^{t_0+r^2}\fint_{B(x_0,3r)}\{|e^{\psi}\nabla^2u_0|^p+|e^{\psi}\partial_t u_0|^p\}\biggr)^{1/p}\\ &\qquad+C\varepsilon e^{2r\|\nabla\psi\|_\infty}\|e^{\psi}S_\varepsilon(\nabla u_0)\|_{L^\infty(Q_{2r}(x_0,t_0))}+C\varepsilon^2e^{2r\|\nabla\psi\|_\infty}\|e^{\psi}S_\varepsilon(\nabla^2u_0)\|_{L^\infty(Q_{2r}(x_0,t_0))},\end{aligned}$$

where $p>d+2$ and we also used the assumption $\varepsilon\le r$. Estimate (5-6) now follows. □

We are now in a position to give the proof of Theorem 1.1.

Proof of Theorem 1.1. We begin by fixing $x_0, y_0\in\mathbb{R}^{d+1}$ and $s_0<t_0$. We may assume that

$$\varepsilon<r=(t_0-s_0)^{1/2}/100.$$

For otherwise the desired estimate (1-8) follows directly from (1-6).

For $f\in C_0^\infty(Q_r(y_0,s_0);\mathbb{R}^m)$, define

$$\begin{aligned}u_\varepsilon(x,t)&=\int_{-\infty}^{t}\int_{\mathbb{R}^d}e^{-\psi(y)}\Gamma_\varepsilon(x,t;y,s)f(y,s)\,dy\,ds,\\ u_0(x,t)&=\int_{-\infty}^{t}\int_{\mathbb{R}^d}e^{-\psi(y)}\Gamma_0(x,t;y,s)f(y,s)\,dy\,ds,\end{aligned}$$

where ψ is a bounded Lipschitz function in $\mathbb{R}^d$ to be chosen later. Then

$$(\partial_t+\mathcal{L}_\varepsilon)u_\varepsilon=(\partial_t+\mathcal{L}_0)u_0=e^{-\psi}f\quad\text{in }\mathbb{R}^{d+1}$$

and $u_\varepsilon(x,t)=u_0(x,t)=0$ if $t\le s_0$. Let w_ε be defined by (3-2). It follows from Lemma 5.2 and Theorem 4.1 that

$$\int_{\mathbb{R}^d}|w_\varepsilon(x,t)|^2e^{2\psi(x)}\,dx\le C\varepsilon^2e^{2\varepsilon\|\nabla\psi\|_\infty+\kappa(t-s_0+\varepsilon^2)\|\nabla\psi\|_\infty^2}\int_{s_0}^{t+\varepsilon^2}\int_{\mathbb{R}^d}|f|^2\,dx\,ds \tag{5-9}$$

for any $t>s_0$.

Next, we use (5-6) to obtain

$$\begin{aligned}\|e^{\psi}(u_\varepsilon-u_0)\|_{L^\infty(Q_r(x_0,t_0))}&\le Ce^{3r\|\nabla\psi\|_\infty}\biggl(\fint_{Q_{2r}(x_0,t_0)}|e^{\psi}w_\varepsilon|^2\biggr)^{1/2}\\ &\quad+C\varepsilon re^{3r\|\nabla\psi\|_\infty}\|e^{\psi}(|\nabla^2u_0|+|\partial_tu_0|)\|_{L^\infty(B(x_0,3r)\times(t_0-5r^2,t_0+r^2))}\\ &\quad+C\varepsilon e^{3r\|\nabla\psi\|_\infty}\|e^{\psi}\nabla u_0\|_{L^\infty(B(x_0,3r)\times(t_0-5r^2,t_0+r^2))}.\end{aligned} \tag{5-10}$$

Since $\operatorname{supp}(f)\subset Q_r(y_0,s_0)$, it follows from the estimate (1-7) for $\Gamma_0(x,t;y,s)$ that

$$|\nabla^2u_0(x,t)|+|\partial_tu_0(x,t)|+r^{-1}|\nabla u_0(x,t)|\le C\exp\biggl\{-\frac{\kappa|x_0-y_0|^2}{t_0-s_0}\biggr\}\fint_{Q_r(y_0,s_0)}|fe^{-\psi}|\,dy\,ds \tag{5-11}$$

for any $x \in B(x_0, 3r)$ and $|t-t_0| \le 5r^2$, where $\kappa > 0$ depends only on μ. Thus, by (5-10), we obtain

$$\begin{aligned} &\|e^{\psi}(u_\varepsilon - u_0)\|_{L^\infty(Q_r(x_0,t_0))} \\ &\quad \le Ce^{3r\|\nabla\psi\|_\infty}\Big(\fint_{Q_{2r}(x_0,t_0)} |e^{\psi} w_\varepsilon|^2\Big)^{1/2} + \varepsilon r e^{c(|x_0-y_0|+r)\|\nabla\psi\|_\infty} \exp\Big\{-\frac{\kappa|x_0-y_0|^2}{t_0-s_0}\Big\}\fint_{Q_r(y_0,s_0)} |f|\,dy\,ds \\ &\quad \le C\varepsilon r e^{cr\|\nabla\psi\|_\infty}\Big\{e^{cr^2\|\nabla\psi\|_\infty^2} + e^{c|x_0-y_0|\|\nabla\psi\|_\infty}\exp\Big\{-\frac{\kappa|x_0-y_0|^2}{t_0-s_0}\Big\}\Big\}\cdot\Big(\fint_{Q_r(y_0,s_0)}|f|^2\Big)^{1/2}, \end{aligned} \tag{5-12}$$

where we have used (5-9) for the last step. By duality this implies

$$\begin{aligned} &\Big(\fint_{Q_r(y_0,s_0)} |e^{\psi(x)-\psi(y)}(\Gamma_\varepsilon(x,t;y,s)-\Gamma_0(x,t;y,s))|^2\,dy\,ds\Big)^{1/2} \\ &\qquad\qquad \le C\varepsilon r^{-d-1}e^{cr\|\nabla\psi\|_\infty}\Big\{e^{cr^2\|\nabla\psi\|_\infty^2} + e^{c|x_0-y_0|\|\nabla\psi\|_\infty}\exp\Big\{-\frac{\kappa|x_0-y_0|^2}{t_0-s_0}\Big\}\Big\} \end{aligned} \tag{5-13}$$

for any $(x,t) \in Q_r(x_0,t_0)$.

To deduce the L^∞ bound for

$$e^{\psi(x)-\psi(y)}(\Gamma_\varepsilon(x,t;y,s)-\Gamma_0(x,t;y,s))$$

from its L^2 bound in (5-13), we apply Theorem 5.3 (with ψ replaced by $-\psi$ and A replaced by $\tilde{A} = \tilde{A}(y,s) = A^*(y,-s)$) to the functions

$$v_\varepsilon(y,s) = \Gamma_\varepsilon(x_0,t_0;y,-s) \quad \text{and} \quad v_0(y,s) = \Gamma_0(x_0,t_0;y,-s).$$

Note that $(\partial_t + \widetilde{\mathcal{L}}_\varepsilon)v_\varepsilon = (\partial_t + \widetilde{\mathcal{L}}_0)v_0 = 0$ in $B(y_0,3r)\times(-s_0-5r^2, -s_0+r^2)$. Since $\tilde{A}$ satisfies the same conditions as A, we obtain

$$\begin{aligned} &|e^{\psi(x_0)-\psi(y_0)}(v_\varepsilon(y_0,-s_0)-v_0(y_0,-s_0))| \\ &\quad \le Ce^{3r\|\nabla\psi\|_\infty}\Big(\fint_{Q_r(y_0,-s_0)} |e^{\psi(x_0)-\psi(y)}(v_\varepsilon-v_0)|^2\,dy\,ds\Big)^{1/2} \\ &\qquad\qquad + C\varepsilon r^{-d-1}e^{cr\|\nabla\psi\|_\infty}e^{\psi(x_0)-\psi(y_0)}\exp\Big\{-\frac{\kappa|x_0-y_0|^2}{t_0-s_0}\Big\} \\ &\quad = Ce^{3r\|\nabla\psi\|_\infty}\Big(\fint_{Q_r(y_0,s_0+r^2)} |e^{\psi(x_0)-\psi(y)}(\Gamma_\varepsilon(x_0,t_0;y,s)-\Gamma_0(x_0,t_0;y,s))|^2\,dy\,ds\Big)^{1/2} \\ &\qquad\qquad + C\varepsilon r^{-d-1}e^{cr\|\nabla\psi\|_\infty}e^{\psi(x_0)-\psi(y_0)}\exp\Big\{-\frac{\kappa|x_0-y_0|^2}{t_0-s_0}\Big\} \\ &\quad \le C\varepsilon r^{-d-1}e^{cr\|\nabla\psi\|_\infty}\Big\{e^{cr^2\|\nabla\psi\|_\infty^2} + e^{c|x_0-y_0|\|\nabla\psi\|_\infty}\exp\Big\{-\frac{\kappa|x_0-y_0|^2}{t_0-s_0}\Big\}\Big\}, \end{aligned} \tag{5-14}$$

where we have used (5-13) for the last inequality.

Finally, as in [Hofmann and Kim 2004; Cho et al. 2008], we let $\psi(y) = \gamma\psi_0(|y-y_0|)$, where $\gamma \ge 0$ is to be chosen, $\psi_0(\rho) = \rho$ if $\rho \le |x_0-y_0|$, and $\psi_0(\rho) = |x_0-y_0|$ if $\rho > |x_0-y_0|$. Note that $\|\nabla\psi\|_\infty = \gamma$

and $\psi(x_0)-\psi(y_0)=\gamma|x_0-y_0|$. It follows from (5-14) that

$$|\Gamma_\varepsilon(x_0,t_0;y_0,s_0)-\Gamma_0(x_0,t_0;y_0,s_0)| \le C\varepsilon r^{-d-1}e^{-\gamma|x_0-y_0|+c\gamma\sqrt{t_0-s_0}}\left\{e^{c\gamma^2(t_0-s_0)}+e^{c\gamma|x_0-y_0|}\exp\left\{-\frac{\kappa|x_0-y_0|^2}{t_0-s_0}\right\}\right\}, \tag{5-15}$$

where $c>0$ depends at most on μ. If $|x_0-y_0|\le 2c\sqrt{t_0-s_0}$, we may simply choose $\gamma=0$. This gives

$$|\Gamma_\varepsilon(x_0,t_0;y_0,s_0)-\Gamma_0(x_0,t_0;y_0,s_0)|\le C\varepsilon r^{-d-1}\le C\varepsilon(t_0-s_0)^{-(d+1)/2}\exp\left\{-\frac{\kappa|x_0-y_0|^2}{t_0-s_0}\right\}.$$

If $|x_0-y_0|>2c\sqrt{t_0-s_0}$, we choose

$$\gamma=\frac{\delta|x_0-y_0|}{t_0-s_0}.$$

Note that

$$\begin{aligned}-\gamma|x_0-y_0|+c\gamma\sqrt{t_0-s_0}+c\gamma^2(t_0-s_0)&=-\delta(1-c\delta)\frac{|x_0-y_0|^2}{t_0-s_0}+c\delta\frac{|x_0-y_0|}{\sqrt{t_0-s_0}}\\&\le\left\{-\delta(1-c\delta)+\tfrac{1}{2}\delta\right\}\frac{|x_0-y_0|^2}{t_0-s_0}\le\frac{-\delta|x_0-y_0|^2}{4(t_0-s_0)}\end{aligned}$$

if $\delta\le\frac{1}{4}c^{-1}$. Also, observe that

$$c\gamma\sqrt{t_0-s_0}+c\gamma|x_0-y_0|-\frac{\kappa|x_0-y_0|^2}{t_0-s_0}\le\left\{\tfrac{1}{2}\delta+c\delta-\kappa\right\}\frac{|x_0-y_0|^2}{t_0-s_0}\le-\frac{\kappa|x_0-y_0|^2}{2(t_0-s_0)},$$

if $\delta\le\frac{1}{2}\left(c+\frac{1}{2}\right)^{-1}\kappa$. Recall that $r=(100)^{-1}\sqrt{t_0-s_0}$. As a result, we have proved that there exists $\kappa_1>0$, depending only on μ, such that

$$|\Gamma_\varepsilon(x_0,t_0;y_0,s_0)-\Gamma_0(x_0,t_0;y_0,s_0)|\le\frac{C\varepsilon}{(t_0-s_0)^{(d+1)/2}}\exp\left\{-\frac{\kappa_1|x_0-y_0|^2}{t_0-s_0}\right\}.$$

This completes the proof of Theorem 1.1. □

6. Proof of Theorem 1.2

Define

$$\|F\|_{C^{\lambda,0}(K)}=\sup\left\{\frac{|F(x,t)-F(y,t)|}{|x-y|^\lambda}:(x,t),(y,t)\in K\text{ and }x\ne y\right\}.$$

The proof of Theorem 1.2 relies on the following Lipschitz estimate.

Theorem 6.1. *Assume that A satisfies conditions (1-2), (1-3) and (1-9). Suppose that*

$$(\partial_t+\mathcal{L}_\varepsilon)u_\varepsilon=(\partial_t+\mathcal{L}_0)u_0$$

in $Q_{2r}(x_0,t_0)$ *for some* $(x_0,t_0)\in\mathbb{R}^{d+1}$ *and* $\varepsilon\le r<\infty$. *Then*

$$\|\nabla u_\varepsilon-\nabla u_0-(\nabla\chi)^\varepsilon\nabla u_0\|_{L^\infty(Q_r(x_0,t_0))}\le Cr^{-1}\biggl(\fint_{Q_{2r}(x_0,t_0)}|u_\varepsilon-u_0|^2\biggr)^{1/2}+C\varepsilon r^{-1}\|\nabla u_0\|_{L^\infty(Q_{2r}(x_0,t_0))}$$
$$+C\varepsilon\ln[\varepsilon^{-1}r+2]\||\nabla^2u_0|+\varepsilon|\partial_t\nabla u_0|+\varepsilon|\nabla^3u_0|\|_{L^\infty(Q_{2r}(x_0,t_0))}$$
$$+C\varepsilon^{1+\lambda}\||\nabla^2u_0|+\varepsilon|\partial_t\nabla u_0|+\varepsilon|\nabla^3u_0|\|_{C^{\lambda,0}(Q_{2r}(x_0,t_0))},\tag{6-1}$$

where C *depends only on* d, m, μ *and* (λ,τ) *in* (1-9).

Proof. Let

$$w_\varepsilon=u_\varepsilon-u_0-\varepsilon\chi_j^\varepsilon\frac{\partial u_0}{\partial x_j}-\varepsilon^2\phi^\varepsilon_{(d+1)ij}\frac{\partial^2u_0}{\partial x_i\partial x_j},\tag{6-2}$$

where $\chi_j^\varepsilon(x,t)=\chi_j(x/\varepsilon,t/\varepsilon^2)$ and $\phi^\varepsilon_{(d+1)ij}(x,t)=\phi_{(d+1)ij}(x/\varepsilon,t/\varepsilon^2)$. It follows by Proposition 3.1 that $(\partial_t+\mathcal{L}_\varepsilon)w_\varepsilon=\varepsilon\,\mathrm{div}(F_\varepsilon)$ in $Q_{2r}(x_0,t_0)$, where F_ε is given by (3-4) with S_ε being the identity operator. Choose a cut-off function $\varphi\in C_0^\infty(\mathbb{R}^{d+1})$ such that

$$0\le\varphi\le1,\qquad \varphi=1\quad\text{in } Q_{3r/2}(x_0,t_0),$$
$$\varphi(x,t)=0\quad\text{if } |x-x_0|\ge\tfrac74r \text{ or } t<t_0-\bigl(\tfrac74r\bigr)^2,$$
$$|\nabla\varphi|\le Cr^{-1},\qquad |\nabla^2\varphi|+|\partial_t\varphi|\le Cr^{-2}.$$

Using

$$(\partial_t+\mathcal{L}_\varepsilon)(\varphi w_\varepsilon)=(\partial_t\varphi)w_\varepsilon+\varepsilon\,\mathrm{div}(\varphi F_\varepsilon)-\varepsilon F_\varepsilon(\nabla\varphi)-\mathrm{div}(A^\varepsilon(\nabla\varphi)w_\varepsilon)-A^\varepsilon\nabla w_\varepsilon(\nabla\varphi),$$

where $A^\varepsilon(x,t)=A(x/\varepsilon,t/\varepsilon^2)$, we may deduce that for any $(x,t)\in Q_r(x_0,t_0)$,

$$\begin{aligned}w_\varepsilon(x,t)&=\int_{-\infty}^t\int_{\mathbb{R}^d}\Gamma_\varepsilon(x,t;y,s)\{(\partial_s\varphi)w_\varepsilon-\varepsilon F_\varepsilon(\nabla\varphi)-A^\varepsilon\nabla w_\varepsilon(\nabla\varphi)\}\,dy\,ds\\&\qquad-\int_{-\infty}^t\int_{\mathbb{R}^d}\nabla_y\Gamma_\varepsilon(x,t;y,s)\{\varepsilon\varphi F_\varepsilon-A^\varepsilon(\nabla\varphi)w_\varepsilon\}\,dy\,ds\\&=I(x,t)+J(x,t),\end{aligned}$$

where

$$J(x,t)=-\varepsilon\int_{-\infty}^t\int_{\mathbb{R}^d}\nabla_y\Gamma_\varepsilon(x,t;y,s)\varphi(y,s)F_\varepsilon(y,s)\,dy\,ds.$$

Since $\varphi=1$ in $Q_{3r/2}(x_0,t_0)$, we see that for $(x,t)\in Q_r(x_0,t_0)$,

$$\begin{aligned}|\nabla I(x,t)|&\le C\int_{-\infty}^t\int_{\mathbb{R}^d}|\nabla_x\Gamma_\varepsilon(x,t;y,s)|\{|\partial_s\varphi||w_\varepsilon|+\varepsilon|F_\varepsilon||\nabla\varphi|+|\nabla w_\varepsilon||\nabla\varphi|\}\,dy\,ds\\&\qquad+C\int_{-\infty}^t\int_{\mathbb{R}^d}|\nabla_x\nabla_y\Gamma_\varepsilon(x,t;y,s)||\nabla\varphi||w_\varepsilon|\,dy\,ds\\&\le C\biggl\{\frac1r\fint_{Q_{2r}(x_0,t_0)}|w_\varepsilon|+\varepsilon\fint_{Q_{2r}(x_0,t_0)}|F_\varepsilon|+\fint_{Q_{7r/4}(x_0,t_0)}|\nabla w_\varepsilon|\biggr\}\\&\le C\biggl\{\frac1r\biggl(\fint_{Q_{2r}(x_0,t_0)}|w_\varepsilon|^2\biggr)^{1/2}+\varepsilon\biggl(\fint_{Q_{2r}(x_0,t_0)}|F_\varepsilon|^2\biggr)^{1/2}\biggr\},\end{aligned}$$

where we have used (parabolic) Caccioppoli's inequality for the last step. In view of (3-4) with S_ε being the identity operator,

$$|F_\varepsilon| \le C\{|\nabla^2 u_0| + \varepsilon|\partial_t \nabla u_0| + \varepsilon|\nabla^3 u_0|\},$$

where we have used the boundedness of $\nabla\phi$ (see Remark 2.4). It follows that $\|\nabla I\|_{L^\infty(Q_r(x_0,t_0))}$ is bounded by the right-hand side of (6-1).

Finally, to estimate $J(x,t)$, we write

$$J(x,t) = -\varepsilon \int_{-\infty}^{t}\int_{\mathbb{R}^d} \nabla_y\{\Gamma_\varepsilon(x,t;y,s)\varphi(y,s)\}(F_\varepsilon(y,s)-F_\varepsilon(x,s))\,dy\,ds + \varepsilon \int_{-\infty}^{t}\int_{\mathbb{R}^d} \Gamma_\varepsilon(x,t;y,s)(\nabla\varphi)(y,s)F_\varepsilon(y,s)\,dy\,ds.$$

It follows that for $(x,t) \in Q_r(x_0,t_0)$

$$\begin{aligned} |\nabla J(x,t)| &\le \varepsilon \int_{Q_{2r}(x_0,t_0)} |\nabla_x\nabla_y\{\Gamma_\varepsilon(x,t;y,s)\varphi(y,s)\}|\,|F_\varepsilon(y,s)-F_\varepsilon(x,s)|\,dy\,ds \\ &\qquad + \varepsilon \int_{Q_{2r}(x_0,t_0)} |\nabla_x\Gamma_\varepsilon(x,t;y,s)|\,|\nabla\varphi(y,s)|\,|F_\varepsilon(y,s)|\,dy\,ds \\ &\le C\varepsilon \int_{Q_{2r}(x_0,t_0)} \frac{|F_\varepsilon(y,s)-F_\varepsilon(x,s)|}{(|x-y|+|t-s|^{1/2})^{d+2}}\,dy\,ds + C\varepsilon \fint_{Q_{2r}(x_0,t_0)} |F_\varepsilon|. \end{aligned} \tag{6-3}$$

To bound the first integral in the right-hand side of (6-3), we subdivide the domain $Q_{2r}(x_0,t_0)$ into $Q_\varepsilon(x,t)$ and $Q_{2r}(x_0,t_0)\setminus Q_\varepsilon(x,t)$. On $Q_{2r}(x_0,t_0)\setminus Q_\varepsilon(x,t)$, we use the bound

$$|F_\varepsilon(y,s)-F_\varepsilon(x,s)| \le 2\|F_\varepsilon\|_{L^\infty(Q_{2r}(x_0,t_0))},$$

while for $Q_\varepsilon(x,t)$, we use

$$|F_\varepsilon(y,s)-F_\varepsilon(x,s)| \le |x-y|^\lambda \|F\|_{C^{\lambda,0}(Q_{2r}(x_0,t_0))}.$$

This leads to

$$\begin{aligned} |\nabla J(x,t)| &\le C\varepsilon \ln[\varepsilon^{-1}r+1]\|F_\varepsilon\|_{L^\infty(Q_{2r}(x_0,t_0))} + C\varepsilon^{1+\lambda}\|F_\varepsilon\|_{C^{\lambda,0}(Q_{2r}(x_0,t_0))} \\ &\le C\varepsilon \ln[\varepsilon^{-1}r+1]\||\nabla^2 u_0| + \varepsilon|\partial_t\nabla u_0| + \varepsilon|\nabla^3 u_0|\|_{L^\infty(Q_{2r}(x_0,t_0))} \\ &\qquad + C\varepsilon^{1+\lambda}\||\nabla^2 u_0| + \varepsilon|\partial_t\nabla u_0| + \varepsilon|\nabla^3 u_0|\|_{C^{\lambda,0}(Q_{2r}(x_0,t_0))}. \end{aligned}$$

Thus, in view of the estimate for $\nabla I(x,t)$, we have proved that $\|\nabla w_\varepsilon\|_{L^\infty(Q_r(x_0,t_0))}$ is bounded by the right-hand side of (6-1). Since

$$\|\nabla w_\varepsilon - \{\nabla u_\varepsilon - \nabla u_0 - (\nabla\chi)^\varepsilon \nabla u_0\}\|_{L^\infty(Q_r(x_0,t_0))} \le C\varepsilon\||\nabla^2 u_0| + \varepsilon|\nabla^3 u_0|\|_{L^\infty(Q_r(x_0,t_0))},$$

the estimate (6-1) follows. □

To prove Theorem 1.2, we fix $x_0, y_0 \in \mathbb{R}^d$ and $s_0 < t_0$. We may assume that $\varepsilon < (t_0 - s_0)/8$. For otherwise the estimate (1-10) follows directly from (2-22). We apply Theorem 6.1 to the functions

$$u_\varepsilon(x,t) = \Gamma_\varepsilon(x,t;y_0,s_0) \quad\text{and}\quad u_0(x,t) = \Gamma_0(x,t;y_0,s_0)$$

in $Q_{2r}(x_0, t_0)$, where $r = (t_0 - s_0)/8$. Note that $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = (\partial_t + \mathcal{L}_0)u_0 = 0$ in $Q_{4r}(x_0, t_0)$. To bound the first term in the right-hand side of (6-1), we use the estimate (1-8) in Theorem 1.1. All other terms in the right-hand side of (6-1) may be handled easily by using the estimates (1-7) for $\Gamma_0(x, t; y, s)$. We leave the details to the reader.

7. Proof of Theorem 1.3

To prove Theorem 1.3, we fix $x_0, y_0 \in \mathbb{R}^d$ and $s_0 < t_0$. As before, we may assume that $\varepsilon < (t_0 - s_0)/8$, for otherwise the estimate (1-15) follows directly from (2-23).

Let $r = (t_0 - s_0)/8$. Fix $1 \le j \le d$ and $1 \le \beta \le m$. We apply Theorem 6.1 to the functions $u_\varepsilon = (u_\varepsilon^\alpha)$ and $u_0 = (u_0^\alpha)$ in $Q_{2r}(x_0, t_0)$, where

$$u_\varepsilon^\alpha(x, t) = \frac{\partial}{\partial y_j}\{\Gamma_\varepsilon^{\alpha\beta}\}(x, t; y_0, s_0),$$

$$u_0^\alpha(x, t) = \frac{\partial}{\partial y_\ell}\{\Gamma_0^{\alpha\sigma}\}(x, t; y_0, s_0) \cdot \left\{\delta^{\beta\sigma}\delta_{j\ell} + \frac{\partial}{\partial y_j}(\tilde{\chi}_\ell^{\beta\sigma})(y_0/\varepsilon, -s_0/\varepsilon^2)\right\},$$

where $\tilde{\chi}$ denotes the correctors for $\partial_t + \widetilde{\mathcal{L}}_\varepsilon$. Observe that $(\partial_t + \mathcal{L}_\varepsilon)u_\varepsilon = (\partial_t + \mathcal{L}_0)u_0 = 0$ in $Q_{4r}(x_0, t_0)$. To bound the first term in the right-hand side of (6-1), we use the estimate (1-14). As in the proof of Theorem 1.1, all other terms in the right-hand side of (6-1) may be handled readily by using estimate (1-7) for $\Gamma_0(x, t; y, s)$.

References

[Aronson 1967] D. G. Aronson, "Bounds for the fundamental solution of a parabolic equation", *Bull. Amer. Math. Soc.* **73** (1967), 890–896. MR Zbl

[Auscher 1996] P. Auscher, "Regularity theorems and heat kernel for elliptic operators", *J. London Math. Soc.* (2) **54**:2 (1996), 284–296. MR Zbl

[Avellaneda and Lin 1991] M. Avellaneda and F.-H. Lin, "L^p bounds on singular integrals in homogenization", *Comm. Pure Appl. Math.* **44**:8-9 (1991), 897–910. MR Zbl

[Bensoussan et al. 1978] A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic analysis for periodic structures*, Studies in Math. Appl. **5**, North-Holland, Amsterdam, 1978. MR Zbl

[Byun 2007] S.-S. Byun, "Optimal $W^{1,p}$ regularity theory for parabolic equations in divergence form", *J. Evol. Equ.* **7**:3 (2007), 415–428. MR Zbl

[Cho et al. 2008] S. Cho, H. Dong, and S. Kim, "On the Green's matrices of strongly parabolic systems of second order", *Indiana Univ. Math. J.* **57**:4 (2008), 1633–1677. MR Zbl

[Davies 1987a] E. B. Davies, "Explicit constants for Gaussian upper bounds on heat kernels", *Amer. J. Math.* **109**:2 (1987), 319–333. MR Zbl

[Davies 1987b] E. B. Davies, "Heat kernel bounds for second order elliptic operators on Riemannian manifolds", *Amer. J. Math.* **109**:3 (1987), 545–569. MR Zbl

[Fabes and Stroock 1986] E. B. Fabes and D. W. Stroock, "A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash", *Arch. Rational Mech. Anal.* **96**:4 (1986), 327–338. MR Zbl

[Geng and Shen 2015] J. Geng and Z. Shen, "Uniform regularity estimates in parabolic homogenization", *Indiana Univ. Math. J.* **64**:3 (2015), 697–733. MR Zbl

[Geng and Shen 2017] J. Geng and Z. Shen, "Convergence rates in parabolic homogenization with time-dependent periodic coefficients", *J. Funct. Anal.* **272**:5 (2017), 2092–2113. MR Zbl

[Hofmann and Kim 2004] S. Hofmann and S. Kim, "Gaussian estimates for fundamental solutions to certain parabolic systems", *Publ. Mat.* **48**:2 (2004), 481–496. MR Zbl

[Jikov et al. 1994] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik, *Homogenization of differential operators and integral functionals*, Springer, 1994. MR Zbl

[Kenig and Shen 2011] C. E. Kenig and Z. Shen, "Layer potential methods for elliptic homogenization problems", *Comm. Pure Appl. Math.* **64**:1 (2011), 1–44. MR Zbl

[Kenig et al. 2014] C. E. Kenig, F. Lin, and Z. Shen, "Periodic homogenization of Green and Neumann functions", *Comm. Pure Appl. Math.* **67**:8 (2014), 1219–1262. MR Zbl

[Krylov 2007] N. V. Krylov, "Parabolic and elliptic equations with VMO coefficients", *Comm. Partial Differential Equations* **32**:1-3 (2007), 453–475. MR Zbl

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

Received 1 Jan 2018. Revised 28 Oct 2018. Accepted 29 Dec 2018.

JUN GENG: gengjun@lzu.edu.cn
School of Mathematics and Statistics, Lanzhou University, Lanzhou, China

ZHONGWEI SHEN: zshen2@uky.edu
Department of Mathematics, University of Kentucky, Lexington, KY, United States

ANALYSIS AND PDE
Vol. 13, No. 1, 2020
dx.doi.org/10.2140/apde.2020.13.171

CAPILLARY SURFACES ARISING IN SINGULAR PERTURBATION PROBLEMS

ARAM L. KARAKHANYAN

We prove some Bernstein-type theorems for a class of stationary points of the Alt–Caffarelli functional in $\mathbb{R}^2$ and $\mathbb{R}^3$ arising as limits of the singular perturbation problem

$$\begin{cases} \triangle u_\varepsilon(x) = \beta_\varepsilon(u_\varepsilon) & \text{in } B_1, \\ |u_\varepsilon| \le 1 & \text{in } B_1, \end{cases}$$

in the unit ball B_1 as $\varepsilon \to 0$. Here $\beta_\varepsilon(t) = (1/\varepsilon)\beta(t/\varepsilon) \ge 0$, $\beta \in C_0^\infty[0,1]$, $\int_0^1 \beta(t)\,dt = M > 0$, is an approximation of the Dirac measure and $\varepsilon > 0$. The limit functions $u = \lim_{\varepsilon_j \to 0} u_{\varepsilon_j}$ of uniformly converging sequences $\{u_{\varepsilon_j}\}$ solve a Bernoulli-type free boundary problem in some weak sense. Our approach has two novelties: First we develop a hybrid method for stratification of the free boundary $\partial\{u_0 > 0\}$ of blow-up solutions which combines some ideas and techniques of viscosity and variational theory. An important tool we use is a new monotonicity formula for the solutions u_ε based on a computation of J. Spruck. It implies that any blow-up u_0 of u either vanishes identically or is a homogeneous function of degree 1, that is, $u_0 = rg(\sigma)$, $\sigma \in \mathbb{S}^{N-1}$, in spherical coordinates (r, θ). In particular, this implies that in two dimensions the singular set is empty at the nondegenerate points, and in three dimensions the singular set of u_0 is at most a singleton. Second, we show that the spherical part g is the support function (in Minkowski's sense) of some capillary surface contained in the sphere of radius $\sqrt{2M}$. In particular, we show that $\nabla u_0 : \mathbb{S}^2 \to \mathbb{R}^3$ is an almost conformal and minimal immersion and the singular Alt–Caffarelli example corresponds to a piece of catenoid which is a unique ring-type stationary minimal surface determined by the support function g.

1. Introduction

In this paper we study the singular perturbation problem

$$\begin{cases} \triangle u_\varepsilon(x) = \beta_\varepsilon(u_\varepsilon) & \text{in } B_1, \\ |u_\varepsilon| \le 1 & \text{in } B_1, \end{cases} \tag{$\mathcal{P}_\varepsilon$}$$

where $\varepsilon > 0$ is a small parameter,

$$\begin{cases} \beta_\varepsilon(t) = (1/\varepsilon)\beta(t/\varepsilon), \\ \beta(t) \ge 0, \quad \operatorname{supp}\beta \subset [0,1], \\ \int_0^1 \beta(t)\,dt = M > 0 \end{cases} \tag{1-1}$$

is an approximation of the Dirac measure, and $B_1 \subset \mathbb{R}^N$ is the unit ball centered at the origin. It is well known that ($\mathcal{P}_\varepsilon$) models propagation of equidiffusional premixed flames with high activation of energy

MSC2010: 49Q05, 35R35, 35B25.
Keywords: singular perturbation problem, free boundary regularity, capillary surfaces, global solutions.

[Caffarelli 1995]. Heuristically, the limit $u_0 = \lim_{\varepsilon_j \to 0} u_{\varepsilon_j}$ (for a suitable sequence $\varepsilon_j \to 0$) solves a Bernoulli-type free boundary problem with the free boundary condition

$$|\nabla u^+|^2 - |\nabla u^-|^2 = 2M.$$

If the functions $\{u_\varepsilon\}$ are also minimizers of

$$J_\varepsilon[u_\varepsilon] = \int_\Omega \frac{|\nabla u_\varepsilon|^2}{2} + \mathcal{B}\left(\frac{u_\varepsilon}{\varepsilon}\right), \quad \mathcal{B}(t) = \int_0^t \beta(s)\, ds, \tag{1-2}$$

then the limits of $\{u_\varepsilon\}$ inherit the generic features of minimizers (e.g., nondegeneracy, rectifiability of $\partial\{u > 0\}$, etc.). Consequently, the limits of uniformly converging sequences $\{u_{\varepsilon_j}\}$ as $\varepsilon_j \to 0$ are minimizers of the Alt–Caffarelli functional $J[u] = \int_{B_1} |\nabla u|^2 + 2M\chi_{\{u>0\}}$. It is known that the singular set of minimizers is empty in dimensions $2, 3$ and 4; see [Alt and Caffarelli 1981; Caffarelli et al. 2004; Jerison and Savin 2015]. However, if u_ε is not a minimizer then the analysis of the limits of u presents a more delicate problem. The main difficulty in carrying out such analysis is that the free boundary may contain degenerate points [Weiss 2003].

This paper is devoted to the study of the blow-ups of the limits of the singular perturbation problem $(\mathcal{P}_\varepsilon)$ and establishes a new and direct connection with *minimal surfaces*. In particular, we show that every blow-up of a limit function $u = \lim_{\varepsilon_j \to 0} u_{\varepsilon_j}$ in $\mathbb{R}^3$ (for an appropriate sequence ε_j) defines an almost conformal and minimal immersion which is perpendicular to the sphere of radius $\sqrt{2M}$, where $M = \int_0^1 \beta(t)\, dt$. In other words, one obtains a capillary surface inside the sphere of radius $\sqrt{2M}$.

Our first result is:

Theorem A. *Let $u_{\varepsilon_j} \to u$ locally uniformly in B_1 for some subsequence ε_j. Then any blow-up of u at a free boundary point $x_0 \in \partial\{u > 0\}$ is either identically zero or a homogeneous function of degree* 1. *In particular, if $N = 2$ and u is not degenerate at $x_0 \in \partial\{u > 0\}$ then every blow-up of u at x_0 must be one of the following functions* (*after some rotation of coordinates*):

(1) $\sqrt{2M}x_1^+$, *a half-plane solution provided that there is a measure-theoretic normal at x_0*,

(2) *a wedge* $\alpha|x_1|$, $0 < \alpha \le \sqrt{2M}$,

(3) *a two-plane solution* $\alpha x_1^+ - \beta x_1^-, \alpha^2 - \beta^2 = 2M$, $\alpha, \beta > 0$.

In order to prove Theorem A we will introduce a monotone quantity based on a computation of Joel Spruck [1983]. From Theorem A it follows that in $\mathbb{R}^2$ the blow-up limits at nondegenerate free boundary points can be explicitly computed. It is worthwhile to note that the minimizers of

$$J[u] = \int_{B_1} |\nabla u|^2 + 2M\chi_{\{u>0\}} \tag{1-3}$$

are nondegenerate; i.e., for each subdomain $\Omega' \Subset B_1$ there is a constant $c_0 > 0$ depending on $\mathrm{dist}(\partial B_1, \partial\Omega')$, N, M, such that

$$\sup_{B_r(x_0)} u^+ \ge c_0 r \quad \text{for all } x_0 \in \partial\{u > 0\} \cap \Omega', \ \ B_r(x_0) \subset B_1. \tag{1-4}$$

However, if u_ε is *any* solution of $(\mathcal{P}_\varepsilon)$ then nondegeneracy may not be true. There is a sufficient condition [Caffarelli et al. 1997, Theorem 6.3] that implies (1-4).

Some well-known examples demonstrate rather strikingly that for the stationary case there are wedge-like global solutions for which the measure-theoretic boundary of $\{u > 0\}$ is empty. This is impossible for minimizers. In fact, the zero set of a minimizer has uniformly positive Lebesgue density. In this respect Theorem A only states that if u is nondegenerate at x_0 then the blow-up is a nontrivial cone.

The existence of wedge solutions, see [Caffarelli et al. 1997, Remark 5.1], suggests that some further assumptions are needed to formulate the free boundary condition. For instance, one may assume that the upper Lebesgue density at $x \in \partial\{u > 0\}$ satisfies $\Theta^*(x, \{u > 0\}) < 1$; i.e., the upper density measure is not covering the full ball. We emphasize that for some solutions the topological and measure-theoretic boundaries may not coincide. Our next result addresses the degeneracy and wedge-formation in $\mathbb{R}^3$ of blow-ups at free boundary points.

Theorem B. *Suppose $N = 3$. Let $u \geq 0$ be a limit of some uniformly converging sequence $\{u_{\varepsilon_j}\}$ solving $(\mathcal{P}_\varepsilon)$ such that u is nondegenerate at $y_0 \in \partial\{u > 0\}$. Let u_0 be a blow-up of u at y_0. If $\mathfrak{C}$ is a component of $\partial\{u_0 > 0\}$ such that the measure-theoretic boundary of $\{u_0 > 0\}$ in $\mathfrak{C}$ is nonempty then*

(1) *all points of $\mathfrak{C}$ are nondegenerate,*

(2) *$\mathfrak{C}$ is a subset of the measure-theoretic boundary of $\{u_0 > 0\}$,*

(3) *$\mathfrak{C} \setminus \{0\}$ is smooth.*

In particular in $\mathbb{R}^3$ the singular set of $\partial\{u_0 > 0\}$ is at most a singleton.

Theorem B implies that the reduced boundary propagates instantaneously in the components of $\partial\{u_0 > 0\}$. Our last result sheds some new light on the characterization of the blow-ups as minimal surfaces inside spheres with contact angle $\frac{\pi}{2}$.

Theorem C. *Let u_0 be as in Theorem B and $u_0 = rg(\sigma)$, $\sigma \in \mathbb{S}^2$, in spherical coordinates. Then the parametrization $\mathcal{X}(\sigma) = \sigma g(\sigma) + \nabla_{\mathbb{S}^2} g(\sigma)$ defines an almost conformal and minimal immersion. If $\{g > 0\}$ is homeomorphic to a disk then u_0 is a half-plane solution $\sqrt{2M}x_1^+$. If $\{g > 0\}$ is homeomorphic to a ring then the only singular cone is the Alt–Caffarelli catenoid.*

Observe that $\triangle u_0 = 0$ implies that the spherical part g satisfies the following equation on the sphere:

$$\triangle_{\mathbb{S}^{N-1}} g + (N-1)g = 0,$$

where $\triangle_{\mathbb{S}^{N-1}}$ is the Laplace–Beltrami operator. If we regard g as the support function of some embedded hypersurface $\mathcal{M}$ then the matrix $[\nabla_{ij} g + \delta_{ij} g]^{-1}$ gives the Weingarten mapping and its eigenvalues are the principal curvatures $k_1, \dots, k_{N-1}$ of $\mathcal{M}$. If $N = 3$ then we have

$$0 = \triangle_{\mathbb{S}^2} g + 2g = \operatorname{trace}[\nabla_{ij} g + \delta_{ij} g] = \frac{1}{k_1} + \frac{1}{k_2} = \frac{k_1 + k_2}{k_1 k_2},$$

implying that the mean curvature is zero at the points where the Gauss curvature $k_1 k_2$ does not vanish. This is how the minimal surfaces enter into the game. One of the main obstacles is to show that the surface parametrized by $\mathcal{X}(\sigma) = \nabla u_0(\sigma)$ is embedded. Then the classification for the disk-type domains $\{g > 0\}$ follows from a result of [Nitsche 1985]. To prove the last statement of Theorem C we will use the moving plane method. It is worthwhile to point out that the results of this paper can be extended to other classes of stationary points. For instance, the weak solutions introduced in [Alt and Caffarelli 1981] can be analyzed in similar way provided that the zero set has uniformly positive Lebesgue density at free boundary points in order to guarantee that the class of weak solution is closed with respect to blow-ups; see [Alt and Caffarelli 1981, Example 5.8].

Related works. In [Hauswirth et al. 2011] L. Hauswirth, F. Hélein, and F. Pacard considered the overdetermined problem

$$\begin{cases} \triangle u(x) = 0 & \text{in } \Omega, \\ u > 0 & \text{in } \Omega, \\ u(x) = 0, \quad |\nabla u| = 1 & \text{on } \partial\Omega, \end{cases} \tag{1-5}$$

where Ω is a smooth domain and the boundary conditions are satisfied in the classical sense. A domain Ω admitting a solution u to (1-5) is called exceptional. Note that every nonnegative smooth solution of the limiting singular perturbation problem solves (1-5) with $M = \frac{1}{2}$. In [Hauswirth et al. 2011] the authors constructed a number of examples of exceptional domains and proposed to classify them. In particular, they proved that if $\Omega \subset \mathbb{R}^2$ is conformal to a half-plane such that u is strictly monotone in one fixed direction then Ω is a half-space [Hauswirth et al. 2011, Proposition 6.1]. However the general problem remained open.

Later M. Traizet [2014, Proposition 1] showed that the smoothness assumption can be relaxed, namely if $\Omega \subset \mathbb{R}^2$ has C^0 boundary and the boundary conditions are still satisfied in the classical sense then Ω is real-analytic. Under various topological conditions on the two-dimensional domain $\Omega \subset \mathbb{R}^2$ (such as finite connectivity and periodicity), M. Traizet classified the possible exceptional domains. One of his remarkable results is that from Ω one can construct a *complete* minimal surface using the Weierstrass representation formula [Traizet 2014, Theorem 9]. Another classification result in $\mathbb{R}^2$, under stronger topological hypotheses than in [Traizet 2014], was given by D. Khavinson, E. Lundberg and R. Teodorescu [Khavinson et al. 2013]. Moreover, their results in the simply connected case are stronger because unlike M. Traizet they do not assume the finite connectivity (i.e., $\partial\Omega$ has finite number of components). As opposed to these results (1) we do not assume any regularity of the free boundary (which plays the role of $\partial\Omega$ in (1-5)), (2) the Neumann condition is not satisfied in the classical sense, (3) the minimal surface we construct in Theorem C is *not complete* and it is a capillary surface inside sphere, and (4) our techniques do not impose any restriction on the dimension. Note that in [Hauswirth et al. 2011] the authors suggested to study more general classes of exceptional domains: if (M, g) is an m-dimensional Riemannian manifold admitting a harmonic function with zero Dirichlet and constant Neumann boundary data then M is called exceptional and u a roof function. In this context Theorem C provides a way of constructing a roof function on the sphere from the blow-ups of stationary points of the Alt–Caffarelli functional.

One may consider higher-order critical points as well, such as mountain passes (which are, in fact, minimizers over some subspace of admissible functions), for which one has nondegeneracy and nontrivial

Lebesgue density properties [Jerison and Perera 2018, Propositions 1.7–5.1]. Observe that neither of these properties is available for our solutions as Theorem 6.3 and Remark 5.1 in [Caffarelli et al. 1997] indicate, and in the present work we do not impose any additional assumptions on our stationary points of this kind.

It seems that the only result in high dimensions that appears in [Hauswirth et al. 2011; Khavinson et al. 2013; Traizet 2014] states that if the complement of Ω is connected and has $C^{2,\alpha}$ boundary, then Ω is the exterior of a ball [Khavinson et al. 2013, Theorem 7.1]. The restriction $\Omega \subset \mathbb{R}^2$ is because the authors have mainly used the techniques from complex analysis. Our approach does not have this restriction since our main tool is the representation of the solution in terms of the Minkowski support function. We remark that using our method in high dimensions we can construct a surface $\mathcal{M}$ inside the sphere of radius $\sqrt{2M}$ such that the sum of its principal radii of curvature is zero, and $\mathcal{M}$ is transversal to the sphere.

Finally, we point out that our approach may lead to a new characterization of global minimizers in $\mathbb{R}^3$ [Caffarelli et al. 2004]. Indeed, [Ros and Vergasta 1995, Theorem 6] implies that the capillary surface $\mathcal{M}$ in Theorem C associated with the blow-up limit must be totally geodesic (i.e., the second fundamental form is identically zero). Consequently, the blow-up must be the half-plane solution.

The paper is organized as follows: In Section 2 we set up some basic notation which will be used throughout the paper. Section 3 is devoted to the study of a new monotone quantity $s(x_0, u, r)$. This interesting quantity is derived from a computation of Spruck [1983]. Among other things, properties of s imply that every blow-up of u is either a homogeneous function of degree 1 or identically zero. Section 4 contains the proof of Theorem A. In Section 5 we develop a new method of stratification of the free boundary points and prove Theorem B. Section 6 contains the proof of Theorem C. For the convenience of the reader, in the Appendix we repeat the relevant material from [Caffarelli et al. 1997] without proofs.

2. Notation

Throughout the paper N will denote the spatial dimension. $B_r(x_0) = \{x \in \mathbb{R}^N : |x - x_0| < r\}$ denotes the open ball of radius $r > 0$ centered at $x_0 \in \mathbb{R}^N$. The s-dimensional Hausdorff measure is denoted by $\mathcal{H}^s$, the unit sphere by $\mathbb{S}^{N-1} \subset \mathbb{R}^N$, and the characteristic function of the set D by χ_D. We also let

$$M = \int_0^1 \beta(t)\, dt.$$

Sometimes we will set $x = (x_1, x')$, where $x' \in \mathbb{R}^{N-1}$. For a given function v, we will define $v^+ = \max(0, v)$ and $v^- = \max(0, -v)$. Finally, we say that $v \in C^{0,1}_{\text{loc}}(\mathcal{D})$ if for every $\mathcal{D}' \Subset \mathcal{D}$, there is a constant $L(\mathcal{D}')$ such that

$$|v(x) - v(y)| \le L(\mathcal{D}')|x - y| \quad \text{for all } x, y \in \mathcal{D}.$$

If $v \in C^{0,1}_{\text{loc}}(\mathcal{D})$ then we say that v is locally Lipschitz continuous in $\mathcal{D}$. For $x = (x_1, \dots, x_N)$ and fixed $x_0 \in \mathbb{R}^N$ we denote by $(x - x_0)_1^+$ the positive part of the first coordinate of $x - x_0$. If $u(x_0) = 0$ then $(u(x))_r = u(x_0 + rx)/r$, $r > 0$, denotes the scaled function at x_0. For given $r_j \to 0$ the sequence $(u(x))_{r_j}$ is called a blow-up sequence and its limit u_0 a blow-up of u at x_0.

3. Monotonicity formula of Spruck

It is convenient to work with a weaker definition of nondegeneracy which only ensures that the blow-up does not vanish identically.

Definition 3.1. We say that u is degenerate at $x_0 \in \partial\{u > 0\}$ if $\liminf_{r\to 0} (1/r) \fint_{B_r(x_0)} u^+ = 0$.

Observe that $u^+(x) = o(|x - x_0|)$ near the degenerate point x_0 because u^+ is subharmonic.

It is known that the solutions of $(\mathcal{P}_\varepsilon)$ are locally Lipschitz continuous; see the Appendix, Proposition A.1. Consequently, there is a subsequence $\varepsilon_j \to 0$ such that $u_{\varepsilon_j} \to u$ locally uniformly. Furthermore, u is a stationary point of the Alt–Caffarelli problem in some weak sense and the blow-up of u can be approximated by some scaled family of solutions to $(\mathcal{P}_\varepsilon)$; see the Appendix, Propositions A.5 and A.6.

Proposition 3.1. *Let u be a limit of some sequence u_{ε_j} as in Proposition A.2. Then any blow-up of u at a nondegenerate point is a homogeneous function of degree* 1.

Proof. To fix the ideas we assume that $0 \in \partial\{u > 0\}$ is a nondegenerate point. We begin with writing the Laplacian in polar coordinates

$$\Delta u = u_{rr} + \frac{N-1}{r} u_r + \frac{1}{r^2} \Delta_{\mathbb{S}^{N-1}} u \tag{3-1}$$

and then introducing the auxiliary function

$$v(t,\sigma) = \frac{u(r,\sigma)}{r}, \qquad r = e^{-t}. \tag{3-2}$$

A straightforward computation yields

$$v_t = -u_r + v, \qquad v_\sigma = \frac{u_\sigma}{r}, \qquad v_{tt} = u_{rr} r + v_t, \qquad \Delta_{\mathbb{S}^{N-1}} v = \frac{1}{r} \Delta_{\mathbb{S}^{N-1}} u,$$

where, with some abuse of notation, v_σ denotes the gradient of v computed on the sphere. Rewriting the equation $\Delta u_\varepsilon = \beta_\varepsilon(u_\varepsilon)$ in t- and σ-derivatives we obtain

$$\frac{1}{r}[(N-1)(v - \partial_t v_\varepsilon) + \partial_t^2 v_\varepsilon - \partial_t v_\varepsilon + \Delta_{\mathbb{S}^{N-1}} v_\varepsilon] = \frac{1}{\varepsilon}\beta\left(\frac{r}{\varepsilon} v_\varepsilon\right).$$

Next, we multiply both sides of the last equation by $\partial_t v_\varepsilon$ to get

$$\partial_t v_\varepsilon[(N-1)(v - \partial_t v_\varepsilon) + \partial_t^2 v_\varepsilon - \partial_t v_\varepsilon + \Delta_{\mathbb{S}^{N-1}} v_\varepsilon] = \partial_t v_\varepsilon \frac{r}{\varepsilon}\beta\left(\frac{r}{\varepsilon} v_\varepsilon\right). \tag{3-3}$$

The right-hand side of (3-3) can be further transformed as follows:

$$\begin{aligned}
\frac{r}{\varepsilon}\beta\left(\frac{e^{-t}}{\varepsilon} v_\varepsilon\right)\partial_t v_\varepsilon &= \beta\left(\frac{e^{-t}}{\varepsilon} v_\varepsilon\right)\left[\frac{e^{-t}}{\varepsilon}\partial_t v_\varepsilon - \frac{e^{-t}}{\varepsilon} v_\varepsilon\right] + \beta\left(\frac{e^{-t}}{\varepsilon} v_\varepsilon\right)\frac{e^{-t}}{\varepsilon} v_\varepsilon \\
&= \partial_t \int_0^{(e^{-t}/\varepsilon) v_\varepsilon} \beta(s)\,ds + \beta\left(\frac{e^{-t}}{\varepsilon} v_\varepsilon\right)\frac{e^{-t}}{\varepsilon} v_\varepsilon \\
&= \partial_t \mathcal{B}\left(\frac{e^{-t}}{\varepsilon} v_\varepsilon\right) + \beta\left(\frac{e^{-t}}{\varepsilon} v_\varepsilon\right)\frac{e^{-t}}{\varepsilon} v_\varepsilon \equiv I_1.
\end{aligned}$$

It is important to note that by our assumption (1-1) the last term is nonnegative; in other words

$$\beta\left(\frac{e^{-t}}{\varepsilon}v_\varepsilon\right)\frac{e^{-t}}{\varepsilon}v_\varepsilon \geq 0. \tag{3-4}$$

Moreover, we have

$$\begin{aligned} I_2 &\equiv [(N-1)v_\varepsilon - N\partial_t v_\varepsilon + \partial_t^2 v_\varepsilon + \triangle_{\mathbb{S}^{N-1}} v_\varepsilon]\partial_t v_\varepsilon \\ &= (N-1)\partial_t\left(\frac{v_\varepsilon^2}{2}\right) - N(\partial_t v_\varepsilon)^2 + \partial_t\left(\frac{(\partial_t v_\varepsilon)^2}{2}\right) + \partial_t v_\varepsilon \triangle_{\mathbb{S}^{N-1}} v_\varepsilon. \end{aligned}$$

Next we integrate the identity

$$I_2 = rI_1$$

over $\mathbb{S}^{N-1}$ and then over $[T_0, T]$ in order to get

$$\begin{aligned} (N-1)\int_{\mathbb{S}^{N-1}} \frac{v_\varepsilon^2}{2}\bigg|_{T_0}^{T} - N\int_{T_0}^{T}\int_{\mathbb{S}^{N-1}}(\partial_t v_\varepsilon)^2 + \int_{\mathbb{S}^{N-1}} \frac{(\partial_t v_\varepsilon)^2}{2}\bigg|_{T_0}^{T} + \int_{T_0}^{T}\int_{\mathbb{S}^{N-1}} \partial_t v_\varepsilon \triangle_{\mathbb{S}^{N-1}} v_\varepsilon \\ = \int_{\mathbb{S}^{N-1}} \mathcal{B}\left(\frac{e^{-t}}{\varepsilon}v_\varepsilon\right)\bigg|_{T_0}^{T} + \int_{T_0}^{T}\int_{\mathbb{S}^{N-1}} \beta\left(\frac{r}{\varepsilon}v_\varepsilon\right)\frac{r}{\varepsilon}v_\varepsilon. \end{aligned}$$

Note that

$$\int_{T_0}^{T}\int_{\mathbb{S}^{N-1}} \partial_t v_\varepsilon \triangle_{\mathbb{S}^{N-1}} v_\varepsilon = -\frac{1}{2}\int_{\mathbb{S}^{N-1}} |\nabla_\sigma v_\varepsilon|^2\Big|_{T_0}^{T}. \tag{3-5}$$

Rearranging the terms and utilizing (3-4) we get the identity

$$\begin{aligned} N\int_{T_0}^{T}\int_{\mathbb{S}^{N-1}}(\partial_t v_\varepsilon)^2 + \int_{T_0}^{T}\int_{\mathbb{S}^{N-1}} \beta\left(\frac{r}{\varepsilon}v_\varepsilon\right)\frac{r}{\varepsilon}v_\varepsilon \\ = (N-1)\int_{\mathbb{S}^{N-1}} \frac{v_\varepsilon^2}{2}\bigg|_{T_0}^{T} + \int_{\mathbb{S}^{N-1}} \frac{(\partial_t v_\varepsilon)^2}{2}\bigg|_{T_0}^{T} - \frac{1}{2}\int_{\mathbb{S}^{N-1}} |\nabla_\sigma v_\varepsilon|^2\Big|_{T_0}^{T} - \int_{\mathbb{S}^{N-1}} \mathcal{B}\left(\frac{e^{-t}}{\varepsilon}v_\varepsilon\right)\bigg|_{T_0}^{T}. \end{aligned} \tag{3-6}$$

From here it follows that

$$\int_{T_0}^{T}\int_{\mathbb{S}^{N-1}}(\partial_t v_\varepsilon)^2 \leq C, \tag{3-7}$$

where C depends on $\|\nabla u_\varepsilon\|_\infty, M, N$ but not on ε, T_0 or T.

Letting $\varepsilon \to 0$ we conclude

$$\int_{T_0}^{T}\int_{\mathbb{S}^{N-1}}(\partial_t v)^2 \leq C, \tag{3-8}$$

where $v(t,\sigma) = u(r,\sigma)/r$. But $\partial_t v = -u_r + u/r$, implying that

$$\int_{T_0}^{\infty}\int_{\mathbb{S}^{N-1}}\left(u_r - \frac{u}{r}\right)^2 dt\, d\sigma \leq C. \tag{3-9}$$

The proof of Theorem A follows if we note that $-u_r + u/r = 0$ is the Euler equation for the homogeneous functions of degree 1. □

In the proof of Proposition 3.1 we used Spruck's original computation [1983]. The identity (3-6) can be interpreted as a local energy balance for u_ε. Moreover, using (3-6) we can construct a monotone quantity which has some remarkable properties.

Corollary 3.2. *Suppose $0 \in \partial\{u > 0\}$ and let (r, σ), $\sigma \in \mathbb{S}^{N-1}$, be the spherical coordinates. Introduce*

$$S_\varepsilon(r) = \int_{\mathbb{S}^{N-1}} \left[2\mathcal{B}\left(\frac{u_\varepsilon(r,\sigma)}{\varepsilon}\right) + \frac{1}{r^2}|\nabla_\sigma u_\varepsilon|^2 - (N-1)\frac{u_\varepsilon^2(r,\sigma)}{r^2} - \left(\partial_r u_\varepsilon(r,\sigma) - \frac{u_\varepsilon(r,\sigma)}{r}\right)^2 \right] d\sigma. \tag{3-10}$$

- *Then $S_\varepsilon(r)$ is nondecreasing in r.*
- *Moreover, if $u_{\varepsilon_j} \to u$ for some subsequence $\varepsilon_j \to 0$, then $S_{\varepsilon_j}(r) \to S(r)$ for a.e. r, where*

$$S(r) = \int_{\mathbb{S}^{N-1}} \left[2M\chi_{\{u>0\}} + \frac{1}{r^2}|\nabla_\sigma u|^2 - (N-1)\frac{u^2(r,\sigma)}{r^2} - \left(\partial_r u(r,\sigma) - \frac{u(r,\sigma)}{r}\right)^2 \right] d\sigma. \tag{3-11}$$

  In particular, $S(r)$ is a nondecreasing function of r.
- *$S(r)$ is constant if and only if u is a homogeneous function of degree* 1.

Proof. By setting $r_1 = e^{-T}$, $r_2 = e^{-T_0}$ and noting that $r_1 < r_2$ if $T > T_0$ we obtain from (3-6)

$$S_\varepsilon(r_2) - S_\varepsilon(r_1) = 2N \int_{T_0}^{T} \int_{\mathbb{S}^{N-1}} (\partial_t v_\varepsilon)^2 + 2\int_{T_0}^{T} \int_{\mathbb{S}^{N-1}} \beta\left(\frac{r}{\varepsilon}v_\varepsilon\right)\frac{r}{\varepsilon}v_\varepsilon \geq 0,$$

where we applied (3-4) and hence the first claim follows. The second part follows from Propositions A.1 and A.2. Indeed, integrating $S_\varepsilon(r) \leq S_\varepsilon(r+t)$, $t \geq 0$, over $[r_1 - \delta, r_1 + \delta]$ we infer

$$\frac{1}{2\delta}\int_{r_1-\delta}^{r_1+\delta} S_\varepsilon(r)\,dr \leq \frac{1}{2\delta}\int_{r_1-\delta}^{r_1+\delta} S_\varepsilon(r+t)\,dr.$$

Then first letting $\varepsilon \to 0$ and utilizing Proposition A.2 together with (A-1) and then sending $\delta \to 0$ we infer that $S(r)$ is nondecreasing for a.e. r. Finally the last part follows as in the proof of Proposition 3.1. □

As one can see we did not use the Pohozhaev identity, as opposed to the monotonicity formula in [Weiss 2003]. Spruck's monotonicity formula enjoys a remarkable property.

Lemma 3.3. *Let u be as in Proposition 3.1. Set $S(x_0, r, u)$ for $S(r)$ defined by the sphere centered at $x_0 \in \partial\{u > 0\}$. Suppose $x_k \in \partial\{u > 0\}$ such that $x_k \to x_0$. Then*

$$\limsup_{x_k \to x_0} S(x_k, 0, u) \leq S(x_0, 0, u).$$

Proof. For given $\delta > 0$ there is $\rho_0 > 0$ such that $S(x_0, \rho, u) \leq S(x_0, 0, u) + \delta$ whenever $\rho < \rho_0$. Fix such ρ and choose k so large that $S(x_k, \rho, u) < \delta + S(x_0, \rho, u)$. From the monotonicity of $S(x_k, \rho, u)$ it follows that

$$\begin{aligned} S(x_k, 0, u) &\leq S(x_k, \rho, u) \leq \delta + S(x_0, \rho, u) \\ &\leq 2\delta + S(x_0, 0, u). \end{aligned}$$

First letting $x_k \to x_0$ and then $\delta \to 0$ the result follows. □

Lemma 3.4. *Let S be the monotone quantity in (3-11). Then the following hold*:

(i) $s(x_0, R, u) = (1/R^N)\int_0^R r^{N-1} S(x_0, r, u)\, dr$ *is monotone nondecreasing and*

$$\frac{d}{dR} s(x_0, u, R) = \frac{1}{R^{N+1}} \int_0^R r^N S'(x_0, u, r)\, dr \geq 0.$$

(ii) *If the solution $u \geq 0$ is degenerate at $x_0 \in \partial\{u > 0\}$ then the set $\{u > 0\}$ has well-defined Lebesgue density $\Theta(x_0, \{u > 0\})$ equal to*

$$\frac{1}{2M|B_1|} s(x_0, 0, u) = \frac{1}{2M|B_1|} \lim_{R\to 0} s(x_0, R, u).$$

(iii) *Suppose $x_k \in \partial\{u > 0\}$ such that $x_k \to x_0$. Then*

$$\limsup_{x_k \to x_0} s(x_k, 0, u) \leq s(x_0, 0, u).$$

Proof. It is easy to compute

$$\begin{aligned} s'(x_0, R, u) &= -\frac{N}{R^{N+1}} \int_0^R r^{N-1} S(x_0, r, u)\, dr + \frac{S(x_0, R, u)}{R} \\ &= -\frac{S(x_0, R, u)}{R} + \frac{1}{R^{N+1}} \int_0^R r^N S'(x_0, u, r)\, dr + \frac{S(x_0, R, u)}{R} \\ &= \frac{1}{R^{N+1}} \int_0^R r^N S'(x_0, r, u)\, dr. \end{aligned}$$

To prove the second claim notice that at the degenerate point x_0 we have $u(x) = o(|x - x_0|)$ by virtue of the subharmonicity of u. Consequently $⨍_{B_R(x_0)} |\nabla u|^2 = o(1)$ as $r \to 0$ by virtue of the Caccioppoli inequality. Therefore the only surviving term in S comes from $2M\chi_{\{u>0\}}$. The proof of the last claim is analogous to that of Lemma 3.3. □

Lemma 3.5. *Let $0 \in \partial\{u > 0\}$ and assume that $u_0 = rg(\sigma)$, $\sigma \in \mathbb{S}^{N-1}$, is a blow-up limit of u at 0 which is homogeneous function of degree 1. Then*

$$\int_{\mathbb{S}^{N-1}} |\nabla_\sigma g|^2 - (N-1)\int_{\mathbb{S}^{N-1}} g^2 \begin{cases} = 0 & \text{if } \partial\{u > 0\} \text{ is flat at } 0, \\ \leq 0 & \text{otherwise.} \end{cases}$$

Proof. Let (r, σ) be the spherical coordinates; then the Laplacian takes the form

$$\Delta u_\varepsilon = \partial^2_{rr} u_\varepsilon + \frac{N-1}{r}\partial_r u_\varepsilon + \frac{1}{r^2}\Delta_{\mathbb{S}^{N-1}} u_\varepsilon.$$

Multiply both sides of Δu_ε by $r^{N-1}u_\varepsilon$ and integrate over $[0, R] \times \mathbb{S}^{N-1}$ to get

$$\begin{aligned} I_1(u_{\varepsilon_j}) &:= \int_0^R \int_{\mathbb{S}^{N-1}} u_\varepsilon \partial^2_{rr} u_\varepsilon r^{N-1}\, d\sigma\, dr \\ &= R^{N-1}\int_{\mathbb{S}^{N-1}} u_\varepsilon \partial_r u_\varepsilon - \int_0^R \int_{\mathbb{S}^{N-1}} \big[(\partial^2_r u_\varepsilon)^2 r^{N-1} + (N-1)\partial_r u_\varepsilon u_\varepsilon r^{N-1}\big]\, d\sigma\, dr, \end{aligned}$$

$$I_2(u_{\varepsilon_j}) := \int_0^R \int_{\mathbb{S}^{N-1}} u_\varepsilon \partial_r^2 u_\varepsilon r^{N-2}\, d\sigma\, dr$$
$$= R^{N-2} \int_{\mathbb{S}^{N-1}} \frac{(u_\varepsilon)^2}{2} - (N-2) \int_0^R \int_{\mathbb{S}^{N-1}} \left[\frac{(u_\varepsilon)^2}{2} r^{N-3}\right] d\sigma\, dr,$$
$$I_3(u_{\varepsilon_j}) := \int_0^R \int_{\mathbb{S}^{N-1}} \triangle_{\mathbb{S}^{N-1}} u_\varepsilon u_\varepsilon r^{N-3}\, d\sigma\, dr = -\int_0^R \int_{\mathbb{S}^{N-1}} |\nabla_\sigma u_\varepsilon|^2 r^{N-3}\, d\sigma\, dr.$$

Choosing a converging sequence u_{ε_j} and letting $\varepsilon_j \to 0$ we get by virtue of Proposition A.2

$$\lim_{\varepsilon_j \to 0} \int_{B_R} \beta_{\varepsilon_j} u_{\varepsilon_j} = \lim_{\varepsilon_j \to 0} [I_1(u_{\varepsilon_j}) + (N-1) I_2(u_{\varepsilon_j}) + I_3(u_{\varepsilon_j})] \to I_1(u) + (N-1) I_2(u) + I_3(u).$$

Suppose that u_{R_k} is a blow-up sequence at the origin and $u_{R_k} \to u_0 = r g(\sigma)$; then

$$I_1(u_0) = R^N \int_{\mathbb{S}^{N-1}} g^2 - \frac{R^N}{N} \int_{\mathbb{S}^{N-1}} g^2 - \frac{N-1}{N} R^N \int_{\mathbb{S}^{N-1}} g^2 = 0,$$
$$I_2(u_0) = R^N \int_{\mathbb{S}^{N-1}} \frac{g^2}{2} - \frac{N-2}{N} R^N \int_{\mathbb{S}^{N-1}} \frac{g^2}{2} = \frac{R^N}{N} \int_{\mathbb{S}^{N-1}} g^2.$$

By Proposition A.5 and (A-2) there is a sequence $\delta_j \to 0$ such that $u_{\delta_j} \to u_0$ and

$$\lim_{\delta_j \to 0} \int_{B_1} \beta_{\delta_j} u_{\delta_j} \le \|\beta\|_\infty |\{x \in B_1 : 0 < u_{\delta_j} < \delta_j\}| \to 0,$$

provided that u is flat at 0. Hence we have

$$\lim_{\delta_j \to 0} \int_{B_1} \beta_{\delta_j} u_{\delta_j} = \frac{R^N}{N} \left[(N-1) \int_{\mathbb{S}^{N-1}} g^2 - \int_{\mathbb{S}^{N-1}} |\nabla_\sigma g|^2\right] \begin{cases} = 0 & \text{if } \partial\{u > 0\} \text{ is flat at } 0, \\ \ge 0 & \text{otherwise.} \end{cases} \qquad \square$$

4. Proof of Theorem A

The first part of the theorem follows from Proposition 3.1. Since u is not degenerate at the origin, by Propositions A.2 and A.5 $u_{\rho_k}(x) \to u_0(x)$ locally uniformly and by Proposition 3.1 u_0 is homogeneous of degree 1. Write $\triangle$ in polar coordinates (r, θ) to obtain

$$\Delta w = \frac{1}{r} \frac{\partial}{\partial r}(r w_r) + \frac{1}{r^2} \frac{\partial}{\partial \theta}(w_\theta).$$

In particular, writing $u_0 = r g(\theta)$, this yields a second-order ODE for g,

$$g + \ddot{g} = 0. \tag{4-1}$$

Suppose $g(0) = g(\theta_0) = 0$, $\theta_0 \in [0, 2\pi)$; then (4-1) implies that $g(\theta) = A \sin\theta$ for some constant A, consequently forcing $\theta_0 = \pi$. Hence, since $N = 2$, we obtain that u_0 must be linear; in other words the free boundary $\partial\{u_0 > 0\}$ is everywhere flat. This in turn implies that in two dimensions the singular set of the free boundary $\partial\{u_0 > 0\}$ is empty. Consequently, u_0 is linear in $\{u_0 > 0\}$ and $\{u_0 < 0\}$. From here parts (2) and (3) of Theorem A follow from [Caffarelli et al. 1997, Propositions 5.3 and 5.1].

So it remains to check (1). For the elliptic problem the only difference is that the limit function $M(x) = \lim_{\delta_j\to 0} \mathcal{B}_{\delta_j}(u_{\delta_j})$ cannot have nontrivial concentration on the free boundary coming from $\{x_1 < 0\}$, as opposed to the parabolic case studied in [Caffarelli et al. 1997]. Observe that $\nabla\mathcal{B}(u_{\delta_j}/\delta_j) = \nabla u_{\delta_j}\beta_{\delta_j}(u_{\delta_j}) = 0$ in $B_1 \setminus \{0 < u_{\delta_j} < \delta_j\}$. By Proposition A.5 and (A-2) there is sequence $0 < \lambda_j \to 0$ such that $(u_{\varepsilon_j})_{\lambda_j} \to u_0, \varepsilon_j/\lambda_j \to 0$ and $M(x) = M\chi_{\{x_1>0\}} + M_0\chi_{\{x_1<0\}}$. It follows from (A-3) that

$$\int_{\{x_1>0\}} M\,\partial_1\phi + \int_{\{x_1<0\}} M_0\,\partial_1\phi = \int_{\{x_1>0\}} \frac{\alpha^2}{2}\,\partial_1\phi \quad \text{for all } \phi \in C_0^\infty(B_1). \tag{4-2}$$

After integration by parts we obtain $M_0\int_{-1}^{1}\phi(0,x_2)\,dx_2 = \left(M - \frac{1}{2}\alpha^2\right)\int_{-1}^{1}\phi(0,x_2)\,dx_2$. This yields

$$M_0 = M - \frac{\alpha^2}{2}.$$

Next we claim that $M_0 = 0$. Suppose $M_0 > 0$; then $I_0 := \{t \in \mathbb{R} : \mathcal{B}(t) = M_0\} \neq \varnothing$ and there is $a \in (0,1)$ such that $I_0 \subset [a,1]$. Since $\mathcal{B}(t)$ is continuous and nondecreasing, it follows that there is $0 < a_0 < a$ such that $u_{\delta_j}(x)/\delta_j \in [a_0, 1]$ provided that j is sufficiently large.

Let

$$\mathcal{C} = \left\{x : \frac{u_{\delta_j}(x)}{\delta_j} \in [a_0, 1]\right\} \cap \{x_1 < 0\} \cap B_1.$$

Then,

$$\mathcal{C} \subset \{x \in B_1 : a_0\delta_j \le u_{\delta_j}(x) \le \delta_j\} \subset \{0 < u_{\delta_j} < 2\delta_j\} \cap B_1.$$

But $|\{0 < u_{\delta_j} < 2\delta_j\} \cap B_1| \to 0$, which implies that M_0 cannot be positive. □

5. The structure of the free boundary of blow-ups in $\mathbb{R}^3$

In this section we assume that $u \ge 0$ is a limit of u_{ε_j} solving $(\mathcal{P}_\varepsilon)$ for some sequence $\varepsilon_j \to 0$, u is nondegenerate at some $y_0 \in \partial\{u > 0\}$ and u_0 is a blow-up of u at y_0. Note that by Corollary 3.2 u_0 is a homogeneous function of degree 1. If u_0 is not a minimizer then it is natural to expect that the solutions of $(\mathcal{P}_\varepsilon)$ develop singularities in $\mathbb{R}^N$, $N \ge 3$.

We first prove a nondegeneracy result.

Lemma 5.1. *Let $x_0 \in \partial\{u_0 > 0\}$ be a free boundary point such that there is a ball $B \subset \{u_0 = 0\}$ touching $\partial\{u_0 > 0\}$ at x_0 and $\Theta(x_0, \{u_0 > 0\}) \ge \frac{1}{2}$. Then u_0 is nondegenerate at x_0 and*

$$u_0(x) = \sqrt{2M}(x - x_0)^+ + o(x - x_0).$$

Proof. Let $(u_0)_r = u_0(x_0 + rx)/r$. There is r_0 such that

$$(u_0)_r = 0 \quad \text{in } \{x_1 < -\delta\} \cap Q_1, \quad \text{for all } r \le r_0, \tag{5-1}$$

for some small $\delta > 0$, where $Q_1 = (-1,1)^3$ is the unit cube. Moreover, there is $\hat{r}_0 > 0$ such that

$$\frac{|\{(u_0)_r > 0\} \cap \{x_1 > 0\} \cap B_1|}{|B_1|} > \tfrac{1}{2} - \delta \quad \text{for all } r \le \hat{r}_0. \tag{5-2}$$

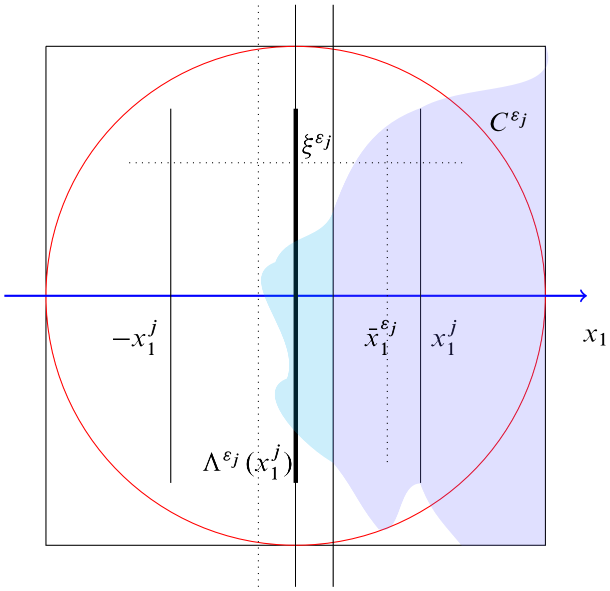

Figure 1. The construction of the point $(\bar{x}_1^{\varepsilon_j}, \xi^{\varepsilon_j})$. The purple region is C^{ε_j}.

Fix r with these two properties (5-1) and (5-2). There exists $\gamma > 0$ such that

$$\frac{|\{(u_0)_r > \gamma\} \cap \{x_1 > 0\} \cap B_1|}{|B_1|} > \tfrac{1}{2} - 2\delta. \tag{5-3}$$

Let $v^{\varepsilon_j} = (u_{\varepsilon_j})_r$, where $u_{\varepsilon_j} \to u_0$ (see Proposition A.5) and $A^{\varepsilon_j} = \{v^{\varepsilon_j} > \gamma/2\} \cap \{x_1 > 0\} \cap B_1$. Since $v^{\varepsilon_j} \to (u_0)_r$ uniformly (see Proposition A.2), it follows that there is $j_0(r)$ such that for $j \geq j_0(r)$ we have

$$|A^{\varepsilon_j}| > |B_1|\big(\tfrac{1}{2} - 2\delta\big). \tag{5-4}$$

Let $B^{\varepsilon_j} = \{x_1 \in (-1, -\delta)\} \cap Q_1$ and $-B^{\varepsilon_j} = \{x_1 \in (\delta, 1)\} \cap Q_1$. Let $C^{\varepsilon_j} = A^{\varepsilon_j} \cap (-B^{\varepsilon_j})$. Then we have

$$|C^{\varepsilon_j}| \geq |B_1|\big(\tfrac{1}{2} - 2\delta\big) > 0.$$

Define $\Lambda^{\varepsilon_j}(x_1) = \{x' : (x_1, x') \in C^{\varepsilon_j}\}$ and $f^{\varepsilon_j}(x_1) = |\Lambda^{\varepsilon_j}(x_1)|$. We claim that

$$\big|\big\{x_1 : f^{\varepsilon_j}(x_1) > |B_1|\big(\tfrac{1}{2} - 3\delta\big)\big\}\big| > 0.$$

Indeed, if the claim fails then we have

$$|B_1|\big(\tfrac{1}{2} - 2\delta\big) \leq |C^{\varepsilon_j}| = \int_\delta^1 f^{\varepsilon_j}(x_1)\, dx_1 \leq |B_1|\big(\tfrac{1}{2} - 3\delta\big),$$

which is a contradiction.

Hence there is $x_1^{\varepsilon_j} \in (\delta, 1)$ such that $f^{\varepsilon_j}(x_1^{\varepsilon_j}) > |B_1|\big(\tfrac{1}{2} - 3\delta\big)$. Now choose $0 < a' < a < b < b' < 1$ such that

$$\beta(s) > \kappa \quad \text{for all } s \in [a', b'].$$

Let $\varepsilon_j' = \varepsilon_j/r$. We claim that there is $\xi^{\varepsilon_j} \in \Lambda^{\varepsilon_j}$ and $\bar{x}_1^{\varepsilon_j}$ such that

$$\frac{v^{\varepsilon_j}}{\varepsilon_j'}(x_1^{\varepsilon_j}, \xi^{\varepsilon_j}) \in (a, b).$$

Indeed, for sufficiently large j we have

$$a > \frac{v^{\varepsilon_j}}{\varepsilon_j'}(-x_1^{\varepsilon_j}, x') = 0 \quad \text{for all } x' \in \Lambda^{\varepsilon_j}(x_1^{\varepsilon_j}),$$
$$\frac{v^{\varepsilon_j}}{\varepsilon_j'}(x_1^{\varepsilon_j}, x') > \frac{\gamma}{2\varepsilon_j'} > b \quad \text{for all } x' \in \Lambda^{\varepsilon_j}(x_1^{\varepsilon_j}),$$

provided that $j > j_1(r)$; see Figure 1. Hence from the mean value theorem we see that the claim is true. From the uniform Lipschitz continuity of the functions v^{ε_j} it follows that there is a constant $c_0 > 0$ such that

$$\frac{v^{\varepsilon_j}}{\varepsilon_j'}(x_1, x') \in (a', b') \quad \text{if } |x_1 - \bar{x}_1^{\varepsilon_j}| < \varepsilon_j' c_0, \;\; x' \in \Lambda^{\varepsilon_j}(x_1^{\varepsilon_j}).$$

Consequently we have

$$\int_{B_1} \beta_{\varepsilon_j'}(v^{\varepsilon_j}) \geq \frac{\kappa}{\varepsilon_j'} \int_{|x_1 - \bar{x}_1^{\varepsilon_j}| < \varepsilon_j' c_0} |\Lambda^{\varepsilon_j}(x_1)| dx_1 \geq \frac{\kappa}{\varepsilon_j'}(1 - 3\delta) 2 c_0 \varepsilon_j' = 2\kappa |B_1|(1 - 3\delta) c_0 := \widetilde{C}.$$

Now the nondegeneracy follows from the proof of Part II of Theorem 6.3 in [Caffarelli et al. 1997]. The asymptotic expansion follows from Theorem A and Proposition 3.1. □

Remark 5.2. Note that under the weaker assumption $\Theta(x_0, \{u_0 > 0\}) > 0$ the argument in the proof of Lemma 5.1 still works. However for a self-crossing free boundary [Weiss 2003] (see Figure 2) the assumptions of Lemma 5.1 may not be satisfied.

As an immediate corollary we have:

Corollary 5.3. *Let $x_0 \in \partial\{u_0 > 0\}$ be a point of reduced boundary. Then u_0 is nondegenerate at x_0.*

Proof. Suppose that $0 \in \partial\{u_0 > 0\}$ and $\partial\{(u_0)_r > 0\} \subset B_2 \cap \{|x \cdot e| < \varepsilon\}$ for some unit vector e and small $\varepsilon > 0$. Here $(u_0)_r = u_0(rx)/r$. Consider the family of balls $B_{1/2}(et), t \in [-\varepsilon, \varepsilon]$. Then there

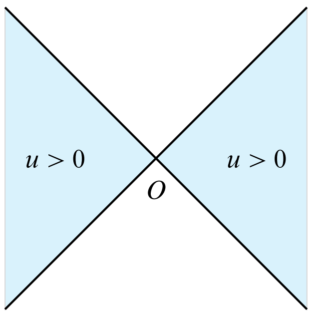

Figure 2. Possible self-crossing free boundary which fails to satisfy the conditions of Lemma 5.1.

is $t_\varepsilon \in [-\varepsilon, \varepsilon]$ such that $B_{1/2}(et_\varepsilon)$ touches the free boundary at some point $z_0 \in B_1$ provided that ε is sufficiently small. Let $\nu_0 = t_\varepsilon e$. Introduce the barrier function

$$w(x) = \frac{\varphi\big(\frac{1}{2}\big) - \varphi(|x-\nu_0|)}{\varphi\big(\frac{1}{2}\big) - \varphi(1)} \sup_{B_1(\nu_0)} u_0,$$

where $\varphi(|x|) = 1/|x|^{N-2}$. We have $\Delta(u_0 - w) = \Delta u_0 \ge 0$ in $D = B_1(\nu_0) \setminus B_{1/2}(\nu_0)$ and $u_0 - w \le 0$ on ∂D. From the maximum principle we infer that $u_0 \le w$ in D. But we have that the maximum of $u_0 - w$ is realized at z_0. Hence from the Hopf lemma we get

$$-\sqrt{2M} = \partial_{\nu_0} u_0(z_0) > \partial_{\nu_0} w(z_0) = -\big|\varphi'\big(\tfrac{1}{2}\big)\big| \frac{\sup_{B_1(\nu_0)} u_0}{\varphi\big(\frac{1}{2}\big) - \varphi(1)}$$

or

$$\sup_{B_1(\nu_0)} (u_0)_r \ge \sqrt{2M} \frac{\varphi\big(\frac{1}{2}\big) - \varphi(1)}{\big|\varphi'\big(\frac{1}{2}\big)\big|}. \qquad \square$$

In the following definition we let $\Omega^+(u) = \{u > 0\}$ and $\Omega^-(u) = \{u < 0\}$. Moreover, let

$$G(u_\nu^+, u_\nu^-) := (u_\nu^+)^2 - (u_\nu^-)^2 - 2M, \tag{5-5}$$

where u_ν^+ and u_ν^- are the normal derivatives in the inward direction ν to $\partial\Omega^+(u)$ and $\partial\Omega^-(u)$, respectively. For more details see [Caffarelli and Salsa 2005, Definition 2.4].

Definition 5.1. Let Ω be a bounded domain of $\mathbb{R}^N$ and let u be a continuous function in Ω. We say that u is a viscosity solution in Ω if:

(i) $\Delta u = 0$ in $\Omega^+(u)$ and $\Omega^-(u)$.

(ii) Along the free boundary $\partial\{u > 0\}$, the function u satisfies the free boundary condition in the sense that:

(a) If at $x_0 \in \partial\{u > 0\}$ there exists a ball $B \subset \Omega^+(u)$ such that $x_0 \in \partial B$ and

$$u^+(x) \ge \alpha\langle x - x_0, \nu\rangle^+ + o(|x - x_0|) \quad \text{for } x \in B, \tag{5-6}$$

$$u^-(x) \le \beta\langle x - x_0, \nu\rangle^- + o(|x - x_0|) \quad \text{for } x \in B^c, \tag{5-7}$$

for some $\alpha > 0$ and $\beta \ge 0$, with equality along every nontangential domain, then the free boundary condition is satisfied:

$$G(\alpha, \beta) \le 0.$$

(b) If at $x_0 \in \partial\{u > 0\}$ there exists a ball $B \subset \Omega^-(u)$ such that $x_0 \in \partial B$ and

$$u^-(x) \ge \beta\langle x - x_0, \nu\rangle^- + o(|x - x_0|) \quad \text{for } x \in B,$$

$$u^+(x) \le \alpha\langle x - x_0, \nu\rangle^+ + o(|x - x_0|) \quad \text{for } x \in \partial B,$$

for some $\alpha \ge 0$ and $\beta > 0$, with equality along every nontangential domain, then

$$G(\alpha, \beta) \ge 0.$$

In our case $\beta = 0$ and we have only u^+. However, one has to check that the free boundary conditions (a) and (b) in Definition 5.1 are satisfied.

Lemma 5.4. *Let u_0 be a blow-up of u at some nondegenerate point such that $\Theta(x, \{u_0 > 0\}) \geq \frac{1}{2}$ for every $x \in \partial\{u_0 > 0\}$. Then u_0 is a viscosity solution in the sense of Definition 5.1.*

Proof. We have to show that the properties (a), (b) in Definition 5.1 hold. Suppose that $B \subset \{u_0 > 0\}$ touches $\partial\{u_0 > 0\}$ at some point x_0. Then it follows from Hopf's lemma that u_0 is nondegenerate at x_0. Consequently, if u_{00} is a blow-up at x_0 then by Theorem A $u_{00}(x) = \alpha x_1^+$ after some rotation of coordinate system. Moreover $0 < \alpha \leq \sqrt{2M}$. Hence $G(\alpha, 0) \leq 0$.

Now suppose that $B \subset \{u_0 = 0\}$ and B touches $\partial\{u_0 > 0\}$ at z_0. By Lemma 5.1 u_0 is nondegenerate at z_0. Theorem A implies that any blow-up u_{00} of u_0 at z_0 must be $u_{00}(x) = \sqrt{2M}x_1^+$ after some rotation of coordinates. Hence $G(\sqrt{2M}, 0) \geq 0$. □

5A. *Properties of $\partial\{u_0 > 0\}$.* We want to study the properties of g. We first prove a Bernstein-type result which is a simple consequence of a refinement of the Alt–Caffarelli–Friedman monotonicity formula [Alt et al. 1984b; Caffarelli et al. 2000].

Lemma 5.5. *Let $u \geq 0$ be a limit of solutions to $(\mathcal{P}_\varepsilon)$. Let $u_0 = rg(\sigma)$, $\sigma \in \mathbb{S}^{N-1}$, be a nontrivial blow-up of u at some free boundary point. If there is a hemisphere containing* supp g *then the graph of u_0 is a half-plane.*

Proof. Without loss of generality we assume supp $g \subset \mathbb{S}_+^{N-1} = \{X \in \mathbb{S}^{N-1} : x_N \geq 0\}$. Let $v(x_1, \dots, x_N) = u(x_1, \dots, -x_N)$ be the reflection of u with respect to the hyperplane $x_N = 0$. Then v is a nonnegative subharmonic function satisfying the requirements of [Caffarelli et al. 2000, Lemma 2.3]. Thus

$$\Phi(r) = \frac{1}{r^4} \int_{B_r} \frac{|\nabla u_0|^2}{|x|^{N-2}} \int_{B_r} \frac{|\nabla v|^2}{|x|^{N-2}}$$

is nondecreasing in r. Moreover

$$\Phi'(r) \geq \frac{2\Phi(r)}{r} A_r, \quad A_r = \frac{C_N}{r^{N-1}} \operatorname{Area}(\partial B_r \setminus (\operatorname{supp} u_0 \cup \operatorname{supp} v)).$$

Thus, if supp g digresses from the hemisphere by size $\delta > 0$ then $A_r \geq c(\delta) > 0$. Hence integrating the differential inequality for Φ we see that Φ grows exponentially, which is a contradiction since in view of Proposition A.3 u_0 is Lipschitz and hence Φ must be bounded. □

It is convenient to define the following subsets of the free boundary:

$$\Gamma_{1/2} = \{x \in \partial\{u_0 > 0\} \text{ such that } \Theta(x, \{u_0 > 0\}) = \tfrac{1}{2}\}, \tag{5-8}$$

$$\Gamma_1 = \{x \in \partial\{u_0 > 0\} \text{ such that } \Theta(x, \{u_0 > 0\}) = 1\}, \tag{5-9}$$

where $\Theta(x, D)$ denotes the Lebesgue density of D at x. We will see that $\Theta(x, \{u_0 > 0\})$ exists at every nondegenerate point and equals either 1 or $\frac{1}{2}$.

Lemma 5.6. *Assume $N = 3$. Let $x_0 \in \partial\{u_0 > 0\} \setminus \{0\}$ be a nondegenerate free boundary point such that the lower Lebesgue density satisfies $\Theta_*(x_0, \{u_0 \equiv 0\}) > 0$. Then there is a unit vector ν_0 such that*

$$u_0(x) = \sqrt{2M}[(x - x_0) \cdot \nu_0]^+ + o(x - x_0). \tag{5-10}$$

In particular, $x_0 \in \Gamma_{1/2}$.

Proof. Set $v_k = u_0(x_0 + \rho_k x)/\rho_k$. Since u_0 is nondegenerate at x_0 it follows from a customary compactness argument that $v_k \to v$ and by virtue of Corollary 3.2 v is a homogeneous function of degree 1. We have

$$\begin{aligned} \frac{u_0(x_0 + \rho_k x)}{\rho_k} &= u_0(\rho_k^{-1} x_0 + x) = \nabla u_0(\rho_k^{-1} x_0 + x)(\rho_k^{-1} x_0 + x) \\ &= \frac{1}{\rho_k} \nabla u_0(x_0 + \rho_k x)(x_0 + \rho_k x), \end{aligned} \tag{5-11}$$

where the last line follows from the zero-degree homogeneity of the gradient; hence

$$\rho_k v_k(x) = \nabla u_0(x_0 + \rho_k x)(x_0 + \rho_k x) = \nabla v_k(x)(x_0 + \rho_k x). \tag{5-12}$$

By the Lipschitz continuity of u_0 it follows that v_k is locally bounded. Consequently, for a suitable subsequence of ρ_k we have $v_{k_j} \to v$ and $\nabla v(x) x_0 = 0$. Without loss of generality we may assume that x_0 is on the x_3-axis, implying that v depends only on x_1 and x_2. Applying Proposition A.6 and Corollary 3.2 we conclude that $S(x_0, r, u_0)$ is nondecreasing and thus v must be homogeneous of degree 1. Indeed, there is a sequence $\delta_j \to 0$ such that $(u_{\varepsilon_j})_{\lambda_j} \to v$, $\delta_j = \varepsilon_j/\lambda_j$ by Proposition A.6.

Finally, applying Theorem A and the assumption $\Theta_*(x_0, \{u_0 = 0\}) > 0$ we see that v must be a half-plane solution. It remains to note that the approximate tangent of $\partial\{u_0 > 0\}$ at x_0 is unique and this completes the proof. □

Lemma 5.7. *We want to show that $\Theta(x, \{u_0 > 0\}) \geq \frac{1}{2}$ in some neighborhood of x_0. Let $x_0 \in \Gamma_{1/2}$. Then there exists $r_0 > 0$ such that $B_{r_0}(x_0) \cap \partial\{u_0 > 0\}$ is a $C^{1,\alpha}$ surface.*

Proof. Let $y_0 \in \partial\{u_0 > 0\}$ be a degenerate point. Suppose there is $\rho > 0$ such that u_0 is degenerate at every point of $B_\rho(y_0) \cap \partial\{u_0 > 0\}$. Since $\operatorname{supp} \Delta u_0 \subset \partial\{u_0 > 0\}$, it follows that $u_0 \equiv 0$ in $B_\rho(y_0)$. Consequently, there is a sequence of nondegenerate points $y_k \to y_0$. Note that if y_k is a nondegenerate point then by Theorem A the Lebesgue density satisfies $\Theta(y_k, \{u_0 > 0\}) \geq \frac{1}{2}$.

Let u_{00}^k be a blow-up of u_0 at y_k. By Proposition A.6 for fixed k there are $\delta_j^k \to 0$ such that $(u_{\varepsilon_j^k})_{\lambda_j^k} \to u_{00}^k$, $\delta_j^k = \varepsilon_j^k/\lambda_j^k$. Thus applying Theorem A it follows that u_{00}^k is a half-plane solution or a wedge.

From scaling properties of Spruck's monotonicity formula and Lemma 3.4 we get

$$s(0, y_k, u_0) = s(1, 0, u_{00}^k) = 2M \operatorname{vol}(B_1 \cap \{u_{00}^k > 0\}) = \begin{cases} 2\pi M & \text{if } y_k \text{ is a wedge point,} \\ \pi M & \text{otherwise.} \end{cases} \tag{5-13}$$

Then applying Corollary 3.2 to $u_{\delta_j^k}$ and using the semicontinuity of S, Lemma 3.3 together with Lemma 3.5, we have

$$2M \operatorname{vol}(B_1 \cap \{u_{00}^k > 0\}) = \limsup_{y_k \to y_0} s(0, y_k, u_0) \leq s(0, y_0, u_0) = 2M\pi \Theta(x_0, \{u_0 > 0\}). \tag{5-14}$$

Therefore we conclude that $\Theta(x, \{u_0 > 0\}) \geq \frac{1}{2}$ for every free boundary point x in some neighborhood of x_0. By virtue of Lemma 5.4 u_0 is a viscosity solution which is flat x_0. Applying the "flatness-implies-$C^{1,\alpha}$" regularity results from [Caffarelli 1987; 1989] the lemma follows. □

Next we prove a representation formula for $\triangle u_0$.

Lemma 5.8. *Let u_0 be as in Lemma 5.5. Then*:

(i) $\mathcal{H}^2(\Gamma_{1/2} \cap B_R) < \infty$ *for any* $R > 0$.

(ii) *Away from* Γ_1 *the following representation formula holds*:

$$\triangle u_0 = \sqrt{2M}\,\mathcal{H}^2 \llcorner \Gamma_{1/2}.$$

Proof. (i) For given $x \in \Gamma_{1/2}$ there is a $\tilde{\rho}_x > 0$ such that

$$\sup_{B_r(x)} u_0 \geq \sqrt{M}\,r, \quad r \in (0, \tilde{\rho}_x). \tag{5-15}$$

This follows from the asymptotic expansion in Lemma 5.6. Consequently, there is $\rho'_x > 0$ such that

$$\int_{B_r(x)} \triangle u_0 \geq \sqrt{M}\,r^2, \quad r \in (0, \rho'_x). \tag{5-16}$$

Indeed, if this inequality is false then there is a sequence $r_j \searrow 0$ such that

$$\int_{B_{r_j}(x)} \triangle u_0 < \sqrt{M}\,r_j^2.$$

Set $v_j(x) = u_0(x + r_j x)/r_j$. By (5-15) $\sup_{B_1} v_j(x) \geq \sqrt{M}$. Moreover, it follows from Lemma 5.6 that $v_j(x) \to \sqrt{2M}\,x_1^+$ in a suitable coordinate system, while $\int_{B_1} \triangle v \leq \sqrt{M}$. However, $\int_{B_1} \triangle x_1^+ = \sqrt{2M}\,\frac{\pi}{2}$ and this is in contradiction with the former inequality. Putting $\bar{\rho}_x = \min(\rho'_x, \tilde{\rho}_x)$ we see that the collection of balls $\mathcal{F} = \bigcup B_{\rho_x}(x),\ x \in \Gamma_{1/2} \cap B_R,\ \rho_x < \bar{\rho}_x$, is a Besicovitch-type covering of $\Gamma_{1/2} \cap B_R$. Consequently, there is a positive integer $m > 0$ and subcoverings $\mathcal{F}_1, \dots, \mathcal{F}_m$ such that the balls in each $\mathcal{F}_i$, $1 \leq i \leq m$, are disjoint and $\Gamma_{1/2} \cap B_R \subset \bigcup_{i=1}^m \mathcal{F}_i$. We have from (5-16)

$$4\pi R^2 \|\nabla u_0\|_\infty \geq \int_{\partial B_R} \partial_\nu u_0 \geq \int_{B_{\rho_x}(x) \in \mathcal{F}_i} \triangle u_0 = \sum_{B_{\rho_x}(x) \in \mathcal{F}_i} \int_{B_{\rho_x}(x)} \triangle u_0 \geq m\sqrt{M} \sum_{B_{\rho_x}(x) \in \mathcal{F}_i} \rho_x^2.$$

This yields

$$\sum_{B_{\rho_x}(x) \in \bigcup_{i=1}^m \mathcal{F}_i} \rho_x^2 \leq \frac{4\|\nabla u_0\|_\infty \pi R^2}{m\sqrt{M}}. \tag{5-17}$$

Given $\delta > 0$ small, suppose there is $x \in \Gamma_{1/2}$ such that $\bar{\rho}_x \geq \delta$. Then we choose $\rho_x < \delta$. Thus, in any case we can assume that $\rho_x < \delta$. In view of (5-17) this implies that the δ-Hausdorff premeasure is bounded independently of δ. This proves (i).

(ii) From the estimate

$$\sqrt{M}r^2 \le \int_{B_r(x)} \Delta u_0 \le 4\pi r^2 \|\nabla u_0\|, \quad r \in (0, \bar{\rho}_x), \ \ B_r(x) \cap \Gamma_{1/2} \subset \Gamma_{1/2},$$

we see that there is a positive bounded function q such that $\Delta u_0 = q\mathcal{H}^2 \llcorner \Gamma_{1/2}$. Using Lemma 5.6 we conclude that $q = \sqrt{2M}$. □

Next we prove the full nondegeneracy of u_0 near $\Gamma_{1/2}$.

Lemma 5.9. *Let u_0 be as above and $x_0 \in \Gamma_{1/2}$. Then for any $B_r(x)$ such that $x \in \partial\{u_0 > 0\}$, $B_r(x) \cap \partial\{u_0 > 0\} \subset \Gamma_{1/2}$, we have*

$$\sup_{B_r(x)} u_0 \ge \sqrt{2M}\pi r.$$

Proof. By a direct computation we have

$$r^{-2} \int_{\partial B_r(x)} u_0 = \int_0^r \frac{dt}{t^2} \int_{B_t(x)} \Delta u_0 \ge \int_0^r \frac{1}{t^2} \sqrt{2M}\pi t^2 = \sqrt{2M}\pi r,$$

where the inequality follows from the representation formula and the fact that $\partial\{u_0 > 0\}$ is a cone; hence for all $t \in (0, r)$ we have $\mathcal{H}^2(B_r(x) \cap \Gamma_{1/2}) \ge \pi t^2$. It remains to note that $r^{-2} \int_{\partial B_r(x)} u_0 \le \sup_{B_r(x)} u_0$. □

5B. *Weak solutions.* Combining Lemmas 5.8 and 5.9 as well as Propositions A.2(iii) and A.3(i) we see that u_0 is a weak solution near $\Gamma_{1/2}$ in the sense of [Alt and Caffarelli 1981, Definition 5.1]. Furthermore, $\partial\{u_0 > 0\} \setminus \{0\}$ is flat at each point.

Lemma 5.10. *The blow-up u_0 is a weak solution in the Alt–Caffarelli sense away from Γ_1. Furthermore, $\Gamma_{1/2}$ is smooth.*

Proof. All conditions in [Alt and Caffarelli 1981, Definition 5.1] are satisfied and u_0 is flat at every point $z_0 \in \partial\{u_0 > 0\} \setminus \{0\}$ thanks to (5-10). Applying Theorem 8.1 of the same paper we infer that $\Gamma_{1/2}$ is smooth at every $z_0 \in \partial\{u_0 > 0\} \setminus \{0\}$. □

5C. *Minimal perimeter.* In this section we prove that the local perturbations $S' \subset \{u_0 > 0\}$ of a portion $S \subset \Gamma_{1/2}$ have larger $\mathcal{H}^2$ measure than S. This can be seen from the estimate $|\nabla u_0(x)| \le \sqrt{2M}$, which follows from Lemma A.7. Since by Lemma 5.10 on $\Gamma_{1/2}$ the free boundary condition $|\nabla u_0| = \sqrt{2M}$ is satisfied in the classical sense, it follows that

$$0 = \int_D \Delta u_0 = \int_S \partial_\nu u_0 + \int_{S'} \partial_\nu u_0 = \sqrt{2M}\mathcal{H}^2(S) + \int_{S'} \partial_\nu u_0,$$

where $D \subset \{u_0 > 0\}$ such that $\partial D = S \cup S'$. But $\left|\int_{S'} \partial_\nu u_0\right| \le \sqrt{2M}\mathcal{H}^2(S')$ and thereby

$$\mathcal{H}^2(S) \le \mathcal{H}^2(S'). \tag{5-18}$$

The estimate for the perimeter can be reformulated as follows:

Theorem 5.11. *Let $N = 3$. Then the components of $\Gamma_{1/2}$ are surfaces of nonpositive outward mean curvature. In particular, $\Gamma_{1/2}$ is a union of smooth convex surfaces.*

Proof. Since u_0 is a weak solution, by Lemma 5.10 $\Gamma_{1/2}$ is smooth. If $z_0 \in \Gamma_{1/2}$ then choosing the coordinate system in $\mathbb{R}^3$ so that x_3-axis has the direction of the inward normal of $\{u_0 > 0\}$ at z_0 and considering the free boundary near z_0 as a graph $x_3 = h(x_1, x_2)$, we can consider the one-sided variations of the surface area functional. Indeed, let $\mathcal{D} \subset \mathbb{R}^2$ be a open bounded domain in the x_1x_2-plane containing z_0 and assume $t > 0,\ 0 \le \psi \in C_0^\infty(\mathcal{D})$. Then from (5-18) we have

$$
\begin{aligned}
0 &\ge \frac{1}{t}\int_{\mathcal{D}}\left[\sqrt{1+|\nabla h|^2} - \sqrt{1+|\nabla(h-t\psi)|^2}\right] \\
&= \int_{\mathcal{D}} \frac{2\nabla h\nabla\psi - t|\nabla\psi|^2}{\sqrt{1+|\nabla h|^2}+\sqrt{1+|\nabla(h-t\psi)|^2}} \\
&\to \int_{\mathcal{D}} \frac{\nabla h\nabla\psi}{\sqrt{1+|\nabla h|^2}} \quad \text{as } t \to 0.
\end{aligned} \tag{5-19}
$$

Therefore

$$
\operatorname{div}\left(\frac{\nabla h}{\sqrt{1+|\nabla h|^2}}\right) \ge 0
$$

and, noting that $\Gamma_{1/2}$ is a cone, the result follows. □

5D. *Full nondegeneracy.*

Lemma 5.12. *Assume that $N = 3$ and let u_0 be a nontrivial blow-up of u such that the measure-theoretic boundary of $\{u_0 > 0\}$ is nonempty. Then $\partial\{u_0 > 0\} \setminus \{0\} \subset \Gamma_{1/2}$. In particular the set of degenerate points of $\partial\{u_0 > 0\}$ is empty.*

Proof. Let u_0 be a blow-up of u at 0. Since u is nondegenerate at 0, it follows that u_0 does not vanish identically. Hence there is a ball $B \subset \{u_0 > 0\}$ touching $\partial\{u_0 > 0\}$ at some point $z_0 \in \partial\{u_0 > 0\} \cap B$. By Hopf's lemma, the Lipschitz estimate Proposition A.3(i) and asymptotic expansion [Caffarelli 1989, Lemma A1] it follows that u_0 is not degenerate at z_0. Consequently, the set of nondegenerate points of u_0 is not empty.

Suppose that S is a component of $\partial\{u_0 > 0\}$ containing a point of measure-theoretic boundary of $\{u_0 > 0\}$. Note that by Lemma 5.7 and Theorem 5.11 S is a smooth convex surface. Let $x_0 \in \partial S,\ x_0 \ne 0$. Then either (a) $x_0 \in \Gamma_1$ or (b) u_0 is degenerate at x_0.

We first analyze the case (a). Let ℓ be the ray passing through x_0 and Π the tangent half-plane to S along ℓ. First note that u_0 is nondegenerate at x_0 because

$$
\int_{B_r(x_0)} \Delta u_0 \ge \int_{B_r(x_0)\cap S} \Delta u_0 \ge \sqrt{2M}\,\mathcal{H}^2(S \cap B_r(x_0)) \ge \sqrt{2M}\,\frac{\pi r^2}{2}
$$

for sufficiently small r. Consequently

$$
\frac{1}{R^2}\int_{\partial B_R} u_0 = \int_0^R \frac{1}{r^2}\int_{B_r(x_0)} \Delta u_0 \ge \sqrt{2M}\,\frac{\pi r^2}{2} R. \tag{5-20}
$$

Let u_{00} be a blow-up of u_0 at x_0. Then from Theorem A it follows that u_{00} is two-dimensional. Moreover $\Pi \subset \partial\{u_{00} > 0\},\ \{u_{00} > 0\}$ has unit density at 0, and the interior of $\{u_{00} = 0\}$ near Π is not empty. Note

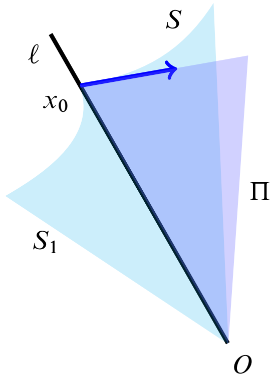

Figure 3. The structure of the free boundary near the point x_0.

that the interior of the set $\{u_0 = 0\}$ propagates to x_0 along another component S_1 of measure-theoretic boundary; see Figure 3. Consequently, near Π, we have $u_{00}(x) = \sqrt{2M}x_1^+$ after some rotation of coordinates. From the unique continuation theorem it follows that $u_{00}(x) = \sqrt{2M}x_1^+$ everywhere, which is in contradiction with the fact that $\{u_{00} > 0\}$ has unit density at 0.

As for the case (b), (5-20) shows that u_0 is nondegenerate at x_0 as long as x_0 is on the boundary of S. □

5E. *Properties of* $\Gamma_{1/2}$.

Lemma 5.13. *Suppose u_0 is not degenerate at $x_0 \in \partial\{u_0 > 0\} \setminus \{0\}$, such that $\Theta_*(x_0, \{u_0 = 0\}) > 0$. Then there is a unique component $\mathfrak{C}$ of $\partial\{u_0 > 0\}$ containing x_0 such that $\mathfrak{C} \subset \Gamma_{1/2}$.*

Proof. We only have to show the uniqueness of $\mathfrak{C}$; the rest follows from Lemmas 5.6 and 5.7. Suppose there are two components of $\partial\{u_0 > 0\} \setminus \{0\}$, $\mathfrak{C}_1$ and $\mathfrak{C}_2$, containing x_0. From the dimension-reduction argument as in the proof of Lemma 5.6, it follows that $\mathfrak{C}_1$ and $\mathfrak{C}_2$ have the same approximate tangent plane at x_0. This is in contradiction with our assumption $\Theta_*(x_0, \{u_0 = 0\}) > 0$. □

Lemma 5.14. *Let $\mathfrak{C}$ be a component of $\partial\{u_0 > 0\}$ such that $\mathfrak{C} \cap \Gamma_{1/2} \neq \varnothing$. Then $\mathfrak{C} \setminus \Gamma_{1/2} = \varnothing$; in other words all points of $\mathfrak{C}$ are in $\Gamma_{1/2}$.*

Proof. By Lemma 5.12 $\mathfrak{C}$ cannot have degenerate points; thus we have to show that $\Gamma_{1/2}$ cannot have limit points in Γ_1. Note that $\Gamma_{1/2}$ is of locally finite perimeter (see Lemma 5.8(i)) and hence locally it is a countable union of convex surfaces. Let $x_0 \in \Gamma_1 \cap \mathfrak{C}$ be a limit point of $\Gamma_{1/2} \cap \mathfrak{C}$. The generatrix of the cone $\partial\{u_0 > 0\}$ passing through x_0 splits $\mathfrak{C}$ into two parts, one of which must be convex near x_0 because by assumption x_0 is a limit point of $\Gamma_{1/2}$; see Theorem 5.11. The set $\{u_0 = 0\}^\circ$ propagates to x_0 because $\Gamma_{1/2}$ is a subset of reduced boundary. Thus, there is another subset of $\Gamma_{1/2}$ approaching x_0, and it is a part of the topological boundary of $\{u_0 = 0\}^\circ$. Therefore, the ray passing through x_0 is on the boundaries of two convex pieces of $\partial\{u_0 > 0\}$ (near x_0). Note that if these pieces of $\Gamma_{1/2}$ contain flat parts then from the unique continuation theorem we infer that $\partial\{u_0 > 0\}$ cannot have singularity at 0. Thus, they cannot contain flat parts and consequently the density of $\{u_0 > 0\}$ at x_0 cannot be 1, because by convexity of $\Gamma_{1/2}$ it follows that $\{u_0 \equiv 0\}^\circ$ has positive density at x_0. But this is in contradiction with the assumption $x_0 \in \Gamma_1$. □

Summarizing we have:

Proposition 5.15. *Let u_0 be as above and $N = 3$. Then $\partial\{u_0 > 0\}\setminus\{0\}$ is a union of smooth convex cones.*

5F. ***Proof of Theorem B.*** The first part of Theorem B follows from Lemma 5.12, while the second part is a corollary of Lemma 5.14 since $\Gamma_{1/2}$ coincides with the reduced boundary. Finally, the last part follows from Lemma 5.10, because by Lemma 5.14 the reduced boundary propagates instantaneously in $\partial\{u_0 > 0\}$.

6. Proof of Theorem C

6A. ***Inverse Gauss map and the support function.*** Suppose $N = 3$ and $u = rg(\theta, \phi)$, where

$$x = r(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta).$$

Then

$$\triangle u_0 = \frac{1}{r}\left(g_{\theta\theta} + \frac{\cos\theta}{\sin\theta}g_\theta + \frac{g_{\phi\phi}}{\sin^2\theta} + 2g\right).$$

Note that

$$\triangle_{\mathbb{S}^2} g = g_{\theta\theta} + \frac{\cos\theta}{\sin\theta}g_\theta + \frac{g_{\phi\phi}}{\sin^2\theta}$$

is the Laplace–Beltrami operator. Thus we get

$$\triangle_{\mathbb{S}^2} g + 2g = 0. \tag{6-1}$$

Let $H(n)$, $n \in \mathbb{S}^{N-1}$, be the Minkowski support function of some hypersurface $\mathcal{M}$. $H(n)$ is the distance between the point on $\mathcal{M}$ with normal n and the origin. It is known [Alexandroff 1939] that the eigenvalues of the matrix

$$\nabla^2_{ij} H(n) + \delta_{ij} H(n)$$

are the principal radii of curvature of the surface determined by H, where the second-order derivatives are taken with respect to an orthonormal frame at $n \in \mathbb{S}^{N-1}$. The support function uses the inverse of the Gauss map to parametrize the surface as

$$H(n) = G^{-1}(n)\cdot n.$$

Furthermore, we have the following formula for the Gauss curvature K [Alexandroff 1939]:

$$\frac{1}{K} = \det(\nabla^2_{ij} H(n) + \delta_{ij} H(n)). \tag{6-2}$$

The Gauss map is a local diffeomorphism whenever $K \neq 0$ [Langevin and Rosenberg 1988]. Since $u_0 = rg$ is harmonic in $\{u_0 > 0\}$, we infer that g is smooth on $\mathbb{S}^2 \cap \{g > 0\}$.

Remark 6.1. In higher dimensions (6-1) becomes

$$\Delta_{\mathbb{S}^{N-1}} g + (N-1)g = \sum_{i=1}^{N-1} \frac{1}{k_i} = \frac{\sigma_{N-2}(k)}{\sigma_{N-1}(k)} = 0, \quad \sigma_m = \sum_{i_1 < i_2 < \cdots < i_m} k_{i_1} k_{i_2} \cdots k_{i_m}, \tag{6-3}$$

where $\sigma_m(k)$ is the m-th elementary symmetric function and k_i, $i=1,\dots,N-1$, are the principal curvatures. Observe that any positive function $g>0$ satisfying the equation $\Delta_{\mathbb{S}^{N-1}}g+(N-1)g=0$ defines an $(N-2)$-minimal surface (i.e., $\sigma_{N-2}(k)=0$) provided that the Gauss curvature satisfies $\sigma_{N-1}\neq 0$. From here we infer that the spherical parts of the homogeneous stationary points of the Alt–Caffarelli functional are support functions of an $(N-2)$-capillary surface in $\mathbb{S}^{N-1}$, because they are solutions to (6-3).

6B. ***Catenoid is a solution.*** Alt and Caffarelli [1981, page 110] constructed a weak solution which is not a minimizer. Their solution can be given explicitly as follows: let

$$x=r(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta)$$

and take

$$u(x)=r\max\left(\frac{f(\theta)}{f'(\theta_0)},0\right),$$

where

$$f(\theta)=2+\cos\theta\log\left(\frac{1-\cos\theta}{1+\cos\theta}\right)=2+\cos\theta\log\left(\tan^2\frac{\theta}{2}\right)$$

and θ_0 is the unique zero of f between 0 and $\frac{\pi}{2}$. The aim of this section is to show that f is the support function of catenoid. Recall that the principal radii of curvature of a smooth surface are the eigenvalues of the matrix $\nabla^2_{\mathbb{S}^{N-1}}H+\delta_{ij}H$, where the Hessian is taken with respect to the sphere $\mathbb{S}^{N-1}$ [Alexandroff 1939]. At each point where the Gauss curvature does not vanish, the zero mean curvature condition for $N=3$ can be written as

$$\triangle_{\mathbb{S}^2}H+2H=0,$$

where $\triangle_{\mathbb{S}^2}$ is the Laplace–Beltrami operator and $H(n)$ is the value of Minkowski's support function corresponding to the normal $n\in\mathbb{S}^2$. From now on let us consider the (x,y)-variables on $\mathbb{R}^2$. Recall that by rotating the graph of $y(x)=a\cosh(x/a)$ around the x-axis one obtains a catenoid for some constant a. Thus it is enough to compute the support function for the graph of y. Let α be the angle the tangent line of y at $(x,y(x))$ forms with the x-axis. If n is the unit normal to the graph of y then $n=(-\sin\alpha,\cos\alpha)$ and

$$H(n)=(x,y(x))\cdot n=-x\sin\alpha+a\cos\alpha\cosh\frac{x}{a}.$$

Noting that the unit tangent at $(x,y(x))$ is $(\cos\alpha,\sin\alpha)$ and equating with the slope of tangent line, which is $(\sinh(x/a),-1)$, we obtain

$$\cos\alpha=\frac{\sinh(x/a)}{\sqrt{1+\sinh^2(x/a)}},\qquad \sin\alpha=-\frac{1}{\sqrt{1+\sinh^2(x/a)}}.$$

From second equation we get that $\sinh(x/a)=\tan\alpha$ and solving the quadratic equation $e^{2(x/a)}-1=2e^{(x/a)}\tan\alpha$ we find that

$$x=a\log\frac{1+\sin\alpha}{\cos\alpha},\qquad \cosh\frac{x}{a}=\frac{1}{\cos\alpha}.$$

Consequently,

$$H(n) = -\frac{a}{2}\sin\alpha \log\left(\frac{1+\sin\alpha}{\cos\alpha}\right)^2 + a.$$

Taking $\alpha = \theta + \frac{\pi}{2}$ we have

$$\frac{1+\sin\alpha}{\cos\alpha} = \frac{1+\cos\theta}{-\sin\theta} = \frac{2\cos^2\frac{\theta}{2}}{-2\sin\frac{\theta}{2}\cos\frac{\theta}{2}} = -\cot\frac{\theta}{2}$$

and thus choosing $a = 2$ the result follows.

6C. ***Almost minimal immersions.*** Consider the parametrization $\mathcal{X} : U_g \to \mathbb{R}^3$, where

$$\mathcal{X}(n) = ng(n) + \nabla_{\mathbb{S}^2} g, \quad U_g = \{g > 0\} \subset \mathbb{S}^2. \tag{6-4}$$

Let $\mathcal{M}$ be the hypersurface determined by $\mathcal{X}$. The spherical part g of u_0 solves (6-1) and by [Reznikov 1992, Theorem 1] $\mathcal{X}$ determines a smooth map which is either constant or a conformal minimal immersion outside a locally finite set of isolated singularities (branch points). Recall that if at some point p

$$\mathcal{X}_{\xi_1} \times \mathcal{X}_{\xi_2} = 0, \quad \mathcal{X} = \mathcal{X}(\xi_1, \xi_2) \text{ in local coordinates } \xi_1, \xi_2, \tag{6-5}$$

then p is called branch point; see [Nitsche 1989, page 314].

Observe that $\mathcal{X}(n)$ is the gradient of the blow-up u_0 at $n = x/|x|$. Indeed,

$$\begin{aligned}\mathcal{X}(n) &= \frac{n}{r} rg + \frac{1}{r}\nabla_{\mathbb{S}^2}(rg) = \frac{n}{r}u_0(x) + \frac{1}{r}\nabla_{\mathbb{S}^2}u_0 \\ &= \frac{n}{r}(\nabla u_0(x)\cdot x) + \frac{1}{r}\nabla_{\mathbb{S}^2}u_0 = n\left(\nabla u_0(x)\cdot\frac{x}{|x|}\right) + \frac{1}{r}\nabla_{\mathbb{S}^2}u_0 = \nabla u_0(x).\end{aligned} \tag{6-6}$$

In particular, the computation above shows that

$$\nabla u_0(x) = \nabla u_0\left(\frac{x}{|x|}\right), \quad \nabla_{\mathbb{S}^2} g(n) \perp n; \tag{6-7}$$

in other words the gradient is homogeneous of degree 0.

The absence of branch points does not rule out the possibility of self-intersection. Therefore we need to prove that under conditions of Theorem C $\mathcal{M}$ is embedded.

6D. ***Dual cones and center of mass.*** If u_0 is a blow-up and the assumptions in Theorem C are satisfied, then by virtue of Proposition 5.15 the free boundary $\partial\{u_0 > 0\} \setminus \{0\}$ is a union of smooth convex cones $\mathcal{C}_1$ and $\mathcal{C}_2$. We define the dual cones as

$$\mathcal{C}_i^\star = \partial\{y \in \mathbb{R}^3 : x \cdot y \le 0,\ x \in \mathcal{C}_i\}, \quad i = 1, 2. \tag{6-8}$$

It is well known that the dual of a convex cone is also convex [Schneider 2014, page 35].

Lemma 6.2. *The largest principal curvature of $\mathcal{C}_i \setminus \{0\}$ is strictly positive.*

Proof. To fix the ideas, we prove the statement for $\mathcal{C}_1$. Note that one of the principal curvatures of $\mathcal{C}_1 \setminus \{0\}$ is zero because $\mathcal{C}_1$ is a cone and $\mathcal{C}_1 \setminus \{0\}$ is smooth; see Theorem B. Let $\kappa(p)$ be the largest principal curvature at $p \in \mathcal{C}_1 \setminus \{0\}$. Suppose there is p such that $\kappa(p) = 0$. Choose the coordinate system at p so that x_1 points in the outward normal direction at p (into $\{u_0 \equiv 0\}$), the x_2-axis is tangential at p and is the principal direction corresponding to $\kappa(p)$. Then we have $\nabla u_0(p) = e_1$, where e_1 is the unit direction of the x_1-axis, and the mean curvature of $\mathcal{C}_1$ at p vanishes because we assumed that $\kappa(p) = 0$. Writing the mean curvature at p in terms of the derivatives of u_0 we have

$$0 = \frac{\nabla u_0 D^2 u_0 (\nabla u_0)^T - |\nabla u_0|^2 \triangle u_0}{|\nabla u_0|^3} = \frac{\partial_{11} u_0}{\sqrt{2M}},$$

implying that $\partial_{11} u_0 = 0$. Moreover, since u_0 is homogeneous of degree 1, $\nabla u_0 = e_1$ along the x_1-axis. This yields $\partial_{13} u_0 = \partial_{23} u_0 = \partial_{33} u_0 = 0$ along the x_1-axis. From the harmonicity of u_0 it follows that $\partial_{22} u_0 = 0$ along the x_3-axis. Summarizing, we have that along the points of the x_3-axis the Hessian of u_0 has the form

$$\begin{pmatrix} 0 & \partial_{12} u_0 & 0 \\ \partial_{12} u_0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Finally, letting $\sigma(t)$, $t \in (-\delta, \delta)$, be the parametrization of the curve along which the $x_1 x_2$-plane intersects $\mathcal{C}_1$ and differentiating $|\nabla u_0(\sigma(t))| = 1$ in t we get that at p one must have

$$0 = e_1 \begin{pmatrix} 0 & \partial_{12} u_0 & 0 \\ \partial_{12} u_0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} e_2 = \partial_{12} u_0(p).$$

Thus, the Hessian $D^2 u_0$ vanishes along the x_1-axis. The function $w = \sqrt{2M} - \partial_1 u_0$ is harmonic in $\{u_0 > 0\}$ and $w \ge 0$ thanks to Lemma A.7. Moreover, $w(e_1) = 0 = \min w$. Since at e_1 the free boundary is regular, by Hopf's lemma $\partial_1 w = -\partial_{11} u_0 \ne 0$. However, $D^2 u_0(te_1) = 0$ for every $t > 0$ and hence $\partial_{11} w(e_1) = 0$, which is a contradiction. □

Remark 6.3. It follows from Lemma 6.2 and Theorem B that there are two positive constants κ_0, κ_1 such that

$$0 < \kappa_0 \le \kappa(p) \le \kappa_1, \quad p \in (\partial\{u_0 > 0\} \setminus \{0\}) \cap \partial B_{\sqrt{2M}},$$

where $\kappa(p)$ is the largest curvature of $\partial\{u_0 > 0\}$ at $p \in (\partial\{u_0 > 0\} \setminus \{0\}) \cap \partial B_{\sqrt{2M}}$.

Let us put $\gamma_i = \mathbb{S}^2 \cap \mathcal{C}_i^\star$.

Lemma 6.4. *Let $\mathcal{C}_1^\star, \mathcal{C}_2^\star$ be the dual cones (6-8). Then we have:*

(i) $\partial\mathcal{M}$ *is differentiable and there are two positive constants $\kappa_0^\star, \kappa_1^\star$ such that the largest curvature $\kappa^\star(p)$ of $(\mathcal{C}_i^\star \setminus \{0\}) \cap \mathbb{S}^2$ satisfies $\kappa_0^\star \le \kappa^\star(p) \le \kappa_1^\star$.*

(ii) *There is $\delta > 0$ small such that every component E_δ of $\partial B_{1-\delta} \cap \mathcal{M}$ defines a convex cone $K_\delta = \{\sigma t : \sigma \in E_\delta,\ t > 0\}$,*

(iii) $\mathcal{M}$ *is star-shaped with respect to the origin and hence embedded.*

Proof. Suppose that $\mathcal{C}_1^\star$ is not differentiable at some $z \neq 0$. Then $\mathcal{C}_i$ must have a flat piece. Indeed, if n_1, n_2 are the normals of two supporting hyperplanes of $\mathcal{C}_i^\star$ at z then the unit vectors

$$n_t = \frac{tn_1 + (1-t)n_2}{|tn_1 + (1-t)n_2|}$$

define a support function at z for every $t \in (0,1)$. Since the vectors n_t lie on the same plane, $\mathcal{C}_1$ must have a flat piece. The unique continuation theorem implies that the free boundary is a hyperplane and cannot have singularities. Now the desired estimate follows from Remark 6.3 and the definition of dual cone. The first claim is proved.

Let k_1, k_2 be the principal curvatures of $\mathcal{M}$. Then $k_1 + k_2 = 0$ and the Gauss curvature is $K = -k_1^2 = -k_2^2$. Since $\mathcal{M}$ is a smooth immersion, from (6-2) and the smoothness of $\mathcal{X} = \nabla u_0$ in U_g we see that $K \neq 0$. Furthermore, there is a tame constant $c_0 > 0$ such that $k_i^2 \geq c_0$, $i = 1, 2$, at every point of $\mathcal{M}$. Thus by virtue of the part (i) $\mathcal{M}$ is fibered by $\partial B_{1-\delta}$ for $\delta > 0$ small. We claim that $|\mathcal{X}(n)| > 0$, $n \in \overline{U}_g$. Clearly this is true if $n \in \partial U_g$, where $|\mathcal{X}(n)| = 1$. Suppose there is $n \in U_g$ such that $\mathcal{X}(n) = 0$. Since $\mathcal{X}(n) = ng + \nabla_{\mathbb{S}^2} g$, it follows that $g(n) = 0$, but this is impossible since $n \in \{g > 0\} = U_g$. From $g(n) = \mathcal{X}(n) \cdot n > 0$, $n \in U_g$, it follows that $\mathcal{M}$ is star-shaped with respect to the origin. Consequently, $\mathcal{M}$ is fibered by ∂B_t for every $t \in (0,1)$ and hence embedded. □

Let $n \in U_g$. Then from $\mathcal{X}(n) = \nabla u_0(n)$ it follows that

$$|\mathcal{X}(n)|_{\partial\{u_0>0\}}| = |\nabla u_0|_{\partial\{u_0>0\}}| = \sqrt{2M}.$$

Since by Lemma 6.4 $\mathcal{M}$ is differentiable along γ_i, we see that the contact angle α between $\mathcal{M}$ and $\mathbb{S}^2$ is

$$\cos\alpha = n \cdot \frac{\mathcal{X}(n)|_{\partial\{u_0>0\}}}{\sqrt{2M}} = g(n)|_{\partial\{u_0>0\}} = 0.$$

Thus, the minimal surface defined by g is inside of the sphere of radius $\sqrt{2M}$ because in view of Lemma A.7 $|\nabla u_0|^2 = g^2 + |\nabla g|^2 \leq 2M$. Moreover, $\mathcal{M}$ is tangential to $\mathcal{C}_1^\star$ and $\mathcal{C}_2^\star$ along $\mathbb{S}^2$ since $n \perp \nabla_{\mathbb{S}^2} g$ by (6-7).

We recall the definition of topological type $[\varepsilon, r, \chi]$ of hypersurface $\mathcal{M} \subset \mathbb{R}^3$ from [Nitsche 1985, page 47].

Definition 6.1. We say that $\mathcal{M}$ is of topological type $[\varepsilon, r, \chi]$ if it has orientation ε, Euler characteristic χ, and r boundary curves. Here $\varepsilon = \pm 1$, where $+1$ means that $\mathcal{M}$ is orientable and $\varepsilon = -1$ is nonorientable. For an orientable surface the Euler characteristic is defined by the relation $\chi = 2 - 2g - r$, where g is the genus of $\mathcal{M}$.

Now the first part of Theorem C follows from Nitsche's theorem [1985, page 2]. Moreover, the only stationary surfaces of disk type are the totally geodesic disks and the spherical cups. From Lemma 5.5 it follows that if $u_0 = rg$ and $\operatorname{supp} g$ is a disk then u_0 is a half-plane.

In view of Lemma 6.4(iii) the proof of Theorem C can be deduced from the result of [Nitsche 1962] but we will sketch a shorter proof based on Aleksandrov's moving plane method and Serrin's boundary lemma. We reformulate Theorem C as follows:

Lemma 6.5. *Let $\mathcal{M}$ be of topological type* $[1, 2, 0]$, *i.e., a ring-type minimal surface. Then $\mathcal{M}$ is a part of a catenoid.*

Proof. By Lemma 6.4(iii) $\mathcal{M}$ is embedded. In particular, $\mathcal{X}$ is a conformal minimal immersion (see the discussion in Section 6C).

Let $\partial\mathcal{M} = \gamma_1 \cup \gamma_2$. Then applying Stokes' formula we have

$$\int_{\mathcal{M}} \triangle_{\mathcal{M}} \mathcal{X} = \int_{\partial\mathcal{M}} n^* \, ds = \int_{\gamma_1} n^* + \int_{\gamma_2} n^* \, ds, \tag{6-9}$$

where n^* is the outward conormal, i.e., n^* is tangent to $\mathcal{M}$ and normal to $\partial\mathcal{M}$; see [Fang 1996, page 81]. Since $\mathcal{X}$ is minimal, $\triangle_{\mathcal{M}}\mathcal{X} = 0$. Thus

$$\int_{\gamma_1} n^* \, ds + \int_{\gamma_2} n^* \, ds = 0. \tag{6-10}$$

Since $\mathcal{M}$ is tangential to $\mathcal{C}_i^\star$ it follows that the conormal n^* on γ_i points in the direction of the generatrix of the dual cone $\mathcal{C}_i^\star$. Observe that if we use the arc-length parametrization of γ_i and let $s_k \in [0, |\gamma_i|]$ be some partition points then the sums $S_m = \sum_{k=0}^m n_k^{*i}(s_{k+1} - s_k)$, $n_k^{*i} \in \mathcal{C}_i^\star$, approximate the boundary integrals in (6-9). Consequently the vector S_m is strictly inside of the cone $\mathcal{C}_i^\star$ and in the limit converges to the center of mass of γ_i computed with respect to the origin (the vertex of the cone). In view of (6-10) there is a diameter of $\mathbb{S}^2$ strictly inside of both dual cones $\mathcal{C}_1^\star$ and $\mathcal{C}_2^\star$.

Without loss of generality we assume that the diameter passes through the north and south poles. Now we can apply Aleksandrov's moving plane method and Serrin's boundary point lemma to finish the proof. Let Π_t be the family of planes containing the x_1-axis where t measures the angle between Π_t and x_3-axis.

Now start rotating Π_t about the x_1-axis starting from a position when Π_t is a support hyperplane to either of the cones $\mathcal{C}_1^\star, \mathcal{C}_2^\star$ and $\Pi_t \cap \mathcal{C}_i^\star \neq \varnothing$, $i = 1, 2$.

<u>Case 1</u>: If the first touch of $\mathcal{M}$ and its reflection $\widetilde{\mathcal{M}}$ with respect to the plane Π_t occurs at some interior point of $\mathcal{M}$, then from the maximum principle it follows that $\mathcal{M} = \widetilde{\mathcal{M}}$.

By Lemma 6.4, both dual cones are strictly convex. Moreover, we claim that for δ small the cones generated by $\mathcal{M} \cap \partial B_{1-\delta}$ are convex, otherwise the inflection point would propagate to $\mathcal{C}_i^\star$.

The two remaining possibilities are:

<u>Case 2</u>: The first touch of $\mathcal{M}$ and its reflection $\widetilde{\mathcal{M}}$ occurs at some boundary point where $\partial\mathcal{M}$ is perpendicular to Π_t.

<u>Case 3</u>: The first touch of $\mathcal{M}$ and its reflection $\widetilde{\mathcal{M}}$ occurs at some boundary point where $\partial\mathcal{M}$ is not lying on Π_t.

We cannot directly apply Serrin's boundary point lemma [1971] because $\partial\mathcal{M}$ is only $C^{1,1}$ by virtue of Lemma 6.4. However, from the fibering of $\mathcal{M}$ near $\partial\mathcal{M}$ we conclude that $\tilde{g} \le g$ near the contact point, where $\tilde{g}$ is the support function of $\widetilde{M}$. Thus $\tilde{u} = r\tilde{g} \le rg = u$. Hence applying Serrin's boundary point lemma to the harmonic functions $\tilde{u}$ and u we conclude that $\mathcal{M} = \widetilde{\mathcal{M}}$.

Choosing Π_t to be an arbitrary family passing through a line perpendicular to the diameter it follows that γ_1, γ_2 are circles and (6-10) forces them to lie on parallel planes. Applying [Schoen 1983, Corollary 2] we infer that $\mathcal{M}$ is a part of catenoid. $\square$

Appendix

This section contains some well-known results about the solutions of the singular perturbation problem $(\mathcal{P}_\varepsilon)$. We begin with the uniform Lipschitz estimates of Luis Caffarelli; see [Caffarelli 1995] for the proof.

Proposition A.1. *Let $\{u_\varepsilon\}$ be a family of solutions of $(\mathcal{P}_\varepsilon)$. Then there is a constant C depending only on N, $\|\beta\|_\infty$ and independent of ε such that*

$$\|\nabla u_\varepsilon\|_{L^\infty(B_{1/2})} \le C. \tag{A-1}$$

As a consequence we get that one can extract converging sequences $\{u_{\varepsilon_n}\}$ of solutions of $(\mathcal{P}_\varepsilon)$ such that the limit functions are stationary points of the Alt–Caffarelli problem.

Proposition A.2. *Let u_ε be a family of solutions to $(\mathcal{P}_\varepsilon)$ in a domain $\mathcal{D} \subset \mathbb{R}^N$. Let us assume that $\|u_\varepsilon\|_{L^\infty(\mathcal{D})} \le A$ for some constant $A > 0$ independent of ε. For every $\varepsilon_n \to 0$ there exists a subsequence $\varepsilon_{n'} \to 0$ and $u \in C^{0,1}_{\text{loc}}(\mathcal{D})$ such that*

(i) *$u_{\varepsilon_{n'}} \to u$ uniformly on compact subsets of $\mathcal{D}$,*

(ii) *$\nabla u_{\varepsilon_{n'}} \to \nabla u$ in $L^2_{\text{loc}}(\mathcal{D})$,*

(iii) *u is harmonic in $\mathcal{D} \setminus \partial\{u > 0\}$.*

Proof. See [Caffarelli et al. 1997, Lemma 3.1]. $\square$

Next, we recall the estimates for the slopes of some global solutions.

Proposition A.3. *Let u be as in Proposition A.2. Then the following statements hold true*:

(i) *u is Lipschitz.*

(ii) *If $u_{\varepsilon_j} \to u = \alpha x_1^+$ locally uniformly, then $0 \le \alpha \le \sqrt{2M}$* (see [Caffarelli et al. 1997, Proposition 5.2]).

(iii) *If $u_{\varepsilon_j} \to u = \alpha x_1^+ - \gamma x_1^- + o(|x|)$ and $\gamma > 0$ then $\alpha^2 - \gamma^2 = \sqrt{2M}$* (see [Caffarelli et al. 1997, Proposition 5.1]). *In this lemma the essential assumption is that $\gamma > 0$.*

Remark A.4. Observe that if $u(x) = \alpha x_1^+ + \bar{\alpha} x_1^-$ then we must necessarily have that $\alpha = \bar{\alpha} \le \sqrt{2M}$; see [Caffarelli et al. 1997, Proposition 5.3]. In this case the interior of the zero set of u is empty. Thus one might have a wedge-like solution.

Using Proposition A.1 we can extract a sequence u_{ε_j} for some sequence ε_j such that $u_{\varepsilon_j} \to u$ uniformly in $B_{1/2}$; see Proposition A.2. Let u be a limit and $0 < \rho_j \downarrow 0$ and $u_j(x) = u(x_0 + \rho_j x)/\rho_j$, $x_0 \in \partial\{u > 0\}$. Thanks to Proposition A.3(i) we can extract a subsequence, still labeled ρ_j, such that u_j converges to some function u_0 defined in $\mathbb{R}^N$. The function u_0 is called a blow-up limit of u at the free boundary point x_0 and it depends on $\{\rho_j\}$.

The two propositions to follow establish an important property of the blow-up limits, namely that the first and second blow-ups of u can be obtained from $(\mathcal{P}_\varepsilon)$ for a suitable choice of parameter ε. Observe that the scaled function $\nabla(u_{\varepsilon_j})_{\lambda_n}$ satisfies the equation

$$\triangle(u_{\varepsilon_j})_{\lambda_j} = \frac{\lambda_j}{\varepsilon_j}\beta\left(\frac{\lambda_j}{\varepsilon_j}(u_{\varepsilon_j})_{\lambda_j}\right). \tag{A-2}$$

Taking $\delta_j = \varepsilon_j/\lambda_j \to 0$ we see that $(u_{\varepsilon_j})_{\lambda_j}$ is solution to $\triangle u_{\delta_j} = \beta_{\delta_j}(u_{\delta_j})$.

Proposition A.5. *Let u_{ε_j} be a family of solutions to $(\mathcal{P}_\varepsilon)$ in a domain $\mathcal{D} \subset \mathbb{R}^N$ such that $u_{\varepsilon_j} \to u$ uniformly on $\mathcal{D}$ and $\varepsilon_j \to 0$. Let $x_0 \in \mathcal{D} \cap \partial\{u > 0\}$ and let $x_n \in \partial\{u > 0\}$ be such that $x_n \to x_0$ as $n \to \infty$. Let $\lambda_n \to 0$, $u_{\lambda_n}(x) = (1/\lambda_n)u(x_n + \lambda_n x)$ and $(u_{\varepsilon_j})_{\lambda_n} = (1/\lambda_n)u_{\varepsilon_j}(x_n + \lambda_n x)$. Assume that $u_{\lambda_n} \to U$ as $n \to \infty$ uniformly on compact subsets of $\mathbb{R}^N$. Then there exists $j(n) \to \infty$ such that for every $j_n \geq j(n)$ it holds that $\varepsilon_j/\lambda_n \to 0$ and*

- *$(u_{\varepsilon_{j_n}})_{\lambda_n} \to U$ uniformly on compact subsets of $\mathbb{R}^N$,*
- *$\nabla(u_{\varepsilon_{j_n}})_{\lambda_n} \to \nabla U$ in $L^2_{\text{loc}}(\mathbb{R}^N)$,*
- *$\nabla u_{\lambda_n} \to \nabla U$ in $L^2_{\text{loc}}(\mathbb{R}^N)$.*

Proof. See [Caffarelli et al. 1997, Lemma 3.2]. □

Finally, recall that the result of the previous proposition extends to the second blow-up.

Proposition A.6. *Let u_{ε_j} be a solution to $(\mathcal{P}_\varepsilon)$ in a domain $\mathcal{D}_j \subset \mathcal{D}_{j+1}$ and $\bigcup_j \mathcal{D}_j = \mathbb{R}^N$ such that $u_{\varepsilon_j} \to U$ uniformly on compact sets of $\mathbb{R}^N$ and $\varepsilon_j \to 0$. Let us assume that for some choice of positive numbers d_n and points $x_n \in \partial\{U > 0\}$, the sequence*

$$U_{d_n}(x) = \frac{1}{d_n}U(x_n + d_n x)$$

converges uniformly on compact sets of $\mathbb{R}^N$ to a function U_0. Let

$$(u_{\varepsilon_j})_{d_n} = \frac{1}{d_n}u_{\varepsilon_j}(x_n + d_n x).$$

Then there exists $j(n) \to \infty$ such that for every $j_n \geq j(n)$, it holds that $\varepsilon_{j_n}/d_n \to 0$ and

- *$(u_{\varepsilon_{j_n}})_{d_n} \to U_0$ uniformly on compact subsets of $\mathbb{R}^N$,*
- *$\nabla(u_{\varepsilon_j})_{d_n} \to \nabla U_0$ in $L^2_{\text{loc}}(\mathbb{R}^N)$.*

Proof. See [Caffarelli et al. 1997, Lemma 3.3]. □

The next lemma contains one of the crucial estimates needed for the proof of Proposition 5.15.

Lemma A.7. *Let $u \geq 0$ be as in Proposition A.2. Then*

$$\limsup_{x \to x_0, u(x) > 0} |\nabla u(x)| \leq \sqrt{2M}.$$

Proof. To fix the ideas we let $x_0 = 0$ and $l = \limsup_{x\to 0, u(x)>0} |\nabla u(x)|$. Suppose $l > 0$, otherwise we are done. Choose a sequence $z_k \to 0$ such that $u(z_k) > 0$ and $|\nabla u(z_k)| \to l$. Setting $\rho_k = |y_k - z_k|$, where $y_k \in \partial\{u > 0\}$ is the nearest point to z_k on the free boundary and proceeding as in the proof of [Alt et al. 1984a, Lemma 3.4] we can conclude that the blow-up sequence $u_k(x) = \rho_k^{-1} u(z_k + \rho_k x)$ has a limit u_0 (at least for a subsequence, thanks to Proposition A.1) such that $u_0(x) = l x_1,\ x_1 > 0$, in a suitable coordinate system. Moreover, by Proposition A.5 it follows that u_0 is a limit of some u_{λ_j} solving $\triangle u_{\lambda_j} = \beta_{\lambda_j}(u_{\lambda_j})$ in $B_{r_j}, r_j \to \infty$. If there is a point $z \in \{x_1 = 0\}$ and $r > 0$ such that $u_0 > 0$ in $B_r(z) \cap \{x_1 < 0\}$ then near z we must have $u_0(x) = l(x-z)_1^+ + l(x-z)_1^- + o(x-z)$; see Remark A.4. Applying the unique continuation theorem to $u_0(x) - u_0(-x_1, x_2, \dots, x_n)$ we see that $u_0 = l(-x_1)^+$, $x_1 < 0$. Thus recalling Remark A.4 again we infer that $l \le \sqrt{2M}$. □

Finally, we mention a useful identity for the solutions u_ε; see [Caffarelli et al. 1997, equation (5.2)]: Let u_ε be a solution of $(\mathcal{P}_\varepsilon)$. Then for any $\phi \in C_0^\infty(B_1)$ there holds

$$\int \left(\frac{|\nabla u_\varepsilon|^2}{2} + \mathcal{B}\left(\frac{u_\varepsilon}{\varepsilon}\right)\right) \partial_1 \phi = \int \sum_k \partial_k u_\varepsilon \, \partial_1 u_\varepsilon \partial_k \phi. \tag{A-3}$$

References

[Alexandroff 1939] A. Alexandroff, "Über die Oberflächenfunktion eines konvexen Körpers", *Rec. Math. N.S.* [*Mat. Sbornik*] **6** (1939), 167–174. In Russian. MR Zbl

[Alt and Caffarelli 1981] H. W. Alt and L. A. Caffarelli, "Existence and regularity for a minimum problem with free boundary", *J. Reine Angew. Math.* **325** (1981), 105–144. MR Zbl

[Alt et al. 1984a] H. W. Alt, L. A. Caffarelli, and A. Friedman, "A free boundary problem for quasilinear elliptic equations", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **11**:1 (1984), 1–44. MR Zbl

[Alt et al. 1984b] H. W. Alt, L. A. Caffarelli, and A. Friedman, "Variational problems with two phases and their free boundaries", *Trans. Amer. Math. Soc.* **282**:2 (1984), 431–461. MR Zbl

[Caffarelli 1987] L. A. Caffarelli, "A Harnack inequality approach to the regularity of free boundaries, I: Lipschitz free boundaries are $C^{1,\alpha}$", *Rev. Mat. Iberoamericana* **3**:2 (1987), 139–162. MR Zbl

[Caffarelli 1989] L. A. Caffarelli, "A Harnack inequality approach to the regularity of free boundaries, II: Flat free boundaries are Lipschitz", *Comm. Pure Appl. Math.* **42**:1 (1989), 55–78. MR Zbl

[Caffarelli 1995] L. A. Caffarelli, "Uniform Lipschitz regularity of a singular perturbation problem", *Differential Integral Equations* **8**:7 (1995), 1585–1590. MR Zbl

[Caffarelli and Salsa 2005] L. Caffarelli and S. Salsa, *A geometric approach to free boundary problems*, Graduate Studies in Math. **68**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[Caffarelli et al. 1997] L. A. Caffarelli, C. Lederman, and N. Wolanski, "Uniform estimates and limits for a two phase parabolic singular perturbation problem", *Indiana Univ. Math. J.* **46**:2 (1997), 453–489. MR Zbl

[Caffarelli et al. 2000] L. A. Caffarelli, L. Karp, and H. Shahgholian, "Regularity of a free boundary with application to the Pompeiu problem", *Ann. of Math.* (2) **151**:1 (2000), 269–292. MR Zbl

[Caffarelli et al. 2004] L. A. Caffarelli, D. Jerison, and C. E. Kenig, "Global energy minimizers for free boundary problems and full regularity in three dimensions", pp. 83–97 in *Noncompact problems at the intersection of geometry, analysis, and topology* (New Brunswick, NJ, 2001), edited by A. Bahri et al., Contemp. Math. **350**, Amer. Math. Soc., Providence, RI, 2004. MR Zbl

[Fang 1996] Y. Fang, *Lectures on minimal surfaces in* $\mathbb{R}^3$, Proc. Centre Math. Appl. Austr. Nat. Univ. **35**, Aust. Nat. Univ. Centre Math. Appl., Canberra, 1996. MR Zbl

[Hauswirth et al. 2011] L. Hauswirth, F. Hélein, and F. Pacard, "On an overdetermined elliptic problem", *Pacific J. Math.* **250**:2 (2011), 319–334. MR Zbl

[Jerison and Perera 2018] D. Jerison and K. Perera, "Higher critical points in an elliptic free boundary problem", *J. Geom. Anal.* **28**:2 (2018), 1258–1294. MR Zbl

[Jerison and Savin 2015] D. Jerison and O. Savin, "Some remarks on stability of cones for the one-phase free boundary problem", *Geom. Funct. Anal.* **25**:4 (2015), 1240–1257. MR Zbl

[Khavinson et al. 2013] D. Khavinson, E. Lundberg, and R. Teodorescu, "An overdetermined problem in potential theory", *Pacific J. Math.* **265**:1 (2013), 85–111. MR Zbl

[Langevin and Rosenberg 1988] R. Langevin and H. Rosenberg, "A maximum principle at infinity for minimal surfaces and applications", *Duke Math. J.* **57**:3 (1988), 819–828. MR Zbl

[Nitsche 1962] J. C. C. Nitsche, "A characterization of the catenoid", *J. Math. Mech.* **11** (1962), 293–301. MR Zbl

[Nitsche 1985] J. C. C. Nitsche, "Stationary partitioning of convex bodies", *Arch. Rational Mech. Anal.* **89**:1 (1985), 1–19. MR Zbl

[Nitsche 1989] J. C. C. Nitsche, *Lectures on minimal surfaces, I: Introduction, fundamentals, geometry and basic boundary value problems*, Cambridge Univ. Press, 1989. MR Zbl

[Reznikov 1992] A. G. Reznikov, "Linearization and explicit solutions of the minimal surface equation", *Publ. Mat.* **36**:1 (1992), 39–46. MR Zbl

[Ros and Vergasta 1995] A. Ros and E. Vergasta, "Stability for hypersurfaces of constant mean curvature with free boundary", *Geom. Dedicata* **56**:1 (1995), 19–33. MR Zbl

[Schneider 2014] R. Schneider, *Convex bodies: the Brunn–Minkowski theory*, 2nd expanded ed., Encyc. Math. Appl. **151**, Cambridge Univ. Press, 2014. MR Zbl

[Schoen 1983] R. M. Schoen, "Uniqueness, symmetry, and embeddedness of minimal surfaces", *J. Differential Geom.* **18**:4 (1983), 791–809. MR Zbl

[Serrin 1971] J. Serrin, "A symmetry problem in potential theory", *Arch. Rational Mech. Anal.* **43** (1971), 304–318. MR Zbl

[Spruck 1983] J. Spruck, "Uniqueness in a diffusion model of population biology", *Comm. Partial Differential Equations* **8**:15 (1983), 1605–1620. MR Zbl

[Traizet 2014] M. Traizet, "Classification of the solutions to an overdetermined elliptic problem in the plane", *Geom. Funct. Anal.* **24**:2 (2014), 690–720. MR Zbl

[Weiss 2003] G. S. Weiss, "A singular limit arising in combustion theory: fine properties of the free boundary", *Calc. Var. Partial Differential Equations* **17**:3 (2003), 311–340. MR Zbl

Received 14 Feb 2018. Revised 5 Sep 2018. Accepted 19 Dec 2018.

ARAM L. KARAKHANYAN: aram6k@gmail.com
School of Mathematics, The University of Edinburgh, Edinburgh, United Kingdom

ANALYSIS AND PDE
Vol. 13, No. 1, 2020
dx.doi.org/10.2140/apde.2020.13.201

A SPIRAL INTERFACE WITH POSITIVE ALT–CAFFARELLI–FRIEDMAN LIMIT AT THE ORIGIN

MARK ALLEN AND DENNIS KRIVENTSOV

We give an example of a pair of nonnegative subharmonic functions with disjoint support for which the Alt–Caffarelli–Friedman monotonicity formula has strictly positive limit at the origin, and yet the interface between their supports lacks a (unique) tangent there. This clarifies a remark of Caffarelli and Salsa (*A geometric approach to free boundary problems*, 2005) that the positivity of the limit of the ACF formula implies unique tangents; this is true under some additional assumptions, but false in general. In our example, blow-ups converge to the expected piecewise linear two-plane function along subsequences, but the limiting function depends on the subsequence due to the spiraling nature of the interface.

1. Introduction

The Alt–Caffarelli–Friedman monotonicity formula (hereafter denoted ACF formula) has been and continues to be a powerful tool in the study of free boundary problems. It was introduced in [Alt et al. 1984] in order to prove that the solutions to a two-phase Bernoulli free boundary problem are Lipschitz continuous. The formula was then adapted to treat more general two-phase problems, and a discussion of the formula, its proof, and its applications to two-phase free boundary problems may be found in [Caffarelli and Salsa 2005]. The ACF formula has also been effective in studying obstacle-type problems, and applications of the formula for obstacle-type problems are found in [Petrosyan et al. 2012]. Further applications also include the study of segregation problems in [Caffarelli et al. 2009]. While the most typical use of the formula is to prove the optimal regularity of solutions or flatness of the free boundary, it can also be used for other purposes, such as to show the separation of phases in free boundary problems; see [Allen and Petrosyan 2012; Allen et al. 2015; Allen and Shi 2016].

The key property of the ACF formula (1-1) is given in the following proposition:

Proposition 1.1. *Let* $u_1, u_2 \geq 0$ *be two continuous subharmonic functions in* B_R *with* $u_1 \cdot u_2 = 0$ *and* $u_1(0) = u_2(0) = 0$. *Then*

$$\Phi(r, u_1, u_2) := \frac{1}{r^4} \int_{B_r(0)} \frac{|\nabla u_1|^2}{|x|^{n-2}} \int_{B_r(0)} \frac{|\nabla u_2|^2}{|x|^{n-2}} \tag{1-1}$$

is nondecreasing for $0 < r < R$. *Consequently, the limit*

$$\Phi(0+, u_1, u_2) := \lim_{r \searrow 0} \Phi(r, u_1, u_2)$$

is well defined.

MSC2010: 35R35, 35J05.
Keywords: ACF monotonicity formula, spiral interface, free boundary, monotonicity formula.

Our paper is motivated by the following claim, which appears as Lemma 12.9 in [Caffarelli and Salsa 2005]:

Claim 1.2. *Let $u \geq 0$ be continuous in B_1 and harmonic in $\{u > 0\}$. Let Ω_1 be a connected component of $\{u > 0\}$ and let $0 \in \partial\Omega_1$. If $u_1 = u|_{\Omega_1}$ and $u_2 = u - u_1$, then if $\Phi(0+, u_1, u_2) > 0$, exactly two connected components Ω_1 and Ω_2 of $\{u > 0\}$ are tangent at 0, and in a suitable system of coordinates,*

$$u(x) = \alpha x_1^+ + \beta x_1^- + o(|x|), \tag{1-2}$$

with $\alpha, \beta > 0$.

As no proof of this Lemma 12.9 is provided in [Caffarelli and Salsa 2005] (it is followed only by some general remarks), it is not entirely clear whether it is meant to be taken at face value. We note, for example, that if u is also assumed to satisfy a two-phase free boundary problem of the type treated in [Caffarelli and Salsa 2005], then the claim is valid, but requires heavy use of the free boundary relation to prove.

Claim 1.2, and in particular the question of whether it is true in the generality stated, drew the authors' interest when the second author was tempted to use it while working on certain eigenvalue optimization problems [Kriventsov and Lin 2019] but was unable to write down a proof. Typically, a monotonicity formula is applied together with other tools making explicit use of the free boundary relation in order to prove regularity of an interface; however, Claim 1.2 would imply that the ACF monotonicity formula, on its own, yields some regularity of the interface. This makes the claim very powerful and useful, especially in problems where the free boundary condition is difficult to exploit, such as the vector-valued free boundary problems arising from spectral optimization [Kriventsov and Lin 2018; 2019].

Unfortunately, it is also not true: the main result of this paper is to provide a counterexample to Claim 1.2.

Theorem 1.3. *For any dimension $n \geq 2$, there exist two continuous subharmonic functions $u, \tilde{u} \geq 0$ with $u, \tilde{u}$ both harmonic in their respective positivity sets and $u \cdot \tilde{u} = 0$. Furthermore, $\Phi(0+, u, \tilde{u}) > 0$. However, $\partial\{u > 0\}$ and $\partial\{\tilde{u} > 0\}$ (which are given by a piecewise smooth, connected hypersurface when restricted to any annulus $B_1 \backslash B_r$) do not admit tangents (or approximate tangents) at the origin, nor do there exist numbers $\alpha, \beta > 0$ and a change of coordinates such that $u + \tilde{u} = \alpha x_1^+ + \beta x_1^- + o(|x|)$.*

In the above, the boundary of a measurable set A is said to *admit a tangent* (*plane*) at the origin if

$$0 < \liminf_{r \searrow 0} \frac{|B_r \cap A|}{|B_r|} \leq \limsup_{r \searrow 0} \frac{|B_r \cap A|}{|B_r|} < 1 \tag{1-3}$$

and there is a unit vector ν such that

$$\lim_{r \searrow 0} \frac{1}{r} \max_{x \in \partial A \cap B_r} |x \cdot \nu| = 0.$$

The boundary is said to admit an *approximate tangent* (*plane*) if (1-3) holds and

$$\lim_{r \searrow 0} \frac{1}{r^{n+1}} \int_{B_r \cap \partial A} |x \cdot \nu|^2 \, d\mathcal{H}^{n-1} = 0.$$

Here $\mathcal{H}^{n-1}$ denotes $(n-1)$-dimensional Hausdorff measure. Note that if $u, \tilde{u}$ are as in Claim 1.2 and A is either $\{u > 0\}$ or $\{\tilde{u} > 0\}$, then (1-3) holds; see Corollary 12.4 in [Caffarelli and Salsa 2005].

It seems that the notion of approximate tangent above (or another similar measure-theoretic notion) is the more meaningful one in this context. Indeed, there are simpler constructions which produce functions $u, \tilde{u}$ as in Claim 1.2 for which $\partial\{u>0\}$ does not admit a tangent at 0 but does admit an approximate tangent.

If one only considers functions u for which $\partial\{u>0\}$ is, say, given by a 1-Lipschitz graph over some plane π_r on every annulus $B_{2r}\setminus B_r$, these two notions of tangent plane are equivalent. This property holds for the example constructed in the proof of Theorem 1.3.

The functions $u, \tilde{u}$ we construct in proving the theorem have $\partial\{u>0\}$ a spiral: while $u+\tilde{u}$ looks more and more like $\alpha(x\cdot\nu)_+ + \beta(x\cdot\nu)_-$ on progressively smaller balls B_r, the choice of ν cannot be made uniformly in r, and the optimal ν rotates (slowly) as r decreases. Some free boundary problems are known to exhibit spiraling patterns for the interface; see [Blank 2001; Terracini et al. 2019] for examples, although the spirals produced there have different properties from ours. We also remark that an example of nonunique tangents for an energy minimization problem is given in [White 1992].

Further questions. Before turning to the proof of Theorem 1.3 we would like to offer some discussion of the further questions raised by this theorem and speculate on what "optimal" results, both positive and negative, might look like.

A standard argument with the ACF formula shows that if $u, \tilde{u}$ are as in Claim 1.2, then for every sequence $r_k \to 0$, there is a subsequence r_{k_j} such that

$$\lim_{j\to\infty} \frac{1}{r_{k_j}^{n+2}} \int_{B_{r_{k_j}}} |u(x) - \alpha(x\cdot\nu)_+ - \beta(x\cdot\nu)_-|^2 = 0,$$

where α, β, ν depend on the subsequence. Let us refer to any such subsequence r_{k_j} as a *blow-up subsequence*. We are interested in whether or not these parameters may be chosen independent of the blow-up subsequence.

In the example constructed below, the functions u and $\tilde{u}$ are rotations of one another around the origin; in particular, this means that for all of the blow-up subsequences, $\alpha = \beta = c\sqrt{\Phi(0+, u, \tilde{u})}$ are the same, while ν depends on the particular subsequence.

This example gives one way for (1-2) to fail. There could, in principle, be another way: say that $\partial\{u>0\} = \partial\{\tilde{u}>0\}$ is given by a C^1 hypersurface (including up to the origin, so that it admits a tangent there), and that $u, \tilde{u}$ are as in Claim 1.2. Can one find a pair $u, \tilde{u}$ like this for which (1-2) fails? This would mean that between the various blow-up subsequences, ν would remain fixed, while α and β would vary. Note that if the hypersurface is more regular near the origin (in particular, if it is a Lyapunov–Dini surface), then this is impossible.

Another set of questions is related to optimality in Theorem 1.3. To clarify the discussion, define, for each r, $\nu(r)$ to be the best approximating normal vector:

$$\int_{B_r\cap\partial\{u>0\}} |x\cdot\nu(r)|^2 \, d\mathcal{H}^{n-1} = \min_{\nu\in S^{n-1}} \int_{B_r\cap\partial\{u>0\}} |x\cdot\nu|^2 \, d\mathcal{H}^{n-1}.$$

It may be verified that $\nu(r)$ is uniquely determined from this relation and depends in a Lipschitz manner on r. The property of having an approximate tangent, then, can be reformulated as saying that $\nu(r)$ has a limit as $r \to 0$, while Theorem 1.3 gives an example where

$$\int_0^1 \left| \frac{d\nu(r)}{dr} \right| = \infty. \tag{1-4}$$

What restrictions on the change in $\nu(r)$, one may ask then, are implied by the conditions in Claim 1.2? We conjecture that under those conditions, one must have

$$\int_0^1 r \left| \frac{d\nu(r)}{dr} \right|^2 < \infty; \tag{1-5}$$

on the other hand, for any $\nu_0(r)$ satisfying (1-4) and (1-5), there is a pair of functions $u, \tilde{u}$ as in Claim 1.2 with $\nu_0(r)$ with

$$\left| \frac{d\nu(r)}{dr} \right| \geq \left| \frac{d\nu_0(r)}{dr} \right|.$$

To explain the source of (1-5), let us point out that in Section 2, we construct a pair of functions $u, \tilde{u}$ for which

$$\int_0^\infty \left| \frac{d\nu(r)}{dr} \right| = \theta \quad \text{and} \quad \frac{\Phi(0+, u, \tilde{u})}{\Phi(\infty, u, \tilde{u})} \geq 1 - \theta^2$$

(and this dependence on θ seems to be sharp up to constants). By gluing truncated and scaled versions of this construction, one might hope to attain functions $u, \tilde{u}$ satisfying the hypotheses of Claim 1.2, and with

$$\int_{2^j}^{2^{j+1}} \left| \frac{d\nu(r)}{dr} \right| \approx \theta_i$$

for any sequence θ_i for which $\prod_i (1-\theta_i^2) > 0$. This restriction is equivalent to (1-5) for such a construction. In the actual proof of Theorem 1.3, we are unable to perform the truncation and gluing steps uniformly in θ, and so do not obtain such a quantitative result.

Finally, over the past two decades enormous progress has been made in understanding the relationship between the behavior of positive harmonic functions with zero Dirichlet condition near the boundaries of domains and the geometric measure-theoretic properties of the boundary; we do not attempt to provide a summary here, but refer the reader to the introduction and references in [Azzam et al. 2016]. We suggest that the questions above can be thought of as a continuation, or extension, of this program, with the goal of relating (finer) geometric properties of a boundary to the simultaneous behavior of positive harmonic functions on a domain and its compliment, using the ACF formula as a crucial tool.

Outline of proof. To prove Theorem 1.3 we will construct a subharmonic function $u \geq 0$ in $\mathbb{R}^2$ such that u is harmonic in its positivity set and $u(0) = 0$. Furthermore, $\partial\{u > 0\}$ will be invariant under a rotation of π. Consequently, if $\tilde{u}(z) := u(-z)$, then the pair $u, \tilde{u}$ will satisfy the assumptions of the ACF formula in Proposition 1.1. Before explaining the construction of u and the outline of the paper, we first give two definitions.

We define the class of functions in $\mathbb{R}^2$

$$\mathcal{K} := \big\{u \in C(B_1) : u \geq 0 \text{ in } B_1,\ \Delta u = 0 \text{ in } \{u>0\},\\ u(0) = 0,\ u(z)\cdot u(-z) = 0,\ \text{and } \partial\{u(z)>0\} = \partial\{u(-z)>0\}\big\}.$$

By working in the class $\mathcal{K}$, we may consider using a one-sided rescaled version of the ACF formula. If $u \in \mathcal{K}$, then

$$J(r,u) := \left(\frac{2}{\pi r^2}\int_{B_r} |\nabla u|^2\right)^{1/2}$$

is monotonically nondecreasing in r since $J(r,u) = \left(\frac{2}{\pi}\right)^2\sqrt{\Phi(r, u(z), u(-z))}$. Furthermore, if u is C^1 up to $\partial\{u>0\}$ near the origin, then $J(0+,u) = |\nabla u(0)|$.

In order to prove Theorem 1.3 we first show in Section 2, working on unbounded domains, that it is possible to turn $\partial\{u>0\}$ so that its asymptotic behavior at infinity differs from its tangent near the origin by an angle of θ while arranging that $J(\infty,u) - J(0+,u) < 1-\theta^2$ (for small θ). In Section 3 we transfer this result to a bounded domain. In Section 4 we inductively construct a sequence of functions in $\mathcal{K}$ and take a limit to obtain the u in Theorem 1.3. Heuristically, the value of $J(0+,u)$ should be $\prod(1-\theta_i^2)$, and this is strictly positive if, say, $\theta_i = i^{-1}$. On successively smaller balls, the interface $\{u=0\}$ will have turned a total amount of $\sum i^{-1} \to \infty$, which implies that the interface spirals towards the origin and therefore lacks a unique tangent there. We make these heuristic ideas rigorous, and then we show how the pair $u, \tilde{u}$ also provide a counterexample in higher dimensions.

2. Conformal mapping

We utilize the Schwarz–Christoffel formula to obtain a conformal mapping. For a fixed angle $0 < \theta < \frac{\pi}{2}$, we map the upper half-plane to the domain Ω_θ (see Figure 1) by the conformal mapping f_θ with derivative

$$f_\theta'(z) = (z-(-1))^{(\pi+\theta)/\pi-1}(z-1)^{(\pi-\theta)/\pi-1} = \left(\frac{z+1}{z-1}\right)^{\theta/\pi}. \tag{2-1}$$

We translate f_θ by a constant z_0, so that the midpoint of the line segment in the image is the origin $0+0i$. We define $t_\theta \in (-1,1) \subset \mathbb{R}$ to be $t_\theta = f_\theta^{-1}(0+0i)$. Clearly, $t_\theta \to 0$ as $\theta \to 0$. What is of importance is how quickly $t_\theta \to 0$. In order to determine this decay rate we use the following result:

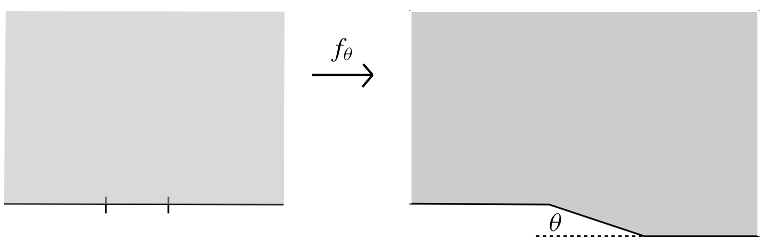

Figure 1. Conformal map.

Lemma 2.1. *Let $f, g > 0$ be integrable functions on an interval I. If f/g is an increasing function, then for any $x_1 < x_2 < x_3 < x_4$ with each $x_i \in I$, we have*

$$\frac{\int_{x_1}^{x_2} f}{\int_{x_1}^{x_2} g} \le \frac{\int_{x_3}^{x_4} f}{\int_{x_3}^{x_4} g}.$$

Proof. Since f/g is increasing we have

$$\int_{x_1}^{x_2} f(x)\,dx \le \int_{x_1}^{x_2} \frac{f(x_2)}{g(x_2)} g(x)\,dx.$$

Consequently, we have

$$\frac{\int_{x_1}^{x_2} f(x)\,dx}{\int_{x_1}^{x_2} g(x)\,dx} \le \frac{f(x_2)}{g(x_2)}.$$

By the same argument, we have

$$\frac{f(x_3)}{g(x_3)} \le \frac{\int_{x_3}^{x_4} f(x)\,dx}{\int_{x_3}^{x_4} g(x)\,dx},$$

and so the conclusion follows. □

We will also need the following:

Lemma 2.2. *Let $f \ge g > 0$ be integrable and continuous on $[0, 1)$ with $f \ge g$ and f/g increasing, and*

$$\int_0^1 f > M \quad \text{and} \quad \int_0^1 g > M.$$

Let x_1, x_2 be the unique values such that

$$M + \int_0^{x_1} g = \int_{x_1}^1 g \quad \text{and} \quad M + \int_0^{x_2} f = \int_{x_2}^1 f. \tag{2-2}$$

Then $x_1 \le x_2$.

Proof. We have

$$\frac{M + \int_0^{x_1} f}{M + \int_0^{x_1} g} \le \frac{\int_0^{x_1} f}{\int_0^{x_1} g} \le \frac{\int_{x_1}^1 f}{\int_{x_1}^1 g},$$

where the second inequality is due to Lemma 2.1. Since x_1 is chosen so that (2-2) holds, we have that the denominator in the inequality above is the same so that

$$M + \int_0^{x_1} f \le \int_{x_1}^1 f.$$

Then $x_1 \le x_2$. □

The two lemmas above allow us to prove:

Lemma 2.3. *Let f_θ be defined as in (2-1) and let $t_\theta = f_\theta^{-1}(0 + 0i)$. Then there exists $\theta_0 > 0$ such that $0 < t_\theta \le 2\theta/\pi$ as long as $0 < \theta \le \theta_0$.*

Proof. To determine the midpoint of a line segment it suffices to find the x-value. Consequently, we focus on the real part of the mapping f_θ. If $t \in (-1, 1)$, then

$$f'(t) = \left((-1)\frac{1+t}{1-t}\right)^{\theta/\pi} = \left(\frac{1+t}{1-t}\right)^{\theta/\pi} e^{i\theta}.$$

Thus, t_θ is the unique value in $(-1, 1)$ such that

$$\int_{-1}^{t_\theta} \left(\frac{1+t}{1-t}\right)^{\theta/\pi} dt = \int_{t_\theta}^{1} \left(\frac{1+t}{1-t}\right)^{\theta/\pi} dt.$$

We now note that

$$\left(\frac{1+t}{1-t}\right)^{\theta/\pi} \ge \left(\frac{1+t}{2}\right)^{\theta/\pi} \quad \text{if } -1 \le t \le 0.$$

Then $t_\theta \le \xi_\theta$, where ξ_θ is the unique value such that

$$\int_{-1}^{0} \left(\frac{1+t}{2}\right)^{\theta/\pi} dt + \int_{0}^{\xi_\theta} \left(\frac{1+t}{1-t}\right)^{\theta/\pi} dt = \int_{\xi_\theta}^{1} \left(\frac{1+t}{1-t}\right)^{\theta/\pi} dt.$$

We also have

$$\left(\frac{1+t}{1-t}\right)^{\theta/\pi} \le \left(\frac{1}{1-t}\right)^{2\theta/\pi} \quad \text{if } 0 \le t \le 1,$$

and

$$\frac{(1/(1-t))^{2\theta/\pi}}{((1+t)/(1-t))^{\theta/\pi}} = \left(\frac{1}{1-t^2}\right)^{\theta/\pi}$$

is an increasing function on $(0, 1)$. If we let

$$M = \int_{-1}^{0} \left(\frac{1+t}{2}\right)^{\theta/\pi} dt,$$

then we may apply Lemma 2.2 and conclude that $t_\theta \le \xi_\theta \le \tau_\theta$, where τ_θ is given by

$$\int_{-1}^{0} \left(\frac{1+t}{2}\right)^{\theta/\pi} dt + \int_{0}^{\tau_\theta} \left(\frac{1}{1-t}\right)^{2\theta/\pi} dt = \int_{\tau_\theta}^{1} \left(\frac{1}{1-t}\right)^{2\theta/\pi} dt.$$

The integrals above have elementary antiderivatives. In order to show that $\tau_\theta \le 2\theta/\pi$ for small θ, we choose $2\theta/\pi$ as the point of integration. By taking explicit antiderivatives and simplifying, it suffices to show that for small enough θ,

$$\frac{\left(\frac{1}{2}\right)^{\theta/\pi}}{1+\theta/\pi} + \frac{1-2(1-2\theta/\pi)^{1-2\theta/\pi}}{1-2\theta/\pi} \ge 0. \tag{2-3}$$

The expression on the left of (2-3) approaches zero as $\theta \to 0$. If we take the derivative of the left side of (2-3) with respect to θ and let $\theta \to 0$ we obtain $\left(1+\ln\left(\frac{1}{2}\right)\right)/\pi > 0$. Then (2-3) is true as long as $0 < \theta \le \theta_0$ for $\theta_0 > 0$ chosen small enough. Hence we conclude that $t_\theta \le \tau_\theta \le 2\theta/\pi$ for any $0 < \theta \le \theta_0$. □

From (2-1) we have $|f'_\theta(z)| \to 1$ as $|z| \to \infty$. We let ϕ_θ be the harmonic function in Ω_θ defined by

$$y^+ = \phi_\theta(u, v),$$

where $f_\theta = u + iv$. Since $1 = |\nabla\phi_\theta||f'(z)|$, we have $|\nabla\phi_\theta| \to 1$ as $|z| \to \infty$. By a rotation of $\frac{\pi}{2}$ of ϕ_θ we have a complementary harmonic function $\tilde{\phi}_\theta$ and can thus apply the ACF monotonicity formula. We have $J(\infty, \phi_\theta, \tilde{\phi}_\theta) = 1$. To find $J(0+, \phi_\theta, \tilde{\phi}_\theta)$ we find $|\nabla\phi_\theta(0)|$. This is given by

$$1 = |\nabla\phi_\theta(0)||f'(t_\theta)|.$$

Thus

$$1 \geq |\nabla\phi_\theta(0)| = \left(\frac{1-t_\theta}{1+t_\theta}\right)^{\theta/\pi},$$

so $|\nabla\phi_\theta(0)|$ is an increasing function of θ, and

$$1 \geq |\nabla\phi_\theta(0)| \geq \left(\frac{1-2\theta/\pi}{1+2\theta/\pi}\right)^{\theta/\pi}.$$

Using L'Hospital's rule we conclude that

$$\lim_{\theta\to 0} \frac{1-((1-2\theta/\pi)/(1+2\theta/\pi))^{\theta/\pi}}{(\theta/\pi)^2} = 4 > 0.$$

As a consequence we have the following result:

Lemma 2.4. *There exists θ_0 such that if $0 < \theta \leq \theta_0$, then*

$$0 < 1-\theta^2 < |\nabla\phi_\theta(0)| \leq 1. \tag{2-4}$$

Since $J(\infty, \phi_\theta) = 1$ and $J(0+, \phi_\theta) = |\nabla\phi_\theta(0)|$, Lemma 2.4 shows that

$$J(\infty, \phi_\theta) - J(0+, \phi_\theta) < 1-\theta^2.$$

3. Bounded domain

The aim of this section is to transfer the inequality in (2-4) to a harmonic function on a bounded domain. We approximate Ω_θ with domains $\Omega_{\theta,M}$; see Figure 2. If $f_{\theta,M}$ is the conformal mapping of the upper half-plane onto $\Omega_{\theta,M}$, then

$$f'_{\theta,M}(z) = \left(\frac{z+1}{z-1}\right)^{\theta/\pi}\left(\frac{z-z_2}{z+z_2}\right)^{1/2}\left(\frac{z+z_1}{z-z_1}\right)^{1/2}, \tag{3-1}$$

where $z_1, z_2 \in \mathbb{R}$ and $1 < z_1 < z_2$. We again translate $f_{\theta,M}$ by a constant so that the domain is centered on the origin as in Figure 2. The points z_1, z_2 are chosen so that $f_{\theta,M}(z_2) = M + 0i$. We point out that $|f'_{\theta,M}| \to 1$ as $|z| \to \infty$. We define $\phi_{\theta,M}(u, v) = y^+$, where $f_{\theta,M} = u + iv$.

Lemma 3.1. *Fix $\theta \leq \theta_0$. There exists $M > 0$, possibly depending on θ, such that $J(\infty, \phi_{\theta,M}) = 1$ and $J(0+, \phi_{\theta,M}) > 1-\theta^2$.*

Proof. That $J(\infty, \phi_{\theta,M}) = 1$ follows from the definition of $\phi_{\theta,M}$ and (3-1). Now from the explicit formulas given for $f'_\theta(z)$ and $f_{\theta,M}$ in (2-1) and (3-1) respectively, we have $\phi_{\theta,M} \to \phi_\theta$ in C^1 up to the boundary in a neighborhood of the origin. Since $|\nabla\phi_\theta(0)| > 1-\theta^2$, the conclusion follows. □

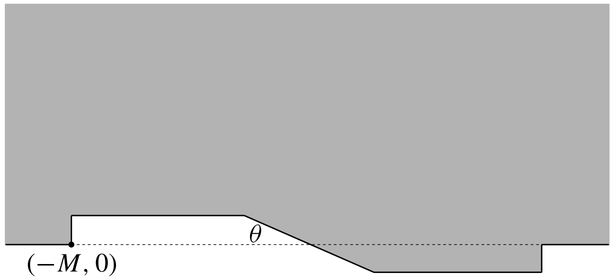

Figure 2. Domain $\Omega_{\theta,M}$.

Remark 3.2. Since $J(r, \phi_{\theta,M})$ is monotonically increasing in r, it follows from Lemma 3.1 that

$$J(\infty, \phi_{\theta,M}) - J(0+, \phi_{\theta,M}) < 1 - \theta^2.$$

For any $\theta \leq \theta_0$, we fix an M that satisfies Lemma 3.1. We now transfer the decrease in energy to a finite domain.

Lemma 3.3. *Let θ and $\phi_{\theta,M}$ be as in Lemma 3.1. Let $\Omega_{\theta,M}$ be defined as before. If we define w_R to be such that*

$$\begin{cases} \Delta w_R = 0 & \text{in } B_R \cap \Omega_{\theta,M}, \\ w_R = 0 & \text{on } \partial\Omega_{\theta,M} \cap B_R, \\ w_R = y & \text{on } (\partial B_R)^+, \end{cases}$$

then $w_R \to \phi_{\theta,M}$ locally uniformly in $\Omega_{\theta,M}$ and in C^1 in $B_\rho \cap \Omega_{\theta,M}$ for small enough ρ.

Proof. Using the rescaling

$$\phi_R := \frac{\phi_{\theta,M}(Rx, Ry)}{R},$$

we have $\phi_R \to y^+$ in C^1 on $(\partial B_1)^+$. Thus, for any $\eta > 0$, there exists $R_0 > 0$ such that if $R \geq R_0$, then

$$(1-\eta)y^+ \leq \phi_R \leq (1+\eta)y^+ \quad \text{on } (\partial B_1)^+.$$

Then rescaling back we obtain that $(1-\eta)y^+ \leq \phi_{\theta,M} \leq (1+\eta)y^+$ on $(\partial B_R)^+$ if $R \geq R_0$. From the maximum principle we then have

$$(1-\eta)w_R \leq \phi_{\theta,M} \leq (1+\eta)w_R \quad \text{for any } R \geq R_0.$$

Then as $R \to \infty$, we have $w_R \to w_\infty$ locally uniformly in $\Omega_{\theta,M}$ and in C^1 in a neighborhood of the origin. Furthermore, we have $(1-\eta)w_\infty \leq \phi_{\theta,M} \leq (1+\eta)w_\infty$. Since η can be taken to be arbitrarily small, we conclude that $w_\infty = \phi_{\theta,M}$. □

We end this section by defining a θ-turn. If $u \in \mathcal{K}$ and for some $\rho > 0$ we have $\partial\{u > 0\} \cap B_\rho$ is a line segment with inward unit normal ν, then a θ-turn in B_ρ gives a new function v with

(i) $v \in \mathcal{K}$,

(ii) $v = u$ on ∂B_1,

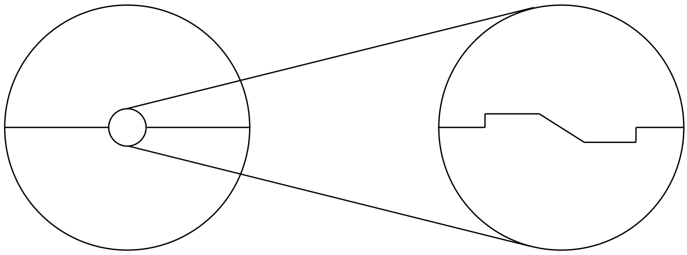

Figure 3. θ-turn when $\nu = i$.

(iii) $\partial\{v > 0\} \cap (B_1 \setminus \bar{B}_\rho) = \partial\{u > 0\} \cap (B_1 \setminus \bar{B}_\rho)$,

(iv) $\partial\{v > 0\} \cap B_\rho = \partial\{\phi_{\theta,M}(e^{i(\nu-\theta)}(2M/\rho)z) > 0\} \cap B_\rho$.

The idea of property (iv) is to shrink $\phi_{\theta,M}$ on B_{2M} to B_ρ and give v the same positivity set; see Figure 3 for when $\nu = i$.

4. Construction of counterexample

As before we let θ_0 be as in Lemma 2.4. This next lemma shows how to apply a θ-turn to a function that is almost linear at the origin.

Lemma 4.1. *Fix $\epsilon > 0$. Assume $u \in \mathcal{K}$, and that there is $s < r_0 < 1$ with*

(1) $B_s \cap \partial\{u > 0\} = B_s \cap \{y_n = 0\}$,

(2) $|u| < 2J(1,u)r_0$ *on* B_{r_0}.

If $\theta \leq \theta_0$, then there exists r, ρ with $s > r > \rho > 0$ with a θ-turn in B_ρ such that if v is the redefined function, then v satisfies

(A) $|v| < 2J(1,v)r$ *on* B_r,

(B) $|v| \leq (1+\theta^2)\sup_{B_t}|u|$ *on* B_t *for* $t \in [r_0, 1]$,

(C) $J(1,v) \leq (1+\theta^2)J(1,u)$,

(D) $J(0+,v) > (1-\theta^2)^2 J(0+,u)$.

Proof. We choose $r < s$ small enough so that

$$\left\| \frac{u(rx)}{r} - J(0+,u)y^+ \right\|_{C^1((\partial B_1)^+)} < \delta, \tag{4-1}$$

and so that $|u| < 2J(1,u)r$. We now apply a θ-turn in B_ρ with $0 < \rho < r$. As $\rho \to 0$, we have $v \to u$ uniformly away from the origin, so that by choosing ρ small enough, v satisfies (B).

We now let $\eta > 0$ be small and use a cut-off function and obtain in the standard way the Caccioppoli inequality

$$\int_{B_1 \setminus B_\eta} |\nabla v - u|^2 \leq C(\eta) \int_{B_1 \setminus B_{\eta/2}} |v-u|^2.$$

Then as $\rho \to 0$, we have $v \to u$ in $H^1(B_1 \setminus B_\eta)$ for any $\eta > 0$. We now use the monotonicity of $J(r, v)$ to prove that $v \to u$ in $H^1(B_1)$ as $\rho \to 0$. We have

$$\int_{B_\eta} |\nabla v|^2 \le \eta^2 \int_{B_1} |\nabla v|^2 = \delta^2 \int_{B_1\setminus B_\eta} |\nabla v|^2 + \int_{B_\eta} |\nabla v|^2,$$

so that

$$\int_{B_\eta} |\nabla v|^2 \le \frac{\eta^2}{1-\eta^2} \int_{B_1\setminus B_\eta} |\nabla v|^2,$$

and we conclude that

$$\int_{B_1} |\nabla v|^2 \le \frac{1}{1-\eta^2} \int_{B_1\setminus B_\eta} |\nabla v|^2.$$

Then $\|v\|_{H^1(B_1)}$ is bounded as $\rho \to 0$, so that $v \rightharpoonup u$ in $H^1(B_1)$ as $\rho \to 0$. We now have

$$\int_{B_1} |\nabla u|^2 \le \lim_{\rho\to 0} \int_{B_1} |\nabla v|^2 \le \lim_{\rho\to 0} \frac{1}{1-\eta^2} \int_{B_1\setminus B_\eta} |\nabla v|^2 = \frac{1}{1-\eta^2} \int_{B_1\setminus B_\eta} |\nabla u|^2.$$

Since η can be chosen arbitrarily small, we have $\nabla v \to \nabla u$ in $L^2(B_1)$ and thus conclude that $v \to u$ in $H^1(B_1)$ as $\rho \to 0$. Consequently, we may choose ρ even smaller so that properties (A) and (C) hold.

From (4-1), if ρ is chosen small enough we have

$$\left\| \frac{v(rz)}{r} - J(0+, u)y^+ \right\|_{C^1((\partial B_1)^+)} < \delta,$$

so that $(1-\delta)J(0+, u)y^+ \le v(rz)/r$ on $(\partial B_1)^+$. We now define w to be the solution to

$$\begin{cases} \Delta w = 0 & \text{in } \{v(rz)/r > 0\} \cap B_1, \\ w = 0 & \text{on } \partial\{v(rz)/r > 0\} \cap B_1, \\ w = (1-\delta)J(0+)y^+ & \text{on } (\partial B_1)^+. \end{cases}$$

We have $w \le v$ in B_1, so that $|\nabla w(0)| \le |\nabla v(0)|$ or $J(0+, w) \le J(0+, v)$. We may rescale w and apply Lemma 3.3 to obtain that for small enough ρ, we have

$$J(0+, w) > (1-\theta^2)(1-\delta)J(0+, u).$$

By choosing $\delta < \theta^2$ we obtain (D). □

Proof of Theorem 1.3 in dimension $n = 2$. We now use Lemma 4.1 to construct a sequence $u_k \in \mathcal{K}$ with $\lim u_k \to u$. The pair u and $\tilde{u}(z) := u(-z)$ will be a counterexample to Claim 1.2. The sequence u_k is constructed inductively as follows. We choose $\theta_k = 1/(k + N_0)$, where $N_0 \in \mathbb{N}$ is chosen large enough so that $\theta_k \le \theta_0$. We then let $u_0 = y^+$ on B_1. By Lemma 4.1 there exists $\rho_1 < r_1$ such that if a θ_1-turn is applied in B_{ρ_1} to obtain u_1, then u_1 will satisfy properties (A)–(D). We now suppose that u_k has been constructed for some $k \ge 1$. By rotating u_k it will satisfy assumption (1) of Lemma 4.1. Assumption (2) will also be satisfied because u_k satisfies (A) for $r = r_k$. By Lemma 4.1 there exists $\rho_{k+1} < r_{k+1}$ with $r_{k+1} < \rho_k$ so that if we apply a θ_{k+1}-turn to u_k to obtain u_{k+1} we have

(i) $|u_{k+1}| < 2J(1, u_{k+1})r$ on B_r,

(ii) $|u_{k+1}| \le \prod_{j=1}^{k}(1+\theta_j^2) \sup_{B_t} |u_0|$ on B_t for $t \in [r_k, 1]$,

(iii) $J(1, u_{k+1}) \le \prod_{j=1}^{k}(1+\theta_j^2) J(1, u_0) = \prod_{j=1}^{k}(1+\theta_j^2)$,

(iv) $J(0+, u_{k+1}) > \prod_{j=1}^{k}(1-\theta_j^2)^2 J(0+, u_0) = \prod_{j=1}^{k}(1-\theta_j^2)^2$.

From the same arguments involving the Caccioppoli inequality as in the proof of Lemma 4.1, there exists u such that $u_k \to u$ in $H^1(B_1)$ and locally uniformly away from the origin. Then u is continuous away from the origin. From (i) we obtain that $|u| \le Cr$ on B_r for $0 < r \le 1$, so that u is continuous up to the origin, and $u(0) = 0$.

Now $0 < \prod_{k=1}^{\infty}(1-\theta_k^2)^2$ if and only if $0 < \prod_{k=1}^{\infty}(1-\theta_k^2)$ if and only if

$$\sum_{k=1}^{\infty}(k+N_0)^{-2} = \sum_{k=1}^{\infty}\theta_k^2 < \infty.$$

Since the inequality above is true, we conclude that

$$0 < \prod_{k=1}^{\infty}(1-\theta_k^2)^2 < \prod_{k=1}^{\infty}(1-\theta_k^2) < 1.$$

The last inequality above is due to the fact that all the terms are less than 1. Since $u_k \to u$ in $H^1(B_1)$ and from properties (ii) and (iii), we conclude that

$$0 < \prod_{k=1}^{\infty}(1-\theta_k^2)^2 \le J(r, u) \le CJ(1, u) < \infty \quad \text{for all } 0 < r \le 1,$$

so that $J(0+, u) > 0$.

If we let $\tilde{u}_k(z) = u_k(-z)$, then $\tilde{u}_k \to \tilde{u}$, where $\tilde{u}(z) = u(-z)$. Furthermore, $u \cdot \tilde{u} = 0$ in B_1. Since also $u, \tilde{u}$ are nonnegative, continuous, and harmonic when positive, they satisfy the assumptions of the ACF monotonicity formula in Proposition 1.1. We now show that $u, \tilde{u}$ are a counterexample to Claim 1.2. We assume by way of contradiction that $\{u > 0\}$ and $\{\tilde{u} > 0\}$ are tangent at the origin and after a rotation $u(z) + \tilde{u}(z) = \alpha x_1^+ + \beta x_1^- + o(|z|)$. Then for any small $\delta > 0$, there exists r_0 such that if $r \le r_0$ and $|z| > \frac{1}{2}$ and $|\mathrm{Arg}(z)| < \delta$, then

$$\frac{u(rz) + \tilde{u}(rz)}{r} > \frac{\alpha x_1^+}{2} > 0. \tag{4-2}$$

We now recall that from the construction

$$\partial\{u > 0\} \cap (B_{r_k} \setminus B_{\rho_k}) = \{z : z = te^{-i\sum_{j=1}^{k}\theta_j} \text{ and } \rho_k \le |t| < r_k\}. \tag{4-3}$$

Since $\sum \theta_k = \infty$ and $\theta_k \to 0$, we obtain from (4-3) there exist infinitely many z_k with $|z_k| \to 0$ and $|\mathrm{Arg}(z)| < \delta$ such that $u(z_k) + \tilde{u}(z_k) = 0$. This contradicts (4-2), and so Claim 1.2 is not true. □

We now show that the pair u and $\tilde{u}$ are also a counterexample in higher dimensions.

Proof of Theorem 1.3 in dimension $n > 2$. For u as in the proof for dimension 2, we let $w_n(x_1, x_2, \dots, x_n) = u(x_1, x_2)$. Since in dimension $n = 2$ we have

$$\frac{1}{r^2}\int_{B_r} |\nabla u|^2 \geq C > 0,$$

it follows that in dimension n,

$$\frac{1}{r^n}\int_{B_r} |\nabla w|^2 \geq C.$$

Then

$$\frac{1}{r^2}\int_{B_r} \frac{|\nabla w|^2}{|x|^{n-2}} \geq \frac{1}{r^2}\int_{B_r} \frac{|\nabla w|^2}{r^{n-2}} = \frac{1}{r^n}\int_{B_r} |\nabla w|^2 \geq C > 0,$$

so that $\Phi(r, w, \tilde{w}) > 0$. We have already shown that $u + \tilde{u}$ cannot satisfy the conclusions in Claim 1.2; consequently, $w + \tilde{w}$ also do not satisfy those conclusions. □

Acknowledgments

Kriventsov was supported by the NSF MSPRF fellowship DMS-1502852. The authors would like to thank Luis Caffarelli for helpful conversations regarding this problem.

References

[Allen and Petrosyan 2012] M. Allen and A. Petrosyan, "A two-phase problem with a lower-dimensional free boundary", *Interfaces Free Bound.* **14**:3 (2012), 307–342. MR Zbl

[Allen and Shi 2016] M. Allen and W. Shi, "The two-phase parabolic Signorini problem", *Indiana Univ. Math. J.* **65**:2 (2016), 727–742. MR Zbl

[Allen et al. 2015] M. Allen, E. Lindgren, and A. Petrosyan, "The two-phase fractional obstacle problem", *SIAM J. Math. Anal.* **47**:3 (2015), 1879–1905. MR Zbl

[Alt et al. 1984] H. W. Alt, L. A. Caffarelli, and A. Friedman, "Variational problems with two phases and their free boundaries", *Trans. Amer. Math. Soc.* **282**:2 (1984), 431–461. MR Zbl

[Azzam et al. 2016] J. Azzam, S. Hofmann, J. M. Martell, S. Mayboroda, M. Mourgoglou, X. Tolsa, and A. Volberg, "Rectifiability of harmonic measure", *Geom. Funct. Anal.* **26**:3 (2016), 703–728. MR Zbl

[Blank 2001] I. Blank, "Sharp results for the regularity and stability of the free boundary in the obstacle problem", *Indiana Univ. Math. J.* **50**:3 (2001), 1077–1112. MR Zbl

[Caffarelli and Salsa 2005] L. Caffarelli and S. Salsa, *A geometric approach to free boundary problems*, Graduate Studies in Math. **68**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[Caffarelli et al. 2009] L. A. Caffarelli, A. L. Karakhanyan, and F.-H. Lin, "The geometry of solutions to a segregation problem for nondivergence systems", *J. Fixed Point Theory Appl.* **5**:2 (2009), 319–351. MR Zbl

[Kriventsov and Lin 2018] D. Kriventsov and F. Lin, "Regularity for shape optimizers: the nondegenerate case", *Comm. Pure Appl. Math.* **71**:8 (2018), 1535–1596. MR Zbl

[Kriventsov and Lin 2019] D. Kriventsov and F. Lin, "Regularity for shape optimizers: the degenerate case", *Comm. Pure Appl. Math.* **72**:8 (2019), 1678–1721. MR

[Petrosyan et al. 2012] A. Petrosyan, H. Shahgholian, and N. Uraltseva, *Regularity of free boundaries in obstacle-type problems*, Graduate Studies in Math. **136**, Amer. Math. Soc., Providence, RI, 2012. MR Zbl

[Terracini et al. 2019] S. Terracini, G. Verzini, and A. Zilio, "Spiraling asymptotic profiles of competition-diffusion systems", *Comm. Pure Appl. Math.* (online publication March 2019).

[White 1992] B. White, "Nonunique tangent maps at isolated singularities of harmonic maps", *Bull. Amer. Math. Soc.* (*N.S.*) **26**:1 (1992), 125–129. MR Zbl

Received 26 Feb 2018. Revised 13 Sep 2018. Accepted 19 Dec 2018.

MARK ALLEN: allen@mathematics.byu.edu
Department of Mathematics, Brigham Young University, Provo, UT, United States

DENNIS KRIVENTSOV: dennisk@cims.nyu.edu
Courant Institute of Mathematical Sciences, New York University, New York, NY, United States

ANALYSIS AND PDE
Vol. 13, No. 1, 2020
dx.doi.org/10.2140/apde.2020.13.215

INFINITE-TIME BLOW-UP FOR THE 3-DIMENSIONAL ENERGY-CRITICAL HEAT EQUATION

MANUEL DEL PINO, MONICA MUSSO AND JUNCHENG WEI

We construct globally defined in time, unbounded positive solutions to the energy-critical heat equation in dimension 3

$$u_t = \Delta u + u^5 \quad \text{in } \mathbb{R}^3 \times (0,\infty), \qquad u(x,0) = u_0(x) \quad \text{in } \mathbb{R}^3.$$

For each $\gamma > 1$ we find initial data (not necessarily radially symmetric) with $\lim_{|x|\to\infty} |x|^\gamma u_0(x) > 0$ such that as $t \to \infty$

$$\|u(\cdot,t)\|_\infty \sim t^{\gamma-\frac{1}{2}} \quad \text{if } 1<\gamma<2, \qquad \|u(\cdot,t)\|_\infty \sim \sqrt{t} \quad \text{if } \gamma>2, \qquad \|u(\cdot,t)\|_\infty \sim \sqrt{t}(\ln t)^{-1} \quad \text{if } \gamma=2.$$

Furthermore we show that this infinite-time blow-up is codimensional-1 stable. The existence of such solutions was conjectured by Fila and King (*Netw. Heterog. Media* **7**:4 (2012), 661–671).

1. Introduction

Let $n \geq 3$. The *energy-critical heat equation* in $\mathbb{R}^n$ is the parabolic Cauchy problem

$$\begin{cases} u_t = \Delta u + |u|^{\frac{4}{n-2}} u & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(\cdot,0) = u_0 & \text{in } \mathbb{R}^n. \end{cases} \tag{1-1}$$

The *energy*

$$E(u) = \frac{1}{2}\int_{\mathbb{R}^n} |\nabla u|^2 - \frac{n-2}{2n}\int_{\mathbb{R}^n} |u|^{\frac{2n}{n-2}}$$

defines a Lyapunov functional for problem (1-1). In fact for classical solutions $u(x,t)$ with sufficient decay in space variable we have that

$$\frac{d}{dt}E(u(\cdot,t)) = -\int_{\mathbb{R}^n} |u_t|^2.$$

Classical parabolic theory yields that the Cauchy problem (1-1) is well-posed in its natural finite-energy space for short time intervals.

In this paper we are interested in *positive finite-energy solutions* of (1-1) which are global in time, namely defined and smooth in the entire time interval $(0,\infty)$. The presence of the Lyapunov functional implies that limits of bounded solutions along sequences $t = t_n \to +\infty$ can only be steady states, namely solutions of the Yamabe equation

$$\Delta u + |u|^{\frac{4}{n-2}} u = 0 \quad \text{in } \mathbb{R}^n. \tag{1-2}$$

MSC2010: 35B33, 35B40, 35K58.
Keywords: blow-up, critical exponents, nonlinear parabolic equations.

All *positive* solutions of (1-2) are given by the *Aubin–Talenti bubbles*

$$U_{\mu,\xi}(x) = \mu^{-\frac{n-2}{2}} w\left(\frac{x-\xi}{\mu}\right),$$

where $\mu > 0,\ \xi \in \mathbb{R}^n$ and

$$w(x) = (n(n-2))^{\frac{n-2}{4}} \left(\frac{1}{1+|x|^2}\right)^{\frac{n-2}{2}}.$$

They are precisely the extremals of Sobolev's embedding. The *criticality* of problem (1-1) refers to the presence of this continuum of steady states which become singular as $\mu \to 0$, in addition to energy invariance. In fact we immediately see that

$$E(U_{\mu,\xi}) = E(U) \quad \text{for all } \xi \in \mathbb{R}^n,\ \mu > 0.$$

A solution $u(x,t)$ of (1-1) which around one or more points of space looks like $u(x,t) \approx U_{\mu(t),\xi(t)}(x)$ with $\mu(t) \to 0$ is called a bubbling blow-up solution. Bubbling phenomena is present in many important time-dependent and stationary settings, usually carrying deep meaning in the global structure of their solutions. Notable examples include the Yamabe and harmonic map flows and the Keller–Segel chemotaxis system; see [Ciraolo et al. 2018; Daskalopoulos et al. 2018; Raphaël and Schweyer 2013; Davila et al. 2017; Ghoul and Masmoudi 2016]. In the last decade or so it has been extensively studied in energy-critical wave equations, Schrödinger maps and other dispersive settings.

Problem (1-1) is a simple-looking model which contains much of the complexity of the bubbling blow-up issue. Basic questions have remain unanswered until today. Existence or nonexistence of infinite-time bubbling positive solutions in problem (1-1) is not known. This question has been explicitly stated for instance in [Poláčik and Yanagida 2014; Quittner and Souplet 2007, Remark 22.10]. Detecting such solutions rigorously is not easy. Usual behaviors in the flow (1-1) are either asymptotic vanishing or blow-up in finite time. Global solutions with nontrivial asymptotic patterns are typically unstable objects and hence harder to be detected.

In a very interesting paper Fila and King [2012] provided insight on the question in the case of a radially symmetric, positive initial condition with an exact power decay rate. Using formal matching asymptotic analysis, they demonstrated that the power decay determines the blow-up rate in a precise manner. Intriguingly enough, their analysis leads them to conjecture that infinite-time blow-up *should only happen* in low dimensions 3 and 4; see Conjecture 1.1 in [Fila and King 2012].

In this paper we rigorously establish the existence of solutions with infinite-time blow-up in dimension 3, confirming the conjecture in [Fila and King 2012]. Thus we consider the Cauchy problem

$$\begin{cases} u_t = \Delta u + u^5 & \text{in } \mathbb{R}^3 \times (0,\infty), \\ u(\cdot\,,0) = u_0 & \text{in } \mathbb{R}^3 \end{cases} \tag{1-3}$$

for an initial datum u_0 which we assume first radially symmetric with an exact power decay of the form

$$\lim_{|x|\to\infty} |x|^\gamma u_0(x) =: A > 0. \tag{1-4}$$

As in [Fila and King 2012] we assume that $\gamma > 1$, which means that u_0 decays faster than the bubble

$$w(x) = 3^{\frac{1}{4}} \left(\frac{1}{1+|x|^2} \right)^{\frac{1}{2}}. \tag{1-5}$$

Theorem 1.1. *Given $\gamma > 1$, there exists a positive, radially symmetric global solution $u(x,t)$ to problem (1-3) whose initial condition $u_0(|x|)$ satisfies (1-4) and as $t \to +\infty$*

$$\|u(\cdot,t)\|_{L^\infty(\mathbb{R}^3)} \sim \begin{cases} t^{\frac{\gamma-1}{2}} & \text{if } 1<\gamma<2, \\ \sqrt{t}/\ln t & \text{if } \gamma = 2, \\ \sqrt{t} & \text{if } \gamma > 2. \end{cases} \tag{1-6}$$

More precisely, the blow-up takes place by bubbling near the origin. The solution of Theorem 1.1 is in the inner self-similar region, $|x| \ll \sqrt{t}$, in leading order of the *bubbling blow-up form*

$$u(x,t) \sim \frac{1}{\mu(t)^{\frac{1}{2}}} w\left(\frac{x}{\mu(t)} \right),$$

where

$$\mu(t) \sim \begin{cases} t^{1-\gamma} & \text{if } 1<\gamma<2, \\ t^{-1}\ln^2 t & \text{if } \gamma = 2, \\ t^{-1} & \text{if } \gamma > 2 \end{cases} \tag{1-7}$$

and w is given by (1-5). In the outer self-similar region $|x| \gg \sqrt{t}$, the solution dissipates in the form of a self-similar solution of heat equation $u_t = \Delta u$ in $\mathbb{R}^3 \times (0,\infty)$.

A surprising feature of the construction is the dynamics discovered for the scaling parameter $\mu(t)$. It has a highly nonlocal character governed by an equation involving a perturbation of the fractional $\frac{1}{2}$-Caputo derivative. In fact, in order to find the precise lower-order corrections needed for the scaling parameter $\mu(t)$ we will need to solve linear equations of the type

$$\int_0^t \frac{\beta'(s)}{\sqrt{t-s}} (1 - e^{-\frac{M^2}{(t-s)}})\, ds = h(t)$$

for suitably decaying right-hand sides $h(t)$. See (6-8) and (6-13) below.

Problem (1-1) is a special case of the *Fujita equation*

$$\begin{cases} u_t = \Delta u + u^p & \text{in } \mathbb{R}^n \times (0,\infty), \\ u(\cdot,0) = u_0 & \text{in } \mathbb{R}^n, \end{cases} \tag{1-8}$$

with $p > 1$. Blow-up phenomena in problem (1-8) is extremely sensitive to the values of the exponent p. A vast literature has been devoted to this problem after the seminal work [Fujita 1966]. We refer the reader for instance to the book [Quittner and Souplet 2007] for background and a comprehensive account of results until 2007 and to the more recent works [Matano and Merle 2004; 2009; 2011]. The case $p = (n+2)/(n-2)$ is special in many ways. Positive steady states do not exist when $p < (n+2)/(n-2)$. Positive radial global solutions must be bounded and go to zero; see [Poláčik and Quittner 2006; Poláčik et al. 2007; Quittner and Souplet 2007]. They exist when $p > (n+2)/(n-2)$ but they have infinite energy; see [Gui et al. 1992]. Infinite-time blow-up exists in that case but it has an entirely different nature; see [Poláčik and Yanagida 2003; 2014].

The study of energy-critical problems has attracted much attention in the last decade. For energy-critical wave equations, blow-up solutions have been characterized and constructed in [Duyckaerts et al. 2012; 2013; 2016a; 2016b; Krieger et al. 2009]. In [Schweyer 2012] Type-II sign changing, finite time blow-up for (1-1) is constructed, first formally predicted in [Filippas et al. 2000]. Threshold dynamics around the steady states of (1-1) has been characterized in large dimensions $n \geq 7$ in [Collot et al. 2017]. Also in large dimensions $n \geq 5$ in [Cortázar et al. 2016] infinite-time bubbling solutions of (1-1) in a bounded domain under Dirichlet boundary conditions are constructed for $n \geq 5$. The cases $n = 3, 4$ are indeed considerably more delicate and not treated there. The solutions in Theorem 1.1 are specially meaningful for the full dynamics since they are *threshold solutions* in the sense that the solution of (1-3) with initial condition λu_0 goes to zero as $t \to \infty$ if $\lambda < 1$, while it blows-up in finite time if $\lambda > 1$. Radial threshold solutions for various ranges of exponents in (1-3) are analyzed in [Quittner and Souplet 2007].

We recall that from [Fila and King 2012], it is not expected to have this blow-up in entire space in dimensions $n \geq 5$. Our approach is entirely different from that in [Schweyer 2012] for $n = 4$ in which a finite-time type-II blow-up solution of (1-1) is constructed on the basis of the modulation equation methods developed for critical dispersive equations in [Donninger and Krieger 2013; Ortoleva and Perelman 2013; Merle et al. 2013; Raphaël and Rodnianski 2012; Raphaël and Schweyer 2013].

Our approach has a parabolic-elliptic flavor, in line with the recent works [Cortázar et al. 2016; Davila et al. 2017]. Since our proofs only rely on elliptic and parabolic estimates, we can easily modify the proof to deal with nonradial and general initial data, in particular establishing *codimension-1 stability* of the solution built. This is concordant with a result of [Krieger et al. 2015] on the corresponding wave analogue. In Section 10 we prove the following:

Theorem 1.2. *Let $\bar{v}_0 = \bar{v}_0(x)$ be a positive continuous function, uniformly bounded for $x \in \mathbb{R}^3$. Let $\gamma > 1$ and $\kappa > \max\{(\gamma+3)/2, \gamma\}$. Then, there exists a positive global solution $u(x,t)$ to problem (1-3) with initial condition*

$$u(x,0) = u_0(|x|) + \frac{\bar{v}_0(x)}{|x|^\kappa}\left[1 - \eta\left(\frac{|x|}{t_0}\right)\right],$$

where u_0 is positive, radially symmetric, satisfies (1-4), $t_0 > 0$ is a fixed large number and η is a smooth cut-off function with $\eta(s) = 1$ for $s < 1$ and $\eta(s) = 0$ for $s > 2$. As $t \to +\infty$, $u(x,t)$ satisfies (1-6).

Furthermore, there exists a codimension-1 manifold of functions in $C^1(\mathbb{R}^3)$ converging to 0 at infinity, with a sufficiently fast decay, that contains $u_0(|x|) + (\bar{v}_0(x)/|x|^\kappa)(1-\eta(|x|/t_0))$ such that if $\bar{u}_0$ lies in that manifold and it is sufficiently close to $u_0(|x|) + (\bar{v}_0(x)/|x|^\kappa)(1-\eta(|x|/t_0))$ in the sense that $\bar{u}_0 = u_0(|x|) + (\bar{v}_0(x)/|x|^\kappa)(1-\eta(|x|/t_0)) + \mathcal{O}(|x|e^{-b|x|})$ for some $b > 0$, then the solution $\bar{u}(x,t)$ to (1-3) with $\bar{u}(x,0) = \bar{u}_0(x)$ is global in time and satisfies (1-6).

In the nonradial setting, the profile of the solution in the inner self-similar regime is

$$u(x,t) \sim \frac{1}{\mu(t)^{\frac{1}{2}}} w\left(\frac{x - p(t)}{\mu(t)}\right), \qquad \frac{|p(t)|}{\mu(t)} \to 0 \quad \text{as } t \to \infty,$$

where w is given by (1-5) and μ satisfies the asymptotics (1-7). Precise description of the dynamics of the center $p = p(t)$ is provided.

A surprising feature of the construction is the dynamics discovered for the scaling parameter $\mu(t)$. It has a highly nonlocal character governed by an equation involving a perturbation of the fractional $\frac{1}{2}$-Caputo derivative. In fact, in order to find the precise lower-order corrections needed for the scaling parameter $\mu(t)$ we will need to solve linear equations of the type

$$\int_0^t \frac{\beta'(s)}{\sqrt{t-s}}(1-e^{-\frac{M^2}{(t-s)}})\,ds = h(t)$$

for suitably decaying right-hand sides $h(t)$. See (6-8) and (6-13) below.

We believe that an approach similar to that in this paper could be used to prove the existence of a global unbounded solution when $N = 4$, $p = 3$ as conjectured in [Fila and King 2012]. We will undertake that issue in a future work.

The proof of Theorem 1.1 starts with the construction of an approximate solution to problem (1-3) with the asymptotic behavior described in (1-6). This is done in full detail in Section 2. We then show the existence of an actual solution to problem (1-3) deforming the approximation, by means of a *inner-outer gluing* procedure. This scheme is described in Section 3, and its proof is addressed in Sections 4–9. In Section 10 we prove Theorem 1.2. Appendices A–C gather some technical results needed to prove the theorems.

In the rest of the paper, we shall denote by C a generic positive constant, whose value may change from line to line, and within the same line. We shall use the notation $\boldsymbol{c}$ to indicate a positive constant, with $\boldsymbol{c} < 1$, whose explicit value may change from line to line. Furthermore, t_0 will denote a large fixed positive number and

$$\eta : \mathbb{R} \to \mathbb{R} \tag{1-9}$$

a smooth cut-off function with $\eta(s) = 1$ for $s < 1$ and $= 0$ for $s > 2$.

2. Construction of an approximate solution and estimate of the associated error

After shifting the initial time to $t_0 > 0$, problem (1-3) takes the form

$$u_t = \Delta u + u^5 \quad \text{in } \mathbb{R}^3 \times (t_0, \infty), \tag{2-1}$$

with initial condition $u_0(r) = u(r, t_0)$ satisfying

$$\lim_{r\to\infty} r^\gamma u_0(r) = A > 0 \quad \text{for some } \gamma > 1. \tag{2-2}$$

This section is devoted to the construction of a first approximation for a solution to (2-1)–(2-2), and to the description of the associated error.

The first approximation is built by matching an inner profile, made upon solving the elliptic problem

$$\Delta u + u^5 = 0 \quad \text{in } \mathbb{R}^3 \tag{2-3}$$

and an outer profile, made upon solving the heat equation in the whole space

$$u_t = \Delta u \quad \text{in } \mathbb{R}^3, \tag{2-4}$$

in the set of functions satisfying the decaying conditions (2-2). It is constructed in Section 2A (for the inner profile) and Section 2B (for the outer profile), and in Section 2C we derive a precise description of the *error* of approximation. In [Fila and King 2012], this approximate solution was already derived. We realize though that, for our rigorous construction to work, we need a further improvement of the approximation. This is done in Section 2D, where we introduce a next correction term, and describe the associated error. It turns out that this next correction term gives the right dynamics for the blow-up rate which turns out to be governed by a nonlocal differential equation with a fractional time-derivative closely related to the so-called $\frac{1}{2}$-Caputo derivative. See (6-13).

2A. ***Construction of the first inner profile.*** We recall that all positive radially symmetric solutions to (2-3) constitute a one-parameter family of functions, which are given explicitly by

$$w(r) = 3^{\frac{1}{4}}\left(\frac{1}{1+r^2}\right)^{\frac{1}{2}}, \quad w_\mu(r) = \mu^{-\frac{1}{2}} w\left(\frac{r}{\mu}\right), \tag{2-5}$$

for any positive number $\mu > 0$; see [Aubin 1976; Caffarelli et al. 1989]. We denote by Z_0 the only bounded and radial function belonging to the kernel of the linear operator

$$L_0(\phi) = \Delta\phi + 5w^4\phi. \tag{2-6}$$

See [Rey 1990]. The function Z_0 is explicitly defined by

$$Z_0(r) = -\left[\frac{w}{2} + w'(r)r\right] = \frac{3^{\frac{1}{4}}}{2}\frac{r^2-1}{(1+r^2)^{\frac{3}{2}}}. \tag{2-7}$$

Given Z_0, we denote by $\Phi_1(r)$ the solution to

$$\Delta\Phi_1 + 5w^4\Phi_1 = Z_0, \tag{2-8}$$

defined as

$$\Phi_1(r) = \Phi_0(r) + \pi_0 + \bar{\Phi}_1(r), \quad \text{where } \Phi_0(r) = \frac{3^{\frac{1}{4}}}{4} r, \tag{2-9}$$

$$\left(5\int_0^\infty w^4 Z_0 r^2\,dr\right)\pi_0 = \int_0^\infty \left(Z_0 - \frac{3^{\frac{1}{4}}}{2r}\right) Z_0 r^2\,dr - 5\int_0^\infty w^4\Phi_0 Z_0 r^2\,dr,$$

and $\bar{\Phi}_1$ is the unique solution to

$$\Delta\phi + 5w^{p-1}\phi = \underbrace{\left(Z_0 - \frac{3^{\frac{1}{4}}}{2r}\right) - 5w^4(\Phi_0 + \pi_0)}_{:=\Pi_0(r)},$$

explicitly given by

$$\bar{\Phi}_1(r) = \tilde{Z}(r)\int_0^r \Pi_0(s)Z_0(s)s^2\,ds - Z_0(r)\int_0^r \Pi_0(s)\tilde{Z}(s)s^2\,ds.$$

In the above expression, $\tilde{Z}$ denotes another solution to $\Delta\phi + 5w^4\phi = 0$, linearly independent to Z_0. $\tilde{Z}$ satisfies the asymptotic behavior $\tilde{Z}(s) \sim s^{-1}$ as $s \to 0$, and $\tilde{Z}(s) \sim 1$ as $s \to \infty$.

A closer look at the expression of $\bar{\Phi}_1$ gives

$$\|r^{2-\sigma}\bar{\Phi}_1(r)\|_\infty < C$$

for some fixed positive constant C and any $\sigma > 0$ small.

Remark 2.1. The solution to (2-8) is not unique. (In fact one can add any multiple of Z_0.) The choice we made in (2-9) is used to match the outer solution in the next section.

We have now the elements to define the first inner profile. We introduce a smooth positive function $\mu(t)$ of the form

$$\mu(t) = \mu_0(t)(1+\Lambda(t))^2, \quad \text{where } \mu_0(t) > 0, \quad \lim_{t\to\infty} \mu_0(t) = 0. \tag{2-10}$$

The function μ_0 will be defined below, see (2-23), (2-32), (2-36), as an explicit function of t depending on the decay rate γ. On the other hand, the function $\Lambda = \Lambda(t)$ will be left as a parameter in the construction, and it will be determined in the final argument to get an actual solution to the problem. In the meanwhile, we shall assume that $\Lambda = \Lambda(t)$ is a smooth function in (t_0, ∞), defined by

$$\Lambda(t) := \int_t^\infty \lambda(s)\,ds,$$

where λ satisfies

$$\|\lambda\|_\sharp := \sup_{t>t_0} \mu_0(t)^{-1} t\,[\|\lambda\|_{\infty,[t,t+1]} + [\lambda]_{0,\sigma,[t,t+1]}] \le \ell \tag{2-11}$$

for $\sigma = \frac{1}{2} + \sigma'$, with $\sigma' > 0$ small, and for some fixed constant ℓ. Here we intend

$$\|f\|_{\infty,[t,t+1]} = \sup_{s\in[t,t+1]} |f(s)|,$$
$$[f]_{0,\sigma,[t,t+1]} = \sup_{s_1\neq s_2\in[t,t+1]} \frac{|f(s_1)-f(s_2)|}{|s_1-s_2|^\sigma}.$$

For a later purpose we introduce the space

$$X_\sharp = \{\lambda \in C(t_0,\infty) : \|\lambda\|_\sharp \text{ is bounded}\}. \tag{2-12}$$

With this in mind, we define the inner approximation to be

$$u_{\text{in}}(r,t) = w_\mu(r) + \mu_0'\psi_1(r,t), \quad \psi_1(r,t) = \mu^{\frac{1}{2}}\Phi_1\left(\frac{r}{\mu}\right). \tag{2-13}$$

A direct computation gives

$$\Delta\psi_1 + 5w_\mu^4\psi_1 = -\mu^{-\frac{3}{2}} Z_0\left(\frac{r}{\mu}\right) = \frac{\partial w_\mu}{\partial\mu}(r).$$

In the region $\{r : r > R\mu_0\}$, where R is any large but fixed positive number, the inner approximation looks like

$$u_{\rm in}(r,t) = 3^{\frac14}\frac{\mu^{\frac12}}{r} - \frac{3^{\frac14}}{4}\mu_0'\mu^{-\frac12} r + \mu_0^{\frac12}\mu_0'\Theta[\mu](r,t) + \frac{\mu_0^{\frac12}}{r}\left(\frac{\mu_0}{r}\right)^2\Theta[\mu](r,t), \tag{2-14}$$

where $\Theta[\mu](r,t)$ denotes a generic function, which depends smoothly on μ, and on (r,t), and which is uniformly bounded, for parameters μ satisfying (2-10), for r in the considered region, and any t large.

2B. *Construction of the first outer profile and choice of $\mu_0(t)$.* The outer profile is chosen to satisfy the heat equation $u_t = \Delta u$, in the whole space $\mathbb{R}^3$, and to fit the requested decaying property for the initial condition (2-2). Its properties and exact definitions change depending on the value of the decay rate γ of the initial condition u_0; see (2-2). We consider three different situations: $1 < \gamma < 2$, $\gamma = 2$, and $\gamma > 2$.

Case 1: $1 < \gamma < 2$. In this case we define $u_{\rm out}$ as

$$u_{\rm out}(r,t) = t^{-\frac{\gamma}{2}} g\left(\frac{r}{\sqrt{t}}\right), \tag{2-15}$$

with g the positive solution to

$$g''(s) + \left(\frac{2}{s} + \frac{s}{2}\right)g'(s) + \frac{\gamma}{2}g(s) = 0, \quad s \in (0,\infty), \tag{2-16}$$

that satisfies the properties

(1) $\lim_{s\to\infty} s^\gamma g(s) = A$,

(2) $\lim_{s\to 0^+} sg(s) = d$ for a certain positive constant d for which $\lim_{s\to 0^+}[g(s) - d/s] = 0$.

Such a function g indeed exists. Let

$$L_\nu(g) = g'' + \left(\frac{2}{s} + \frac{s}{2}\right)g' + \nu g, \quad s \in (0,\infty).$$

In Appendix A, we prove the following:

Lemma 2.2. *If $\frac12 < \nu < 1$, there exist two positive linearly independent solutions $y_1 = y_1(s)$ and $y_2 = y_2(s)$ to*

$$L_\nu(g) = 0, \quad s \in (0,\infty), \tag{2-17}$$

that satisfy respectively

$$y_1(s) = \frac{1}{s} + (\nu - 1)\left(\int_0^\infty s y_1(s)\,ds\right) + \frac{1-2\nu}{4}s + O(s^2) \quad \text{if } s \to 0^+, \tag{2-18}$$

$$y_2(s) = c_2 + o(s) \quad \text{if } s \to 0^+, \tag{2-19}$$

$$y_1(s) = c_1 e^{-\frac{s^2}{4}} s^{4\nu-3}, \quad y_2(s) = \frac{1}{s^{2\nu}}\left(1 + o\left(\frac{1}{s}\right)\right) \quad \text{if } s \to \infty \tag{2-20}$$

for some positive constants c_1, c_2.

Thanks to the lemma, which we apply to solve (2-16) when $\nu = \gamma/2$, we get that the function g we are looking for in (2-15) is thus given by

$$g(s) = dy_1(s) + Ay_2(s), \quad \text{with } d = \frac{2Ay_2(0)}{(2-\gamma)\big(\int_0^\infty sy_1(s)\,ds\big)} > 0. \tag{2-21}$$

We observe that, in a region like $r < R^{-1}\sqrt{t}$, for some large but fixed R, we get

$$u_{\text{out}}(r,t) = d\frac{t^{-\frac{\gamma-1}{2}}}{r} + t^{-\frac{\gamma+1}{2}} A \frac{(1-\gamma)y_2(0)}{2(2-\gamma)\int_0^\infty zy_1(z)\,dz} r + t^{-\frac{\gamma}{2}} O\Big(\frac{r^2}{t}\Big). \tag{2-22}$$

We next choose the function $\mu_0(t)$ in the definition of $\mu(t)$, (2-10), in such a way that the functions u_{in} and u_{out} automatically match in the whole region $R\mu_0 < r < R^{-1}\sqrt{t}$, for some R large, but fixed independent of t. This is possible if

$$\mu_0(t) = \frac{d^2}{\sqrt{3}} t^{1-\gamma}. \tag{2-23}$$

Indeed, with this choice for $\mu_0(t)$, and given the bound (2-11), there exists a constant C so that

$$|u_{\text{in}}(r,t) - u_{\text{out}}(r,t)| \le C\frac{\mu_0^{\frac{1}{2}}}{r}, \quad |\nabla u_{\text{in}}(r,t) - \nabla u_{\text{out}}(r,t)| \le C\frac{\mu_0^{\frac{1}{2}}}{r^2} \tag{2-24}$$

for any $R\mu_0 < r < R^{-1}\sqrt{t}$, and t large enough.

Case 2: $\gamma = 2$. In this case, we define u_{out} as

$$u_{\text{out}}(r,t) = t^{-1}(\log t)kAg_0\Big(\frac{r}{\sqrt{t}}\Big) + t^{-1}h\Big(\frac{r}{\sqrt{t}}\Big), \tag{2-25}$$

where $g_0(s) = s^{-1}e^{-\frac{s^2}{4}}$ is a solution to

$$g''(s) + \Big(\frac{2}{s} + \frac{s}{2}\Big)g'(s) + g(s) = 0 \tag{2-26}$$

and h solves

$$h''(s) + \Big(\frac{2}{s} + \frac{s}{2}\Big)h'(s) + h(s) = kAg_0(s), \tag{2-27}$$

with $\lim_{s\to\infty} s^\gamma h(s) = A$, and $\lim_{s\to 0^+} sh(s) = d$, so that $\lim_{s\to 0^+}[h(s) - d/s] = 0$. The function h can be described explicitly. Let $g_1(s) = s^{-1}e^{-\frac{s^2}{4}} \int_0^s e^{\frac{z^2}{4}}\,dz$. This function solves (2-26). Since g_1 and g_0 are linearly independent, the variation of parameters formula gives that, for any constants d and b,

$$h(s) = g_0(s)\Big[d - kA\int_0^s zg_1(z)\,dz\Big] + g_1(s)\Big[b + kA\int_0^s zg_0(z)\,dz\Big] \tag{2-28}$$

solves (2-27). In order to have $\lim_{s\to\infty} s^\gamma h(s) = A$, we need $2\big[b + kA\int_0^\infty zg_0(z)\,dz\big] = A$. Furthermore, to have $\lim_{s\to 0^+}[h(s) - d/s] = 0$, we need $b = 0$. Thus we select

$$b = 0, \quad k = \frac{1}{2\int_0^\infty zg_0(z)\,dz}. \tag{2-29}$$

Observe that, up to this moment, the constant d is arbitrary. Nevertheless, recall that we want u_{out} to be a solution to $u_t = \Delta u = u_{rr} + (2/r)u_r$. Multiplying this equation by r, and integrating in $(0, R)$, for some fixed, large R, we get

$$\frac{d}{dt}\left(\int_0^R ru(r,t)\,dr\right) = Ru_r(R,t) + u(R,t),$$

where we use the fact that $\lim_{r\to 0}[ru_r(r,t) + u(r,t)] = 0$. Next, we integrate the above equation in t, from 0 to ∞, and using the fact that $\lim_{t\to\infty}\int_0^R ru(r,t)\,dt = 0$, we get

$$-\int_0^R ru(r,0)\,dr = \int_0^\infty [Ru_r(R,t) + u(R,t)]\,dt. \tag{2-30}$$

Take now $u = u_{\text{out}}$ and compute the right-hand side of (2-30):

$$\begin{aligned}\int_0^\infty &[Ru_r(R,t) + u(R,t)]\,dt \\ &= Ak\int_0^\infty t^{-1}(\log t)\left[\frac{R}{\sqrt{t}}g_0'\left(\frac{R}{\sqrt{t}}\right) + g_0\left(\frac{R}{\sqrt{t}}\right)\right]dt + \int_0^\infty t^{-1}\left[\frac{R}{\sqrt{t}}h'\left(\frac{R}{\sqrt{t}}\right) + h\left(\frac{R}{\sqrt{t}}\right)\right]dt \\ &= \left(4Ak\int_0^\infty s^{-1}[sg_0'(s) + g_0(s)]\,ds\right)\log R + \bar{d} + \left(2\int_0^\infty s^{-1}[sh'(s) + h(s)]\,ds\right),\end{aligned}$$

where $s := R/\sqrt{t}$ and $\bar{d}$ is the constant defined by

$$\bar{d} = -\left(4Ak\int_0^\infty s^{-1}(\log s)[sg_0'(s) + g_0(s)]\,ds\right).$$

We can simplify the expression of the constant in front of $\log R$. Indeed, multiplying (2-26) against s, we get that $(sg'(s) + g + (s^2/2)g)' = 0$. For $g = g_0$, and using the fact that g_0 decays very fast as $s \to \infty$, we get that $sg_0'(s) + g_0(s) = -(s^2/2)g_0(s)$ for any s; thus

$$4Ak\int_0^\infty s^{-1}[sg_0'(s) + g_0(s)]\,ds = Ak\left(-2\int_0^\infty sg_0(s)\,ds\right) = -A$$

since (2-29). On the other hand, the decaying condition $\lim_{r\to\infty} r^2u(r,0) = A$ gives

$$-\int_0^R ru(r,0)\,dr = -A\log R + B(R),$$

with $\lim_{R\to\infty} B(R) = B$, where B is a real constant. Plugging this information in (2-30), we get

$$\bar{d} + \left(2\int_0^\infty s^{-1}[sh'(s) + h(s)]\,ds\right) = B.$$

This last relation defines in a unique way the constant $d > 0$ in the definition of h, (2-28). Indeed, a direct computation gives

$$\int_0^\infty s^{-1}[sh'(s) + h(s)]\,ds = -\frac{d}{2}\left(\int_0^\infty sg_0(s)\,ds\right) + \omega,$$

with

$$\omega = \frac{kA}{2}\int_0^\infty s g_0(s)\left(\int_0^s z g_1(z)\,dz\right)ds + \int_0^s s^{-1}[s g_1' + g_1]\left(kA\int_0^s z g_0(z)\,dz\right)ds,$$

from which we deduce that

$$d = \frac{\bar{d} - 2\omega - B}{\int_0^\infty s g_0(s)\,ds}.$$

With this choice for the function h in (2-25), we get

$$h(s) = \frac{d}{s} - \frac{s}{4}[d + 10kA] + O(s^3) \quad \text{as } s \to 0^+$$

and

$$u_{\text{out}}(r,t) = \frac{t^{-\frac{1}{2}}}{r}[kA(\log t) + d] + t^{-1}\left[-\frac{kA(\log t)}{4} - \frac{d + 10kA}{4}\right]\frac{r}{\sqrt{t}} + O\left((\log t)\frac{r^3}{t^3\sqrt{t}}\right) \tag{2-31}$$

in the region $r < R^{-1}\sqrt{t}$, for some large but fixed R, as $t \to \infty$.

In this case, namely when $\gamma = 2$, we choose μ_0 in (2-10) as

$$\mu_0(t) = \frac{[d + kA(\log t)]^2}{\sqrt{3}}t^{-1}, \tag{2-32}$$

and thanks to this choice, and to the bound (2-11) on λ, we find a constant C so that

$$|u_{\text{in}}(r,t) - u_{\text{out}}(r,t)| \le C\frac{\mu_0^{\frac{1}{2}}}{r}, \quad |\nabla u_{\text{in}}(r,t) - \nabla u_{\text{out}}(r,t)| \le C\frac{\mu_0^{\frac{1}{2}}}{r^2} \tag{2-33}$$

for any $R\mu_0 < r < R^{-1}\sqrt{t}$, for some fixed and large R, and for all t large enough.

Case 3: $\gamma > 2$. In this case, we define u^1_{out} as

$$u^1_{\text{out}}(r,t) = t^{-1}\,d g_0\left(\frac{r}{\sqrt{t}}\right), \quad d = \left(\frac{\int_0^\infty r u_0(r)\,dr}{\int_0^\infty s g_0(s)\,ds}\right),$$

where $g_0(s) = s^{-1}e^{-\frac{s^2}{4}}$ solves (2-26), and $u_0(r)$ is the initial condition for (2-1)–(2-2). Observe that, in a region like $r < R^{-1}\sqrt{t}$, for some large but fixed R, we get

$$u^1_{\text{out}}(r,t) = d\,\frac{t^{-\frac{1}{2}}}{r} - t^{-1}\frac{d}{4}\frac{r}{\sqrt{t}} + t^{-1}O\left(\frac{r^2}{t^{\frac{3}{2}}}\right). \tag{2-34}$$

For a given time t, the function u^1_{out} is decaying very fast as $r \to \infty$. For this reason, we modify u^1_{out} with a function that has the right decay to match the initial condition $u_0(r)$, for r large. Define

$$u_{\text{out}}(r,t) = \eta\left(\frac{r}{t}\right)u^1_{\text{out}}(r,t) + \left(1 - \eta\left(\frac{r}{t}\right)\right)u^2_{\text{out}}(r), \quad \text{with } u^2_{\text{out}}(r) = \frac{A}{r^\gamma}, \tag{2-35}$$

where η is the cut-off function defined in (1-9).

In this case, $\gamma > 2$, we choose μ_0 in (2-10) as

$$\mu_0(t) = \frac{d^2}{\sqrt{3}} t^{-1}. \tag{2-36}$$

With this choice for $\mu_0(t)$, and thanks to (2-11), given any large but fixed number $R > 0$, there exists a constant C so that

$$|u_{\text{in}}(r,t) - u_{\text{out}}(r,t)| \le C \frac{\mu_0^{\frac{1}{2}}}{r}, \quad |\nabla u_{\text{in}}(r,t) - \nabla u_{\text{out}}(r,t)| \le C \frac{\mu_0^{\frac{1}{2}}}{r^2} \tag{2-37}$$

for any $R\mu_0 < r < R^{-1}\sqrt{t}$, and for all t large.

2C. *Construction of the first global approximation and estimate of the error.* Let $r_0 > 0$ be a small and fixed number, and define

$$U_1(r,t) = \eta\Big(\frac{r}{r_0\sqrt{t}}\Big) u_{\text{in}}(r,t) + \Big(1 - \eta\Big(\frac{r}{r_0\sqrt{t}}\Big)\Big) u_{\text{out}}(r,t), \tag{2-38}$$

where η is given by (1-9). For any smooth function $u = u(r,t)$, we define the error function as

$$\mathcal{E}[u](r,t) = \Delta u + u^5 - u_t. \tag{2-39}$$

Our next purpose is to describe

$$\mathcal{E}_1(r,t) = \mathcal{E}[U_1](r,t), \tag{2-40}$$

with U_1 given by (2-38). To this end, we introduce the function $\alpha = \alpha(t)$, $t > t_0$,

$$\alpha(t) = 3^{\frac{1}{4}} \mu_0^{-\frac{1}{2}} (\mu_0 \Lambda)'. \tag{2-41}$$

Since Λ satisfies (2-11), definition (2-41) defines a linear homeomorphism $\mathcal{A} : X_\sharp \to X_\flat$, $\mathcal{A}(\lambda) = \alpha$, where

$$X_\flat = \{\alpha \in C(t_0,\infty) : \|\alpha\|_\flat \text{ is bounded}\}, \tag{2-42}$$

and

$$\|\alpha\|_\flat := \sup_{t > t_0} \mu_0^{-\frac{3}{2}}(t) t [\|\alpha\|_{\infty,[t,t+1]} + |\alpha|_{0,\sigma,[t,t+1]}]. \tag{2-43}$$

Here σ is the number introduced in (2-11). Let us denote by $h_0 : (0,\infty) \to (0,\infty)$ a smooth function with the property that

$$h_0(s) = \begin{cases} 1/s & \text{for } s \to 0, \\ 1/s^3 & \text{for } s \to \infty, \end{cases} \tag{2-44}$$

and define the following norm for any function $f : \mathbb{R}^3 \times (t_0,\infty) \to \mathbb{R}$:

$$\|f\|_* := \sup_{x \in \mathbb{R}^3,\, t > t_0} \mu_0^{-\frac{1}{2}} t^{\frac{3}{2}} h_0^{-1}\Big(\frac{r}{\sqrt{t}}\Big) [\|f\|_{\infty, B(x,1)\times[t,t+1]} + [f]_{0,\sigma,B(x,1)\times[t,t+1]}], \quad r = |x|. \tag{2-45}$$

Here σ is defined in (2-11),

$$\|f\|_{\infty,B(x,1)\times[t,t+1]} = \sup_{y\in B(x,1),\quad s\in[t,t+1]} |f(y,s)| \tag{2-46}$$

and

$$[f]_{0,\sigma,B(x,1)\times[t,t+1]} = \sup_{y_1\neq y_2\in B(x,1),\quad s_1\neq s_2\in[t,t+1]} \frac{|f(y_1,s_1)-f(y_2,s_2)|}{|y_1-y_2|^{2\sigma}+|s_1-s_2|^{\sigma}}. \tag{2-47}$$

The following estimates are valid, and we delay the proof to Appendix B since it is quite technical.

Lemma 2.3. *Assume $\lambda=\lambda(t)$ satisfies (2-11). The error function defined in (2-40) can be described as*

$$\mathcal{E}_1(r,t) = \frac{\alpha(t)}{\mu+r}\eta\left(\frac{r}{r_0\sqrt{t}}\right) + \mathcal{E}_{1,*}[\lambda](r,t), \tag{2-48}$$

where η is the smooth cut-off function defined in (1-9), α is the function defined in (2-41), and r_0 is a given fixed small number. The function $\mathcal{E}_{1,*}[\lambda](r,t)$ depends smoothly on λ. Furthermore, there exists $C>0$ such that

$$\|\mathcal{E}_{1,*}\|_* \le C. \tag{2-49}$$

If the initial time t_0 in problem (2-1) is large enough, there exist $\boldsymbol{c}\in(0,1)$ so that, for any λ_1,λ_2 satisfying (2-11), we have

$$\|\mathcal{E}_{1,*}[\lambda_1]-\mathcal{E}_{1,*}[\lambda_2]\|_{\infty,B(x,1)\times[t,t+1]} \le \boldsymbol{c}\,\mu_0^{\frac12} t^{-\frac32} h_0\left(\frac{r}{\sqrt{t}}\right)\|\lambda_1-\lambda_2\|_\sharp, \tag{2-50}$$

$$[\mathcal{E}_{1,*}[\lambda_1]-\mathcal{E}_{1,*}[\lambda_2]]_{0,\sigma,B(x,1)\times[t,t+1]} \le \boldsymbol{c}\,\mu_0^{\frac12} t^{-\frac32} h_0\left(\frac{r}{\sqrt{t}}\right)\|\lambda_1-\lambda_2\|_\sharp \tag{2-51}$$

for any $r=|x|$ and any t. The definitions of the function h_0 and of the norm $\|\cdot\|_*$ are given respectively in (2-44) and in (2-45). Furthermore the constant $\boldsymbol{c}$ in (2-50) and (2-51) can be made as small as one needs, provided that the initial time t_0 is chosen large enough.

2D. *Construction of the second global approximation and estimate of the new error.* Taking into account the expression of the error function given in (2-48), we introduce a correction function ϕ_0 to partially get rid of the term $\alpha(t)/(\mu+r)$. More precisely, let

$$\bar\alpha(t) = \begin{cases}\alpha(t_0) & \text{for } t<t_0,\\ \alpha(t) & \text{for } t\ge t_0,\end{cases} \tag{2-52}$$

and introduce the function ϕ_0, a solution to

$$\partial_t\phi_0 = \Delta\phi_0 + \frac{\bar\alpha(t)}{\mu+r}\mathbf{1}_{\{r<M\}} \quad \text{in } \mathbb{R}^3\times(0,\infty), \qquad \phi_0(x,t_0-1)=0 \quad \text{in } \mathbb{R}^3, \qquad M^2=t_0. \tag{2-53}$$

Here, for a set K, we mean

$$1_K(x) = \begin{cases}1 & \text{if } x\in K,\\ 0 & \text{if } x\notin K.\end{cases}$$

Duhamel's formula provides an explicit expression for ϕ_0:

$$\phi_0(x,t) = \int_{t_0-1}^{t} \frac{1}{(4\pi(t-s))^{\frac32}} \int_{\mathbb{R}^3} e^{-\frac{|x-y|^2}{4(t-s)}} \frac{\bar\alpha(s)}{\mu+|y|}\mathbf{1}_{\{|y|<M\}}\,dy\,ds. \tag{2-54}$$

Since λ satisfies (2-11), classical parabolic estimates give that ϕ_0 is locally $C^{2+2\sigma,1+\sigma}$, where σ is the Hölder exponent in (2-11). In the interval (t_0,∞), the function ϕ_0 solves

$$\partial_t\phi_0 = \Delta\phi_0 + \frac{\alpha(t)}{\mu+r}\mathbf{1}_{\{r<M\}} \quad \text{in } \mathbb{R}^3\times(t_0,\infty), \tag{2-55}$$

and at time $t=t_0$, the function $\phi_0(x,t_0)$ is radial in x and decays fast as $|x|\to\infty$; that is,

$$|\phi_0(x,t_0)|\le ce^{-a|x|^2} \quad \text{as } |x|\to\infty \tag{2-56}$$

for some positive, fixed constants a and c. Indeed, let $x=\ell e$, with $\|e\|=1$, and assume that $\ell>\max\{1,2M\}$. Thus $|x-y|^2>\ell^2/4$ for any $|y|<M$, and

$$|\phi_0(x,t_0)|\le C|\alpha(t_0)|\left(\int_{t_0-1}^{t_0}\frac{e^{-\frac{\ell^2}{16(t_0-s)}}}{(t_0-s)^{\frac{3}{2}}}\,ds\right)\left(\int_{|y|<M}\frac{dy}{|y|}\right)\le C|\alpha(t_0)|M^2e^{-\frac{\ell^2}{16}}.$$

Taking $\ell\to\infty$, estimate (2-56) thus follows from (2-41).

The second approximation is given by

$$U_2[\lambda](r,t)=U_1(r,t)+\phi_0(r,t), \tag{2-57}$$

where U_1 is in (2-38). Observe that U_2 satisfies the decaying conditions (2-2) at the initial time t_0 as consequence of (2-56). The new error function

$$\mathcal{E}_2[\lambda](r,t)=\mathcal{E}[U_2](r,t)$$

is thus

$$\mathcal{E}_2[\lambda](r,t)=\underbrace{\mathcal{E}_{1,*}+\frac{\alpha(t)}{r}\left(\eta\left(\frac{r}{r_0\sqrt{t}}\right)-\mathbf{1}_{\{r<2M\}}\right)}_{:=\mathcal{E}_{21}}+\underbrace{(U_1+\phi_0)^5-U_1^5}_{\mathcal{E}_{22}}. \tag{2-58}$$

The function $\mathcal{E}_{1,*}$ is defined in (2-48). For a later purpose, it is useful to estimate, in the $\|\cdot\|_*$-norm introduced in (2-45), the function

$$\bar{\mathcal{E}}_2:=\mathcal{E}_{21}+(1-\eta_R(x,t))\mathcal{E}_{22}, \quad \text{where } \eta_R(x,t)=\eta\left(\frac{x}{R\mu_0}\right). \tag{2-59}$$

Here $\eta(s)$ is given by (1-9), while the number R is a large number, whose definition will depend on t_0 but will not depend on t.

We have the validity of the following lemma, whose proof is given in Appendix C.

Lemma 2.4. *Assume $\lambda=\lambda(t)$ satisfies (2-11). The error function defined in (2-58) depends smoothly on λ and it satisfies the following estimates*: *There exists $C>0$ such that*

$$\|\bar{\mathcal{E}}_2\|_*\le C. \tag{2-60}$$

If the initial time t_0 is large enough, there exists small positive number $\boldsymbol{c} \in (0,1)$ such that, for any λ_1, λ_2 satisfying (2-11), we have

$$\|\bar{\mathcal{E}}_2[\lambda_1]-\bar{\mathcal{E}}_2[\lambda_2]\|_{\infty,B(x,1)\times[t,t+1]} \le \boldsymbol{c}\mu_0^{\frac{1}{2}}t^{-\frac{3}{2}}h_0\left(\frac{r}{\sqrt{t}}\right)\|\lambda_1-\lambda_2\|_\sharp, \quad r=|x|, \tag{2-61}$$

$$[\bar{\mathcal{E}}_2[\lambda_1](r,t)-\bar{\mathcal{E}}_2[\lambda_2](r,t)]_{0,\sigma,[t,t+1]} \le \boldsymbol{c}\mu_0^{\frac{1}{2}}t^{-\frac{3}{2}}h_0\left(\frac{r}{\sqrt{t}}\right)\|\lambda_1-\lambda_2\|_\sharp \tag{2-62}$$

for any x and $t > t_0$, provided the initial time t_0 in problem (2-1) is chosen large enough. The definition of the function h_0 is given in (2-44), and the definition of the $\|\cdot\|_*$-norm is given in (2-45).

Remark 2.5. From the proof of the result, we also get that the constant $\boldsymbol{c}$ in (2-61) and (2-62) can be made as small as one needs, provided that the initial time t_0 is chosen large enough.

3. The inner-outer gluing

Recall that our ultimate purpose is to construct a global unbounded solution u to (2-1)–(2-2) of the form

$$u = U_2[\lambda](r,t)+\tilde{\phi}, \quad t>t_0, \tag{3-1}$$

where U_2 is defined in (2-57), while $\tilde{\phi}(x,t)$ is a smaller perturbation. The rest of the paper is thus devoted to finding $\tilde{\phi}(x,t)$. The construction of $\tilde{\phi}(x,t)$ is done by means of an *inner-outer gluing* procedure. This procedure consists in writing

$$\tilde{\phi}(x,t)=\psi(x,t)+\phi^{\text{in}}(x,t), \quad \text{where } \phi^{\text{in}}(x,t):=\eta_R(x,t)\hat{\phi}(x,t), \tag{3-2}$$

with

$$\hat{\phi}(x,t):=\mu_0^{-\frac{1}{2}}\phi\left(\frac{x}{\mu_0},t\right), \quad \eta_R(x,t)=\eta\left(\frac{x}{R\mu_0}\right), \tag{3-3}$$

where $\eta(s)$ is given in (1-9).

In terms of $\tilde{\phi}$, problem (2-1)–(2-2) reads as

$$\partial_t\tilde{\phi}=\Delta\tilde{\phi}+5U_2^4\tilde{\phi}+N(\tilde{\phi})+\mathcal{E}_2 \quad \text{in } \mathbb{R}^3\times[t_0,\infty), \tag{3-4}$$

where $\mathcal{E}_2$ is defined in (2-58) and

$$N(\tilde{\phi})=(U_2+\tilde{\phi})^5-U_2^5-5U_2^4\tilde{\phi}.$$

Recalling that $w_\mu=\mu^{-\frac{1}{2}}w(r/\mu)$, we let

$$V[\lambda](r,t)=5(U_2^4-w_\mu^4)\eta_R+5U_2^4(1-\eta_R) \tag{3-5}$$

and write $5U_2^4=5w_\mu^4\eta_R+V[\lambda](r,t)$. A main observation we make is that $\tilde{\phi}$ solves problem (3-4) if the tuple (ψ,ϕ) solves the following coupled system of nonlinear equations:

$$\partial_t\psi=\Delta\psi+V[\lambda]\psi+[2\nabla\eta_R\nabla_x\hat{\phi}+\hat{\phi}(\Delta_x-\partial_t)\eta_R]+N[\lambda](\tilde{\phi})+\mathcal{E}_{21}+\mathcal{E}_{22}(1-\eta_R) \quad \text{in } \mathbb{R}^3\times[t_0,\infty), \tag{3-6}$$

and

$$\partial_t\hat{\phi}=\Delta\hat{\phi}+5w_\mu^4\hat{\phi}+5w_\mu^4\psi+\mathcal{E}_{22} \quad \text{in } B_{2R\mu_0}(0)\times[t_0,\infty). \tag{3-7}$$

We refer to (2-58) for the definition of $\mathcal{E}_{21}$ and $\mathcal{E}_{22}$. In terms of ϕ, see (3-3), (3-7) becomes

$$\mu_0^2\partial_t\phi = \Delta_y\phi + 5w^4\phi + \mu_0^{\frac{5}{2}}\mathcal{E}_{22}(\mu_0 y,t) + 5\frac{\mu_0^{\frac{1}{2}}}{(1+\Lambda)^4}w^4\Big(\frac{y}{(1+\Lambda)^2}\Big)\psi(\mu_0 y,t) + B[\phi] + B^0[\phi] \quad \text{in } B_{2R}(0)\times[t_0,\infty), \tag{3-8}$$

where

$$B[\phi] := \mu_0(\partial_t\mu_0)\Big(\frac{\phi}{2} + y\cdot\nabla_y\phi\Big), \tag{3-9}$$

$$B^0[\phi] := 5\Big[w^4\Big(\frac{y}{(1+\Lambda)^2}\Big) - w^4(y)\Big]\phi + 5\Big(\frac{1-(1+\Lambda)^4}{(1+\Lambda)^4}\Big)w^4\Big(\frac{y}{(1+\Lambda)^2}\Big)\phi. \tag{3-10}$$

We call (3-6) the *outer problem* and (3-8) the *inner problem(s)* .

We next describe precisely our strategy to solve (3-6)–(3-8). For given parameter λ satisfying (2-11), and function ϕ fixed in a suitable range, we first solve for ψ the outer problem (3-6), in the form of a (nonlocal) nonlinear operator $\psi = \Psi(\lambda,\phi)$. This is done in full detail in Section 4.

We then replace this ψ in (3-8). At this point we consider the change of variable

$$t = t(\tau), \qquad \frac{dt}{d\tau} = \mu_0^2(t),$$

which reduces (3-8) to

$$\partial_\tau\phi = \Delta_y\phi + 5w^4\phi + H[\psi,\lambda,\phi](y,t(\tau)), \quad y\in B_{2R}(0), \quad \tau\ge\tau_0, \tag{3-11}$$

where τ_0 is such that $t(\tau_0)=t_0$, and

$$H[\psi,\lambda,\phi](y,t(\tau)) = \mu_0^{\frac{5}{2}}\mathcal{E}_{22}(\mu_0 y,t) + 5\frac{\mu_0^{\frac{1}{2}}}{(1+\Lambda)^4}w^4\Big(\frac{y}{(1+\Lambda)^2}\Big)\psi(\mu_0 y,t) + B[\phi] + B^0[\phi]. \tag{3-12}$$

Next step is to construct a solution ϕ to problem (3-11). We can do this for functions ϕ which furthermore satisfy

$$\phi(y,\tau_0) = e_0 Z(y), \quad y\in B_{2R}(0), \tag{3-13}$$

for some constant e_0. Here Z is the positive radially symmetric bounded eigenfunction associated to the only negative eigenvalue λ_0 to the problem

$$L_0(\phi) + \lambda\phi = 0, \quad \phi\in L^\infty(\mathbb{R}^3). \tag{3-14}$$

Here L_0 is the linear operator around the standard bubble w in $\mathbb{R}^3$. We refer to (2-6) for the definition of L_0. Furthermore, it is known that λ_0 is simple and Z decays like

$$Z(y) \sim |y|^{-1}e^{-\sqrt{|\lambda_0|}|y|} \quad \text{as } |y|\to\infty.$$

To be more precise, we prove that problem (3-11)–(3-13) is solvable in ϕ, provided that in addition the parameter λ is chosen so that $H[\psi,\lambda,\phi](y,t(\tau))$ satisfies the orthogonality condition

$$\int_{B_{2R}} H[\psi,\lambda,\phi](y,t(\tau))Z_0(y)\,dy = 0 \quad \text{for all } t > t_0. \tag{3-15}$$

We recall that $Z_0(y)$, defined in (2-7), is the only bounded radial element in the kernel of the linear elliptic operator L_0.

Equation (3-15) becomes a nonlinear, nonlocal problem in λ for any fixed ϕ. We attack this problem in Sections 5, 6, and 7. In Section 5, we get the precise form of (3-15) as a nonlocal nonlinear operator in λ. The principal part of the operator in λ defined by (3-15) is a linear nonlocal operator which turns out to be a perturbation of the $\frac{1}{2}$-Caputo derivative. We refer to [Caputo 1967] for the original definition of Caputo derivatives. In Section 6 we develop an invertibility theory for such a linear operator. In Section 7 we fully solve (3-15) in λ, by means of a Banach fixed-point argument. The solution $\lambda = \lambda[\phi]$ is a nonlinear operator in ϕ, and we also describe the Lipschitz dependence of λ with respect to ϕ, which is a key property for our final argument.

At this point, one realizes that a central point of our complete proof is to design a linear theory that allows us to solve in ϕ problem (3-11)–(3-13). For this purpose, we shall construct a solution to an initial value problem of the form

$$\phi_\tau = \Delta\phi + 5w^4\phi + h(y,\tau) \quad \text{in } B_{2R}\times(\tau_0,\infty), \qquad \phi(y,\tau_0) = e_0 Z(y) \quad \text{in } B_{2R}. \tag{3-16}$$

And then we solve problem (3-11)–(3-13) by means of a contraction mapping argument.

Let a be a fixed number with $a \in (0,2)$, and let $\nu > 0$ so that, for t large,

$$\tau^{-\nu} \sim \mu_0^{\frac{3}{2}} t^{-1} \quad \text{if } \gamma \neq 2 \qquad \text{and} \qquad \tau^{-\nu} \sim \mu_0^{\frac{3}{2}} t^{-1+\nu'} \quad \text{if } \gamma = 2$$

for some $\nu' > 0$ that can be fixed arbitrarily small. We solve (3-16) for functions h with $\|h\|_{\nu,2+a}$-norm bounded, where

$$\|h\|_{\nu,2+a} := \sup_{\tau>\tau_0,\, y\in\mathbb{R}^3} \tau^\nu (1+|y|^{2+a})[\|h\|_{\infty,B(y,1)\times[\tau,\tau+1]} + [h]_{0,\sigma,B(y,1)\times[\tau,\tau+1]}], \tag{3-17}$$

and we construct solutions ϕ in the class of functions with $\|\phi\|_{\nu,a}$-norm bounded, where

$$\begin{aligned}\|\phi\|_{\nu,a} := {} & \sup_{\tau>\tau_0,\, y\in\mathbb{R}^3} \tau^\tau (1+|y|^{a})[\|\phi\|_{\infty,B(y,1)\times[\tau,\tau+1]} + [\phi]_{0,\sigma,B(y,1)\times[\tau,\tau+1]}] \\ & + \sup_{\tau>\tau_0,\, y\in\mathbb{R}^3} \tau^\nu (1+|y|^{1+a})[\|\nabla\phi\|_{\infty,B(y,1)\times[\tau,\tau+1]} + [\nabla\phi]_{0,\sigma,B(y,1)\times[\tau,\tau+1]}].\end{aligned} \tag{3-18}$$

We have the validity of the following result:

Proposition 3.1. *Let ν, a be given positive numbers with $0 < a < 2$. Then, for all sufficiently large $R > 0$ and function $h = h(y,\tau)$, with $h(y,\tau) = h(|y|,\tau)$ and $\|h\|_{\nu,2+a} < +\infty$, that satisfies*

$$\int_{B_{2R}} h(y,\tau) Z_0(y)\, dy = 0 \quad \textit{for all } \tau \in (\tau_0,\infty) \tag{3-19}$$

there exist $\phi \in C^{2+2\sigma,1+\sigma}$-loc., which is radial in y, and e_0 which solve problem (3-16). Moreover, $\phi = \phi[h]$ and $e_0 = e_0[h]$ define linear operators of h that satisfy the estimates

$$|\phi(y,\tau)| \le C\tau^{-\nu} \frac{R^{4-a}}{1+|y|^3} \|h\|_{\nu,2+a}, \quad |\nabla_y\phi(y,\tau)| \le C\tau^{-\nu} \frac{R^{4-a}}{1+|y|^4} \|h\|_{\nu,2+a}, \tag{3-20}$$

and

$$|e_0[h]| \leq C\|h\|_{\nu,2+a}$$

for some fixed constant C.

We postpone the proof of this proposition to Section 9. Section 8 is devoted to solving problem (3-11)–(3-13) and this concludes the proof of Theorem 1.1.

4. Solving the outer problem

The aim of this section is to solve the *outer problem* (3-6) for given parameter λ satisfying (2-11), and for given small functions ϕ, in the form of a nonlinear nonlocal operator

$$\psi(x,t) = \Psi[\lambda,\phi](x,t).$$

We recall that $\phi^{\text{in}}(x,t) = \eta_R(x,t)\hat{\phi}(x,t)$, with

$$\hat{\phi}(x,t) := \mu_0^{-\frac{1}{2}}\phi\left(\frac{x}{\mu_0},t\right) \quad \text{and} \quad \eta_R(x,t) = \eta\left(\frac{x}{R\mu_0}\right).$$

Here $\eta(s)$ is defined in (1-9), and R is a sufficiently large number, independent of t. We assume that

$$\|\phi\|_{\nu,a} \text{ is bounded.} \tag{4-1}$$

Let $\varphi_0 : (0,\infty) \to (0,\infty)$ be a smooth and bounded given function with the property that

$$\varphi_0(s) = \begin{cases} s & \text{for } s \to 0^+, \\ 1/s^3 & \text{for } s \to \infty. \end{cases} \tag{4-2}$$

We introduce the following L^∞-weighted norms for functions $f = f(r,t)$:

$$\|f\|_{**} := \|f\|_1 + \|Df\|_2, \tag{4-3}$$

$$\|f\|_1 := \sup_{x\in\mathbb{R}^3,\, t>t_0} \mu_0^{-\frac{1}{2}} t^{\frac{1}{2}} \varphi_0^{-1}\left(\frac{r}{\sqrt{t}}\right)[\|f\|_{\infty,B(x,1)\times[t,t+1]} + [f]_{0,\sigma,B(x,1)\times[t,t+1]}], \qquad r = |x|, \tag{4-4}$$

$$\|f\|_2 := \sup_{x\in\mathbb{R}^3,\, t>t_0} \mu_0^{-\frac{1}{2}} t (\varphi_0')^{-1}\left(\frac{r}{\sqrt{t}}\right)[\|f\|_{\infty,B(x,1)\times[t,t+1]} + [f]_{0,\sigma,B(x,1)\times[t,t+1]}], \quad r = |x|. \tag{4-5}$$

Refer to (2-46) and (2-47) for the definitions of $\|f\|_{\infty,B(x,1)\times[t,t+1]}$ and $[f]_{0,\sigma,B(x,1)\times[t,t+1]}$.

Proposition 4.1. *Assume that λ satisfies (2-11), and that the function ϕ satisfies the bound (4-1). Let $\psi_0 \in C^2(\mathbb{R}^3)$, radially symmetric so that*

$$|y||\psi_0(y)| + |y||\nabla\psi_0(y)| \leq t_0^{-a} e^{-b|y|} \tag{4-6}$$

for some positive constants a and b. There exists t_0 large so that problem (3-6) has a unique solution $\psi = \Psi[\lambda,\phi]$ so that

$$\psi(r,t_0) = \psi_0(r), \quad \|\psi\|_1 + \|D\psi\|_2 \leq C. \tag{4-7}$$

Proof. Let f be a given function with $\|f\|_*$-norm bounded. Classical parabolic estimates give that any solution to $\partial_t \psi = \Delta\psi + f$ is locally $C^{2+2\sigma,1+\sigma}$. Furthermore, a consequence of Lemma A.1 is that the function $\bar{\varphi}_0(r,t) = \mu_0^{\frac{1}{2}} t^{-\frac{1}{2}} \varphi_0(r/\sqrt{t})$ is a positive supersolution for $\partial_t\psi \geq \Delta\psi + f(r,t)$. Observe also that $\bar{\varphi}_0(r,t_0) \geq \psi_0(r)$. Combining these facts with the maximum principle, we see that, for a function f with $\|f\|_*$-norm bounded, the unique solution to $\partial_t\psi = \Delta\psi + f$, with $\psi(r,t_0) = \psi_0$, has $\|\psi\|_{**}$-norm bounded. We claim that a possibly large multiple of $\bar{\varphi}_0$ works as a supersolution also for the problem

$$\partial_t\psi \geq \Delta\psi + V(r,t)\psi + f(r,t). \tag{4-8}$$

Indeed, recalling the definition of V in (3-5), we write

$$V = V_1 + V_2, \quad V_1 = 5(U_2^4 - w_\mu^4)\eta_R, \quad V_2 = 5U_2^4(1-\eta_R).$$

In the region where $\eta_R \neq 0$, namely when $r < 2R\mu_0$, we expand in Taylor the function V_1 and we find $s^* \in (0,1)$ so that

$$V_1(r,t) = 20\big(w_\mu + s^*(\mu_0'\Psi_1(r,t) + \phi_0(r,t))\big)^3[\mu_0'\Psi_1(r,t) + \phi_0(r,t))]\eta_R.$$

From here, we see that, in this region, $|V_1(r,t)| \lesssim Rt^{-1}\eta_R$, so that

$$|V_1(r,t)\psi_0(r,t)| \lesssim \mu_0^{\frac{1}{2}} t^{-\frac{3}{2}} h_0\left(\frac{r}{\sqrt{t}}\right). \tag{4-9}$$

Let us now consider V_2. This function is not zero only when $r > R\mu_0$, and in this region we have that $|V_2(r,t)| \lesssim (\mu_0^2/r^4)(1-\eta_R)$, so that

$$|V_2(r,t)\psi_0(r,t)| \lesssim \frac{\mu^2}{r^4}\mu_0^{\frac{1}{2}} t^{-\frac{1}{2}}\varphi_0\left(\frac{r}{\sqrt{t}}\right)(1-\eta_R) \lesssim R^{-2}\mu_0^{\frac{1}{2}} t^{-\frac{3}{2}} h_0\left(\frac{r}{\sqrt{t}}\right). \tag{4-10}$$

Choosing R large, but independent of t, we thus find that a multiple of $\bar{\varphi}_0$ is a supersolution for (4-8).

We call $T_o : (f,\psi_0) \to \psi$ the linear operator that to any f with $\|f\|_*$-norm bounded and any initial condition ψ_0 satisfying (4-6) associates the unique solution to

$$\partial_t\psi = \Delta\psi + V[\lambda](r,t)\psi + f(r,t), \quad \psi(r,t_0) = \psi_0(r), \tag{4-11}$$

which has bounded $\|\psi\|_{**}$-norm. Define $\bar{\psi} = T_o(0,\psi_0)$. We observe that $\psi + \bar{\psi}$ is a solution to (3-6) if ψ is a fixed point for the operator

$$\mathcal{A}_o(\psi) = T_o\big([2\nabla\eta_R\nabla_x\hat{\phi} + \hat{\phi}(\Delta_x - \partial_t)\eta_R] + N[\lambda](\tilde{\phi} + \bar{\psi}) + \mathcal{E}_{21} + \mathcal{E}_{22}(1-\eta_R)\big). \tag{4-12}$$

We shall show the existence and uniqueness of such a fixed point as consequence of the contraction mapping theorem. We perform a fixed-point argument in the set of functions ψ in

$$B_o = \{\psi \in L^\infty : \|\psi\|_{**} < r\} \tag{4-13}$$

for some $r > 0$.

From Lemma 2.3 we have that there exists a constant c_1 so that

$$\|\mathcal{E}_{21} + \mathcal{E}_{22}(1-\eta_R)\|_* \leq c_1. \tag{4-14}$$

We now claim that there exists a constant c_2 such that, if the parameter λ satisfies (2-11), and if the function ϕ satisfies the bound (4-1), then

$$\|2\nabla\eta_R\nabla_x\hat{\phi}+\hat{\phi}(\Delta_x-\partial_t)\eta_R\|_*+\|N(\tilde{\phi}+\bar{\psi})\|_*\le c_2. \tag{4-15}$$

Furthermore, we claim that there exists a constant $\boldsymbol{c}\in(0,1)$ so that, for any $\psi_1,\psi_2\in B_0$,

$$\|\mathcal{A}_o(\psi_1)-\mathcal{A}_o(\psi_2)\|_{**}\le \boldsymbol{c}\|\psi_1-\psi_2\|_{**}. \tag{4-16}$$

If we assume, for the moment, the validity of (4-14), (4-15) and (4-16), we get the existence of a fixed point for problem (4-12) in the set (4-13), provided r is chosen large enough.

Proof of (4-15): We start with the estimate of the first term in (4-15). Since we assume the validity of the bound (4-1) on ϕ, we write

$$|\hat{\phi}\Delta_x\eta_R|\lesssim\frac{|\eta''(|x|/R\mu_0)|}{R^2\mu_0^2}|\hat{\phi}|\lesssim\frac{|\eta''(|x|/R\mu_0)|}{R^2\mu_0^2}\frac{\mu_0^{\frac{3}{2}}t^{-1}}{(1+|x/\mu_0|^a)}\|\phi\|_{\nu,a};$$

see (3-18) for the notation $\|\phi\|_{\nu,a}$. Thus, we get

$$\begin{aligned}|\hat{\phi}\Delta_x\eta_R|&\lesssim\frac{|\eta''(|x|/R\mu_0)|}{R^{2+a}}\mu_0^{-\frac{1}{2}}t^{-1}\frac{r}{\sqrt{t}}h_0\Big(\frac{r}{\sqrt{t}}\Big)\|\phi\|_{\nu,a}\\&\lesssim\frac{|\eta''(|x|/R\mu_0)|}{R^{1+a}}\mu_0^{\frac{1}{2}}t^{-\frac{3}{2}}h_0\Big(\frac{r}{\sqrt{t}}\Big)\|\phi\|_{\nu,a}\lesssim\mu_0^{\frac{1}{2}}t^{-\frac{3}{2}}h_0\Big(\frac{r}{\sqrt{t}}\Big)\frac{\|\phi\|_{\nu,a}}{R^{1+a}}.\end{aligned}$$

Arguing similarly, we get

$$|\hat{\phi}\partial_x\eta_R|\lesssim\mu_0^{\frac{1}{2}}t^{-\frac{3}{2}}h_0\Big(\frac{r}{\sqrt{t}}\Big)\frac{\|\phi\|_{\nu,a}}{R^{1+a}}\quad\text{and}\quad|\nabla\hat{\phi}\nabla\eta_R|\lesssim\mu_0^{\frac{1}{2}}t^{-\frac{3}{2}}h_0\Big(\frac{r}{\sqrt{t}}\Big)\frac{\|\phi\|_{\nu,a}}{R^{1+a}},$$

which proves the L^∞ bound in the first estimate in (4-15). To check the Hölder bound for this term, we focus the analysis on the term $g(x,t):=\hat{\phi}\Delta_x\eta_R$. The other terms can be treated in a similar way. We write

$$\begin{aligned}&\frac{|g(x_1,t_1)-g(x_2,t_2)|}{|x_1-x_2|^{2\sigma}+|t_1-t_2|^\sigma}\\&\qquad=|\Delta_x\eta_R(x_1,t_1)|\frac{|\hat{\phi}(x_1,t_1)-\hat{\phi}(x_2,t_2)|}{|x_1-x_2|^{2\sigma}+|t_1-t_2|^\sigma}+|\hat{\phi}(x_2,t_2)|\frac{|\Delta_x\eta_R(x_1,t_1)-\Delta_x\eta_R(x_2,t_2)|}{|x_1-x_2|^{2\sigma}+|t_1-t_2|^\sigma}.\end{aligned}$$

In order to control the first term, we use the definition in (3-18) of $\|\phi\|_{\nu,a}$ and we argue as before. The second term can be easily treated using the L^∞-bound on $\hat{\phi}$ and the smoothness of the function $\Delta_x\eta_R$. This completes the analysis of the first estimate in (4-15).

We continue with the proof of the second estimate in (4-15). We recall that

$$N(\tilde{\phi})=(U_2+\tilde{\phi})^5-U_2^5-5U_2^4\tilde{\phi}.$$

It is convenient to estimate this function in three different regions: where $r<\overline{M}^{-1}\mu_0$, where $\overline{M}^{-1}\mu_0<r<\overline{M}\sqrt{t}$ and where $r>\overline{M}\sqrt{t}$, with $\overline{M}$ a large positive number.

From the definition of U_2 in (2-57), we see that, if $r < \overline{M}^{-1}\mu_0$, then

$$|N(\tilde{\phi})| \lesssim \mu_0^{-\frac{3}{2}}|\tilde{\phi}|^2 \lesssim \mu_0^{-\frac{3}{2}}[|\psi|^2 + |\eta_R\hat{\phi}|^2].$$

We recall that

$$|\psi| \lesssim \|\psi\|_{**}\mu_0^{\frac{1}{2}}t^{-\frac{1}{2}}\varphi_0\Big(\frac{r}{\sqrt{t}}\Big), \quad |\eta_R\hat{\phi}| \lesssim \mu_0^{\frac{3}{2}}t^{-1}|\eta_R|\|\phi\|_{\mu,a}, \tag{4-17}$$

so that we get, for $r < \overline{M}^{-1}\mu_0$,

$$|N(\tilde{\phi}+\bar{\psi})| \lesssim \mu_0^2 t^{-1}[\|\psi+\bar{\psi}\|_{**}^2 + \|\phi\|_{\mu,a}^2]\Big(\mu_0^{\frac{1}{2}}t^{-\frac{3}{2}}h_0\Big(\frac{r}{\sqrt{t}}\Big)\Big). \tag{4-18}$$

Let us now consider the region $\overline{M}^{-1}\mu_0 < r < \overline{M}\sqrt{t}$. Here, after a Taylor expansion, we get

$$|N(\tilde{\phi}+\bar{\psi})| \lesssim w_\mu^3[|\psi+\bar{\psi}|^2 + |\eta_R\hat{\phi}|^2] \lesssim \frac{\mu_0^{\frac{3}{2}}}{r^3}[|\psi|^2 + |\eta_R\hat{\phi}|^2].$$

Using again (4-17), we obtain, for $\overline{M}^{-1}\mu_0 < r < \overline{M}\sqrt{t}$,

$$|N(\tilde{\phi}+\bar{\psi})| \lesssim \mu_0^2 t^{-1}[\|\psi+\bar{\psi}\|_{**}^2 + \|\phi\|_{\mu,a}^2]\Big(\mu_0^{\frac{1}{2}}t^{-\frac{3}{2}}h_0\Big(\frac{r}{\sqrt{t}}\Big)\Big). \tag{4-19}$$

Let us now consider $r > \overline{M}\sqrt{t}$. Observe that in this region $\eta_R = 0$, $|(\psi+\bar{\psi})(r,t)| \lesssim \mu_0^{\frac{1}{2}}t^{-\frac{1}{2}}\varphi_0(r/\sqrt{t})$ and, from (C-3), also $|U_2(r,t)| \lesssim \mu_0/r$. Thus we have

$$|N(\tilde{\phi}+\bar{\psi})| \lesssim \Big(\frac{\mu_0}{r}\Big)^5 \lesssim \mu_0^{\frac{9}{2}}t^{-\frac{1}{2}}\Big(\mu_0^{\frac{1}{2}}t^{-\frac{3}{2}}h_0\Big(\frac{r}{\sqrt{t}}\Big)\Big). \tag{4-20}$$

From (4-18)–(4-20), we get the L^∞ bound for the second estimate in (4-15).

Proof of (4-16): For any $\psi_1, \psi_2 \in B_o$, we have

$$\mathcal{A}_o(\psi_1) - \mathcal{A}_o(\psi_2) = T_0(N(\psi_1+\bar{\psi}+\phi^{\text{in}}) - N(\psi_2+\bar{\psi}+\phi^{\text{in}}));$$

thus

$$\|\mathcal{A}_o(\psi_1) - \mathcal{A}_o(\psi_2)\|_{**} \le C\|N(\psi_1+\bar{\psi}+\phi^{\text{in}}) - N(\psi_2+\bar{\psi}+\phi^{\text{in}})\|_*.$$

We write

$$\begin{aligned} &N(\psi_1+\phi^{\text{in}}) - N(\psi_2+\phi^{\text{in}}) \\ &\quad = (U_2+\psi_1+g)^5 - (U_2+\psi_2+g)^5 - 5U_2^4(\psi_1-\psi_2) \\ &\quad = \underbrace{(U_2+\psi_1+g)^5 - (U_2+\psi_2+g)^5 - 5(U_2+g)^4(\psi_1-\psi_2)}_{:=N_1} + \underbrace{5[(U_2+g)^4 - U_2^4](\psi_1-\psi_2)}_{:=N_2}, \end{aligned}$$

where $g := \phi^{\text{in}} + \bar{\psi}$. In the region where $r < \overline{M}\sqrt{t}$, we have

$$|N_1(x,t)| \lesssim w_\mu^3|\psi_1-\psi_2|^2,$$

which yields

$$|N_1(x,t)| \lesssim \mu_0^2 t^{-1}[\|\psi_1-\psi_2\|_{**}^2]\Big(\mu_0^{\frac{1}{2}}t^{-\frac{3}{2}}h_0\Big(\frac{r}{\sqrt{t}}\Big)\Big),$$

while N_2 can be estimated as

$$|N_2(x,t)| \lesssim [\mu_0^2 t^{-1}\|\bar\psi\|_{**} + \mu_0^2 t^{-1}\|\phi^{\text{in}}\|_{\nu,a}]\|\psi_1-\psi_2\|_{**}\left(\mu_0^{\frac12} t^{-\frac32} h_0\left(\frac{r}{\sqrt t}\right)\right).$$

On the other hand, if $r > \overline{M}\sqrt{t}$, we have $\phi^{\text{in}} \equiv 0$, so that

$$|N_2(x,t)| \lesssim \mu_0^2\|\bar\psi\|_{**}\|\psi_1-\psi_2\|_{**}\left(\mu_0^{\frac12} t^{-\frac32} h_0\left(\frac{r}{\sqrt t}\right)\right).$$

On the other hand N_1 can be estimated as follows:

$$|N_1(x,t)| \lesssim |\psi_1-\psi_2|^5, \quad \text{from which } |N_1(x,t)| \lesssim \mu_0^2\mu_0^{\frac12} t^{-\frac12} h_0\left(\frac{r}{\sqrt t}\right)\|\psi_1-\psi_2\|_{**}.$$

In summary, we get

$$\|N(\psi_1+\phi^{\text{in}}+\bar\psi) - N(\psi_2+\phi^{\text{in}}+\bar\psi)\|_{*,\beta} \le C\mu_0^2\|\psi_1-\psi_2\|_{**},$$

where $C = \max\{\|\psi_1-\psi_2\|_{**}, \|\phi^{\text{in}}\|_{\nu,a}\}$. Thus we get the validity of (4-16) provided that t_0 is large enough. □

Remark 4.2. Proposition 4.1 defines the solution to problem (3-6) as a function of the initial condition ψ_0, in the form of an operator $\psi = \bar\Psi[\psi_0]$, from a small neighborhood of 0 in the Banach space $L^\infty(\Omega)$ equipped with the norm

$$\sup_{y\in\mathbb{R}^3} [|y||e^{b|y|}\psi_0(y)| + |y||e^{b|y|}\nabla\psi_0(y)|] \tag{4-21}$$

into the Banach space of functions $\psi \in L^\infty(\Omega)$ equipped with the norm $\|\psi\|_{**}$, defined in (4-3). A closer look to the proof of Proposition 4.1, and the implicit function theorem give that $\psi_0 \to \bar\Psi[\psi_0]$ is a diffeomorphism, and that

$$\|\bar\Psi[\psi_0^1] - \bar\Psi[\psi_0^2]\|_{**} \le c\Big[\sup_{y\in\mathbb{R}^3}\big||y|e^{b|y|}[\psi_0^1-\psi_0^2]\big| + \sup_{y\in\mathbb{R}^3}\big||y|e^{b|y|}[\nabla\psi_0^1-\nabla\psi_0^2]\big|\Big]$$

for some positive constant c.

Proposition 4.3. *Assume the validity of the assumptions of Proposition 4.1. Then the function $\psi = \Psi(\lambda,\phi)$ depends smoothly on λ and ϕ, and we have the validity of the following estimates: for any initial time t_0 in problem (2-1) sufficiently large, and any sufficiently large radius R in the cut-off function η_R introduced in (3-2), there exists $\boldsymbol{c}$ such that, given λ_1, λ_2 satisfying (2-11), one has*

$$\|\Psi[\lambda_1,\phi] - \Psi[\lambda_2,\phi]\|_{**} \le \boldsymbol{c}\|\lambda_1-\lambda_2\|_\sharp \tag{4-22}$$

for any ϕ satisfying (4-1). Moreover, given ϕ_1, ϕ_2 satisfying (4-1), one has

$$\|\Psi[\lambda,\phi_1] - \Psi[\lambda,\phi_2]\|_{**} \le \boldsymbol{c}\|\phi_1-\phi_2\|_{\nu,a} \tag{4-23}$$

for any λ satisfying (2-11).

Proof. Fix ϕ and define $\bar\psi = \psi[\lambda_1,\phi]-\psi[\lambda_2,\phi]$ for λ_1 and λ_2 satisfying (2-11). Then $\bar\psi$ solves

$$\partial_t\bar\psi = \Delta\bar\psi + (V[\lambda_1]+N'[\lambda])(\bar\psi)+F, \quad \mathbb{R}^3\times(t_0,\infty), \quad \bar\psi(r,t_0)=0,$$

for $\lambda = s\lambda_1+(1-s)\lambda_2$, $s\in(0,1)$, where

$$F=\mathcal{E}_{21}[\lambda_1]-\mathcal{E}_{21}[\lambda_2]+(1-\eta_R)[\mathcal{E}_{22}[\lambda_1]-\mathcal{E}_{22}[\lambda_2]]+[V[\lambda_1]-V[\lambda_2]]\psi_2+[N[\lambda_1]-N[\lambda_1]](\psi_2+\phi^{\text{in}}),$$

where $\psi_j=\psi[\lambda_j,\phi]$, $j=1,2$. From Lemma 2.3 and estimates (2-61)–(2-62), we get

$$\begin{aligned}\|\mathcal{E}_{21}[\lambda_1]-\mathcal{E}_{21}[\lambda_2]\|_* &\le \boldsymbol{c}\|\lambda_1-\lambda_2\|_\sharp,\\ \|(1-\eta_R)[\mathcal{E}_{22}[\lambda_1]-\mathcal{E}_{22}[\lambda_2]]\|_* &\le \boldsymbol{c}\|\lambda_1-\lambda_2\|_\sharp,\end{aligned}$$

provided t_0 is large enough. One also checks that, for some $\boldsymbol{c}\in(0,1)$,

$$\|[V[\lambda_1]-V[\lambda_2]]\psi_2\|_*\le \boldsymbol{c}\|\lambda_1-\lambda_2\|_\sharp, \quad \|[N[\lambda_1]-N[\lambda_2]](\psi_2+\phi^{\text{in}})\|_*\le \boldsymbol{c}\|\lambda_1-\lambda_2\|_\sharp.$$

The constant c_1 can be made arbitrarily small provided t_0 is large. Arguing as in (4-9) and (4-10), one can show that a certain multiple of the function $\|\lambda_1-\lambda_2\|_\sharp\bar\varphi_0(r,t)$, where $\bar\varphi_0=\mu_0^{\frac12}t^{-\frac12}\varphi_0(r/\sqrt{t})$, serves as supersolution for $\bar\psi$. This proves (4-22).

Let us now fix λ, and take ϕ_1,ϕ_2 satisfying (4-1). Define $\phi_j^{\text{in}}=\eta_R\hat\phi_j$ and $\hat\phi_j(x,t)=\mu_0^{-\frac12}\phi_j(x/\mu_0,t)$, for $j=1,2$, as natural. Let $\bar\psi=\psi(\lambda,\phi_1)-\psi(\lambda,\phi_2)$. We have $\bar\psi(r,t_0)=0$ and

$$\begin{aligned}\partial_t\bar\psi = \Delta\bar\psi + V[\lambda]\bar\psi + (\psi_1+\phi_1^{\text{in}})^5-(\psi_2+\phi_1^{\text{in}})^5+[2\nabla\eta_R\nabla_x(\hat\phi_1-\hat\phi_2)+(\hat\phi_1-\hat\phi_2)(\Delta_x-\partial_t)\eta_R]\\ +(\psi_2+\phi_1^{\text{in}})^5-(\psi_2+\phi_2^{\text{in}})^5-5U_2^4(\phi_1^{\text{in}}-\phi_2^{\text{in}}).\end{aligned}$$

Arguing as in (4-6)–(4-21), we get

$$\begin{aligned}\big|[2\nabla\eta_R\nabla_x(\hat\phi_1-\hat\phi_2)+(\hat\phi_1-\hat\phi_2)(\Delta_x-\partial_t)\eta_R]\big| &\le \mu_0^{\frac12}t^{-\frac32}h_0\Big(\frac{r}{\sqrt t}\Big)\frac{\|\phi_1-\phi_2\|_{\nu,a}}{R^{1+a}}\\ &\le \boldsymbol{c}\mu_0^{\frac12}t^{-\frac32}h_0\Big(\frac{r}{\sqrt t}\Big)\|\phi_1-\phi_2\|_{\nu,a}\end{aligned}$$

and also

$$\begin{aligned}|(\psi_2+\phi_1^{\text{in}})^5-(\psi_2+\phi_2^{\text{in}})^5-5U_2^4(\phi_1^{\text{in}}-\phi_2^{\text{in}})| &\le \mu_0^{\frac12}t^{-\frac32}h_0\Big(\frac{r}{\sqrt t}\Big)\frac{\|\phi_1-\phi_2\|_{\nu,a}}{R^{1+a}}\\ &\le \boldsymbol{c}\mu_0^{\frac12}t^{-\frac32}h_0\Big(\frac{r}{\sqrt t}\Big)\|\phi_1-\phi_2\|_{\nu,a}.\end{aligned}$$

The constant c_1 in the last two formulas can be made arbitrarily small provided R is chosen large enough. □

5. Choice of λ: part I

Let $\psi=\Psi[\lambda,\phi]$ be the solution to problem (3-6) predicted by Proposition 4.1, and satisfying the properties described in Proposition 4.3. We substitute ψ in (3-11) and (3-12), and we want to solve, in ϕ, problem

(3-11), satisfying the initial condition (3-13). As we stated in Proposition 3.1, problem (3-11)–(3-13) can be solved for functions ϕ satisfying (4-1), provided that

$$\int_{B_{2R}} H[\psi,\lambda,\phi](y,t(\tau))Z_0(y)\,dy = 0 \quad \text{for all } t > t_0, \tag{5-1}$$

where $H[\psi,\lambda,\phi]$ is defined in (3-12).

Next lemma states that (5-1) is a nonlinear, nonlocal equation in λ, at any fixed ϕ.

Lemma 5.1. *Assume that λ satisfies (2-11), and that the function ϕ satisfies the bound (4-1). Let $\psi = \Psi[\lambda,\phi]$ be the solution to problem (3-6) predicted by Proposition 4.1. Then (5-1) is equivalent to*

$$[1+\mu_0\mu_0' b(t) + q_1(\lambda)]\phi_0(0,t) = g(t) + G[\lambda,\phi](t). \tag{5-2}$$

Here ϕ_0 is the function defined in (2-53) and also in (2-54); thus

$$\phi_0(0,t) = \int_{t_0-1}^{t} \frac{1}{(4\pi(t-s))^{\frac{3}{2}}} \int_{\mathbb{R}^3} e^{-\frac{|y|^2}{4(t-s)}} \frac{\bar{\alpha}(s)}{\mu+|y|} \mathbf{1}_{\{r<M\}}\,dy\,ds. \tag{5-3}$$

The function $b = b(t)$ is a smooth function in (t_0,∞). By $q_1(s)$ we denote a smooth function such that $q_1(0) = 0$ and $q_1'(0) \neq 0$. Moreover,

$$\|b\|_\infty < C, \quad \|g\|_\flat \le C, \quad \|G[\lambda,\phi]\|_\flat \le C. \tag{5-4}$$

Furthermore, if the initial time t_0 in problem (2-1) is chosen large enough, there exists R in the definition of the cut-off function in (3-2) sufficiently large and there exists a constant $\boldsymbol{c} \in (0,1)$ so that, for any ϕ,

$$\|G[\lambda_1,\phi] - G[\lambda_2,\phi]\|_\flat \le \boldsymbol{c}\|\lambda_1-\lambda_2\|_\sharp \tag{5-5}$$

and, for any λ,

$$\|G[\lambda,\phi_1] - G[\lambda,\phi_2]\|_\flat \le \boldsymbol{c}\|\phi_1-\phi_2\|_{\nu,a}. \tag{5-6}$$

The constants $\boldsymbol{c}$ in (5-5) and (5-6) can be made as small as one needs, provided that the initial time t_0 is chosen large enough. We refer to (2-43) and (3-18) for the definitions of $\|\cdot\|_\flat$ and $\|\cdot\|_{\nu,a}$ respectively.

Proof. Throughout the proof, we denote by $q_i = q_i(s)$, for any integer i, a smooth real function, with the property that

$$\frac{d}{(ds)^j} q_i(0) = 0 \quad \text{for } j<i \qquad \text{and} \qquad \frac{d}{(ds)^i} q_i(0) \neq 0.$$

We have the decomposition

$$\begin{aligned}
&\int_{B_{2R}} H[\psi,\lambda,\phi](y,t(\tau))Z_0(y)\,dy \\
&\quad = \mu_0^{\frac{5}{2}} \int_{B_{2R}} \mathcal{E}_{22}(\mu_0 y,t)Z_0(y)\,dy + 5\int_{B_{2R}} \frac{\mu_0^{\frac{1}{2}}}{(1+\lambda)^2} w^4\Big(\frac{y}{1+\lambda}\Big)\psi(\mu_0 y,t)Z_0(y)\,dy \\
&\qquad\qquad + \int_{B_{2R}} B[\phi]Z_0(y)\,dy \int_{B_{2R}} B^0[\phi]Z_0(y)\,dy \\
&\quad = i_1+i_2+i_3+i_4.
\end{aligned}$$

For any $j = 1, \dots, 4$, i_j is a function of t and depends also on λ and ϕ. To emphasize this dependence, we write $i_j = i_j[\lambda, \phi](t)$.

We claim that

$$\mu_0^{-\frac{1}{2}} i_1[\lambda, \phi](t) = \mu_0^2 \mu^{-2} \Big[\Big(5 \int_{B_{2R}} w^4(y) Z_0(y)\, dy \Big) \phi_0(0,t) + (q_1(\lambda) + \mu_0 \mu_0' q_0(\lambda)) \phi_0(0,t) + \mu_0^\sigma \alpha(t) b(t) \Big], \tag{5-7}$$

where $b(t)$ is a smooth function in (t_0, ∞) which is uniformly bounded as $t \to \infty$.

Observe that i_1 does not depend on ϕ. From (2-53) satisfied by ϕ_0, and Lemma A.1, we get the existence of a positive constant c so that $|\phi_0(\mu_0 y, t)| \le c\alpha(t)\mu_0(t)$ for any $y \in B_{2R}$. Thus, we Taylor expand $\mathcal{E}_{22}$ in the region $y \in B_{2R}$ as follows:

$$\mathcal{E}_{22}(\mu_0 y, t) = 5U_1^4 \phi_0 + 4(U_1 + s\phi_0)^3 \phi_0^2 = a + b$$

for some $s \in (0,1)$. Let us first analyze a. We write

$$a = 5\mu^{-2} w^4(y) \phi_0(0,t) + \underbrace{5[U_1^4(\mu_0 y) - \mu^{-2} w^4(y)] \phi_0(0,t)}_{:=a_1} + \underbrace{5U_1^4[\phi_0(\mu_0 y, t) - \phi_0(0,t)]}_{:=a_2}.$$

Observe that, by the definition of U_1 in (2-38), and (2-13), we have

$$\begin{aligned} U_1^4(\mu_0 y) - \mu^{-2} w^4(y) &= \Big[w_\mu(\mu_0 y) + \mu_0' \mu^{\frac{1}{2}} \Phi_1\Big(\frac{\mu_0 r}{\mu}\Big) \Big]^4 - \mu^{-2} w^4(y) \\ &= \mu^{-2} \Big[w(y) + \Big(w\Big(\frac{y}{(1+\Lambda)^2}\Big) - w(y) \Big) + \mu_0' \mu \Phi_1\Big(\frac{\mu_0 r}{\mu}\Big) \Big]^4 - \mu^{-2} w^4(y) \\ &= 4\mu^{-2} w^3(y) s \Big[\Big(w\Big(\frac{y}{(1+\Lambda)^2}\Big) - w(y) \Big) + \mu_0' \mu \Phi_1\Big(\frac{\mu_0 r}{\mu}\Big) \Big] \end{aligned}$$

for some $s \in (0,1)$. Observe that

$$w\Big(\frac{y}{(1+\Lambda)^2}\Big) - w(y) = \nabla w(y) \cdot y + \nabla w(y) \cdot yz[-2\Lambda - \Lambda^2] \tag{5-8}$$

for some $z \in (0,1)$. Taking into account also the description of Φ_1 in (2-9), we get

$$\int_{B_{2R}} a_1 Z_0 \, dy = \mu^{-2} [q_1(\Lambda) + \mu_0 \mu_0' q_0(\Lambda)] \phi_0(0,t). \tag{5-9}$$

We next claim that, for $y \in B_{2R}$, we have

$$\phi_0(\mu_0 y, t) - \phi_0(0,t) = \alpha(t) |\mu_0 y|^\sigma \Pi(t) \Theta(|y|) \tag{5-10}$$

for some $\sigma \in (0,1)$. We postpone the proof of (5-10) to the Appendix. We thus get

$$\int_{B_{2R}} a_2 Z_0 \, dy = \mu^{-2} \mu_0^\sigma \alpha(t) b(t). \tag{5-11}$$

Collecting estimates (5-9)–(5-11) we get (5-7).

We claim that

$$\mu_0^{-\frac{1}{2}} i_2[\lambda,\phi](t) = g(t) + G[\lambda,\phi](t), \tag{5-12}$$

with

$$\|g\|_\flat \le c, \quad \|G[\lambda,\phi]\|_\flat \le c$$

for some constant c. We refer to (2-43) for the definition of $\|\cdot\|_\flat$. Furthermore, we claim that G satisfies estimates (5-5) and (5-6) for some constant $c_1 \in (0,1)$. To prove the above assertion, we write

$$\begin{aligned}
\mu_0^{-\frac{1}{2}} i_2[\lambda,\phi](t) = 5\int_{B_{2R}} & w^4(y)\psi[0,0](\mu_0 y,t) Z_0(y)\,dy \\
& + 5\int_{B_{2R}} w^4(y)[\psi[\lambda,0]-\psi[0,0]](\mu_0 y,t) Z_0(y)\,dy \\
& + 5\int_{B_{2R}} w^4(y)[\psi[\lambda,\phi]-\psi[\lambda,0]](\mu_0 y,t) Z_0(y)\,dy \\
& + 5\int_{B_{2R}} \left[w^4\left(\frac{y}{(1+\Lambda)^2}\right) - w^4(y)\right]\psi[\lambda,\phi](\mu_0 y,t) Z_0(y)\,dy \\
& + 5\left[\frac{1}{(1+\Lambda)^4}-1\right]\int_{B_{2R}} w^4\left(\frac{y}{(1+\Lambda)^2}\right)\psi[\lambda,\phi](\mu_0 y,t) Z_0(y)\,dy = \sum_{j=1}^{5} g_j.
\end{aligned}$$

The first term,

$$g_1(t) = 5\int_{B_{2R}} w^4(y)\psi(\mu_0 y,t)[0,0] Z_0(y)\,dy,$$

is an explicit smooth function, globally defined in (t_0,∞), which satisfies the bound

$$\|g_1\|_\flat \le c\left(5\int_{B_{2R}} w^4(y)|y| Z_0(y)\,dy\right) \tag{5-13}$$

for some constant $c > 0$, as direct consequence of (4-7). Let us analyze the term g_5. We see that $g_5 = g_5[\lambda,\phi](t)$. Let us first assume that λ and ϕ are fixed. From (4-7), we get

$$|g_5(t)| \le c q_1(\lambda)\int_{B_{2R}} |w^4(y)\psi[\lambda,\phi](\mu_0 y,t) Z_0(y)|\,dy \le c\mu_0^{\frac{3}{2}} t^{-1} q_1(\lambda)\int \frac{|y|}{(1+|y|^5)}\,dy.$$

Using again (4-7) and the assumptions on λ and on ϕ, we get $[g_5]_{0,\sigma,[t,t+1]} \le c\mu_0^{\frac{3}{2}} t^{-1}$, from which we conclude that $\|g_5\|_\flat \le c$, for some constant $c > 0$. Let us now fix ϕ and take λ_1, λ_2 satisfying (2-11). We write

$$\begin{aligned}
& g_5[\lambda_1,\phi] - g_5[\lambda_2,\phi] \\
& = 5\left[\frac{1}{(1+\Lambda_1)^4} - \frac{1}{(1+\Lambda_2)^4}\right]\int_{B_{2R}} w^4\left(\frac{y}{(1+\Lambda_1)^2}\right)\psi[\lambda_1,\phi](\mu_0 y,t) Z_0(y)\,dy \\
& \quad + 5\left[\frac{1}{(1+\Lambda_2)^4}-1\right]\int_{B_{2R}}\left[w^4\left(\frac{y}{(1+\Lambda_1)^2}\right) - w^4\left(\frac{y}{(1+\Lambda_2)^2}\right)\right]\psi[\lambda_1,\phi](\mu_0 y,t) Z_0(y)\,dy \\
& \quad + 5\left[\frac{1}{(1+\Lambda_2)^4}-1\right]\int_{B_{2R}} w^4\left(\frac{y}{(1+\Lambda_2)^2}\right)[\psi[\lambda_1,\phi]-\psi[\lambda_2,\phi]](\mu_0 y,t) Z_0(y)\,dy = e_1+e_2+e_3.
\end{aligned}$$

Thanks to (2-11), and arguing as before, we see that

$$\begin{aligned}|e_1(t)| &\le c|\Lambda_1(t)-\Lambda_2(t)|\int_{B_{2R}}|w^4(y)\psi[\lambda_1,\phi](\mu_0 y,t)Z_0(y)|\,dy\\ &\le c\mu_0(t)^{\frac{3}{2}}t^{-1}\left(\int_t^\infty s^{-1}\mu_0(s)\,ds\right)\|\lambda_1-\lambda_2\|_\sharp\\ &\le[\mu_0(t_0)]\mu_0(t)^{\frac{3}{2}}t^{-1}\|\lambda_1-\lambda_2\|_\sharp\le c_1\mu_0(t)^{\frac{3}{2}}t^{-1}\|\lambda_1-\lambda_2\|_\sharp,\end{aligned}$$

where c_1 is a positive number, which can be chosen arbitrarily small, in particular $c_1<1$, provided t_0 is chosen large enough. Similarly one can show that, thanks to (2-11),

$$[e_1]_{0,\sigma,[t,t+1]}\le c_1\mu_0(t)^{\frac{3}{2}}t^{-1}\|\lambda_1-\lambda_2\|_\sharp.$$

We thus can conclude that there exists a positive small number $c_1<1$ so that

$$\|e_1\|_\flat\le c_1\|\lambda_1-\lambda_2\|_\sharp.$$

A similar argument allow us to say that also $\|e_2\|_\flat\le c_1\|\lambda_1-\lambda_2\|_\sharp$. We next analyze e_3. From (4-22) we get

$$\begin{aligned}|e_3(t)| &\le \mu_0^{\frac{3}{2}}t^{-1}\left(\int w^4(y)\frac{|y|}{1+|y|}\,dy\right)\|\psi[\lambda_1,\phi]-\psi[\lambda_2,\phi]\|_{**}\\ &\le c_1\mu_0^{\frac{3}{2}}t^{-1}\|\lambda_1-\lambda_2\|_\sharp,\end{aligned}$$

and also

$$[e_3]_{0,\sigma,[t,t+1]}\le c_1\mu_0^{\frac{3}{2}}t^{-1}\|\lambda_1-\lambda_2\|_\sharp$$

for some constant $c_1\in(0,1)$. We can conclude that

$$\|g_5[\lambda_1,\phi]-g_5[\lambda_2,\phi]\|_\flat\le c_1\|\lambda_1-\lambda_2\|_\sharp.$$

The same estimate can be obtained for g_4, arguing in a similar way.

Let us now consider g_2. This term does not depend on ϕ; namely $g_2[\lambda,\phi](t)=g_2[\lambda](t)$. From Proposition 4.3, we get

$$|g_2(t)|\le\mu_0^{\frac{3}{2}}t^{-2}\left(\int w^4\frac{|y|}{1+|y|}\,dy\right)\|\lambda\|_\sharp\le c\mu_0^{\frac{3}{2}}t^{-2}\|\lambda\|_\sharp,$$

and similarly

$$[g_2(t)]_{0,\sigma,[t,t+1]}\le c\mu_0^{\frac{3}{2}}t^{-2}\|\lambda\|_\sharp.$$

Furthermore, if t_0 is large enough, there exists $c_1\in(0,1)$ so that

$$\begin{aligned}|g_2[\lambda_1](t)-g_2[\lambda_2](t)| &\le 5\int_{\mathbb{R}^3}w^4(y)\|[\psi[\lambda_1,0]-\psi[\lambda_2,0]](\mu_0 y,t)\|Z_0\,dy\\ &\le Ct_0^{-1}\mu_0^{\frac{3}{2}}t^{-2}\|\lambda_1-\lambda_2\|_\sharp\le c_1\mu_0^{\frac{3}{2}}t^{-2}\|\lambda_1-\lambda_2\|_\sharp\end{aligned}$$

and also

$$[g_1[\lambda_2]-g_2[\lambda_2]]_{0,\sigma,[t,t+1]}\le c_1\mu_0^{\frac{3}{2}}t^{-2}\|\lambda_1-\lambda_2\|_\sharp$$

thanks to the results of Proposition 4.3. Arguing in the same way, one gets similar estimates for g_3.

Collecting all the above arguments, we conclude that $\mu_0^{-\frac{1}{2}} i_2[\lambda,\phi](t)$ can be written as in (5-12), with g and G satisfying (5-4), (5-5) and (5-6).

Next we claim that

$$\mu_0^{-\frac{1}{2}} i_j[\lambda,\phi](t) = G[\lambda,\phi](t), \quad j=3,4, \tag{5-14}$$

and G satisfies (5-4), (5-5) and (5-6). We start with $j=3$. First, we see that i_3 does not depend on λ, and it is linear in ϕ. Since we are assuming that ϕ satisfies (4-1), we have

$$|\mu_0^{-\frac{1}{2}} i_3(t)| \le (\mu_0\mu_0' R^{2-a})\mu_0^{\frac{3}{2}}(t)t^{-1}\|\phi\|_{\nu,a} \le c\mu_0^{\frac{3}{2}}(t)t^{-1}\|\phi\|_{\nu,a},$$
$$[\mu_0^{-\frac{1}{2}} i_3(t)]_{0,\sigma,[t,t+1]} \le c\mu_0^{\frac{3}{2}}(t)t^{-1}\|\phi\|_{\nu,a}$$

for some constant $c>0$. Let us now take ϕ_1, and ϕ_2, and we get that, if $\mu_0(t_0)\mu_0'(t_0)R^{2-a}$ is small enough,

$$|\mu_0^{-\frac{1}{2}}(i_3[\phi_1]-i_3[\phi_2])(t)| \le c_1\mu_0^{\frac{3}{2}}(t)t^{-1}\|\phi_1-\phi_2\|_{\nu,a},$$
$$[\mu_0^{-\frac{1}{2}}(i_3[\phi_1]-i_3[\phi_2])(t)]_{0,\sigma,[t,t+1]} \le c_1\mu_0^{\frac{3}{2}}(t)t^{-1}\|\phi_1-\phi_2\|_{\nu,a}$$

for some $c_1\in(0,1)$. Estimate (5-14) for $j=4$ can be proved in a very similar way. We leave the details to the interested reader. Combining (5-7), (5-12) and (5-14), we complete the proof of (5-2). ☐

6. Solving a nonlocal linear problem

Let ϕ_0 be the function introduced in (2-53). Later in our argument we will need to solve in λ, a nonlocal equation of the form

$$\phi_0(0,t) = h(t), \quad t\in(t_0,\infty), \tag{6-1}$$

for a certain right-hand side h. We see from (5-3) that $\phi_0(0,t)$, defined as

$$\phi_0(0,t) = \int_{t_0-1}^{t}\int_{\mathbb{R}^3} \frac{\bar\alpha(s)}{(4\pi(t-s))^{\frac{3}{2}}}\frac{e^{-\frac{|y|^2}{4(t-s)}}}{\mu+|y|}\mathbf{1}_{\{|y|<M\}}\,dy\,ds,$$

defines a nonlocal nonlinear operator in λ. For convenience we recall that

$$\alpha(t) = 3^{\frac{1}{4}}\mu_0^{-\frac{1}{2}}(\mu_0\Lambda)', \quad \bar\alpha(t) = \begin{cases}\alpha(t_0) & \text{for } t<t_0,\\ \alpha(t) & \text{for } t\ge t_0,\end{cases} \quad \Lambda(t) = \int_t^\infty \lambda(s)\,ds.$$

We write

$$\phi_0(0,t) = T[\lambda](t) + \widehat{T}[\lambda](t), \tag{6-2}$$

where T is

$$T[\lambda](t) = \int_{t_0-1}^{t}\int_{\mathbb{R}^3} \frac{\bar\alpha(s)}{(4\pi(t-s))^{\frac{3}{2}}}\frac{e^{-\frac{|y|^2}{4(t-s)}}}{|y|}\mathbf{1}_{\{|y|<M\}}\,dz\,ds. \tag{6-3}$$

We shall see that $\widehat{T}$ is a small perturbation of T, in a sense we will make precise later. In this section, we start with the analysis of problem

$$T[\lambda](t) = h(t), \quad t>t_0. \tag{6-4}$$

Straightforward computations give

$$T[\lambda](t) = -\frac{\bar\omega_3}{4}\int_{t_0-1}^{t} \frac{\bar\alpha(s)}{\sqrt{t-s}}\big(1-e^{-\frac{M^2}{(t-s)}}\big)\,ds. \tag{6-5}$$

Indeed, letting $z = y/(2\sqrt{t-s})$, one gets

$$\begin{aligned} T[\lambda](t) &= \int_{t_0-1}^{t}\int_{\mathbb{R}^3} \frac{\bar\alpha(s)}{2\sqrt{t-s}}\,\frac{e^{-|z|^2}}{|z|}\mathbf{1}_{\{|z|<\frac{M}{\sqrt{t-s}}\}}\,dz\,ds \\ &= \frac{\bar\omega_3}{2}\int_{t_0-1}^{t}\int_0^{\infty} \frac{\bar\alpha(s)}{\sqrt{t-s}} e^{-\rho^2}\rho\mathbf{1}_{\{\rho<\frac{M}{\sqrt{t-s}}\}}\,d\rho\,ds = \frac{\bar\omega_3}{4}\int_{t_0-1}^{t}\frac{\bar\alpha(s)}{\sqrt{t-s}}\int_0^{\frac{M}{\sqrt{t-s}}} e^{-\rho^2}2\rho\,d\rho \\ &= -\frac{\bar\omega_3}{4}\int_{t_0-1}^{t}\frac{\bar\alpha(s)}{\sqrt{t-s}}\big(1-e^{-\frac{M^2}{(t-s)}}\big)\,ds. \end{aligned} \tag{6-6}$$

Introduce the function $\beta = \beta(t)$ as

$$\beta(t) = \frac{\bar\omega_3}{4}\int_t^{\infty}\bar\alpha(s)\,ds. \tag{6-7}$$

If $\beta = \beta(t)$ solves

$$\int_{t_0-1}^{t}\frac{\beta'(s)}{\sqrt{t-s}}\big(1-e^{-\frac{M^2}{(t-s)}}\big)\,ds = h(t), \tag{6-8}$$

then the function $\Lambda(t) = \int_t^{\infty}\lambda(s)\,ds$, defined as

$$\bar\omega\Lambda(t) = \mu_0^{-\frac{1}{2}}(t)\beta(t) + \frac{\mu_0^{-1}(t)}{2}\int_t^{\infty}\beta(s)\mu_0^{-\frac{1}{2}}\mu_0'(s)\,ds, \quad \bar\omega = \frac{\bar\omega_3}{4}3^{\frac{1}{4}}, \tag{6-9}$$

solves (6-4).

The next lemma constructs a solution to (6-8). If we formally let $M\to\infty$ in (6-8), we get that the left-hand side of (6-8) is nothing but the $\frac{1}{2}$-Caputo derivative of β. This fact inspires the proof of the following:

Lemma 6.1. *Let $\sigma = \frac{1}{2}+\sigma'$, with $\sigma' > 0$ small, be the number fixed in (2-11), and $h : (t_0,\infty)\to\mathbb{R}$ a smooth function satisfying*

$$\sup_{t>t_0}\mu_0^{-\frac{3}{2}} t[\|h\|_{0,[t,t+1]} + [h]_{0,\sigma,[t,t+1]}] \le C \tag{6-10}$$

for some constant C. Then there exist a constant C_1 and a unique smooth function $\beta : (t_0-1,\infty)\to\mathbb{R}$ which solves (6-8), $\beta\in C^1$, that satisfy the bound

$$\sup_{t>t_0}\mu_0^{-\frac{3}{2}} t[\|\beta'\|_{0,[t,t+1]} + [\beta']_{0,\sigma,[t,t+1]}] \le C_1 M^{-1}. \tag{6-11}$$

We recall that $M^2 = t_0$ was first introduced in (2-53).

Observe that a direct consequence of this lemma, together with (6-9) and (2-41), is the invertibility theory for problem (6-4) that will be used in next section to solve (5-1). This is contained in the following:

Proposition 6.2. *The function $T : X_\sharp \to X_\flat$, defined in (6-3) is a linear, nonlocal, homeomorphism such that*

$$\|T^{-1}(h)\|_\sharp \le CM^{-1}\|h\|_\flat \quad \text{for any } h \in X_\flat, \tag{6-12}$$

for some fixed positive constant C. We refer to (2-11) and to (2-12) for the definition of the $\|\cdot\|_\sharp$-norm and of the set $X_\sharp$, and to (2-43) and (2-42) for the definition of the norm $\|\cdot\|_\flat$ and of the space $X_\flat$.

We devote the rest of the section to the following:

Proof of Lemma 6.1. We start performing a change of variables, to transform problem (6-8) into an equivalent one with simpler form: Let

$$s = t_0 - 1 + M^2 a, \quad t = t_0 - 1 + M^2 b, \quad \tilde\beta(a) = \beta(s), \quad \tilde h(b) = h(t).$$

After this change of variables, problem (6-8) takes the form

$$\int_0^b \frac{\tilde\beta'(a)}{\sqrt{b-a}}(1 - e^{-\frac{1}{b-a}})\, da = M\tilde h(b). \tag{6-13}$$

Let

$$K(\eta) = \frac{1 - e^{-\frac{1}{\sqrt{\eta}}}}{\sqrt{\eta}}$$

and take the Laplace transform of both sides in (6-13), thus getting

$$\mathcal{L}(\tilde\beta')(\xi)\mathcal{L}(K)(\xi) = M\mathcal{L}(\tilde h)(\xi).$$

Since $\mathcal{L}(\tilde\beta') = \xi\mathcal{L}(\tilde\beta)(\xi) - \tilde\beta(0)$, we get

$$\mathcal{L}(\tilde\beta)(\xi) = \frac{\tilde\beta(0)}{\xi} + M\frac{\mathcal{L}(\tilde h)(\xi)}{\xi\mathcal{L}(K)(\xi)}. \tag{6-14}$$

Observe now that

$$\mathcal{L}(K)(\xi) = \int_0^\infty e^{-\xi\eta}\left(\frac{1 - e^{-\frac{1}{\eta}}}{\sqrt{\eta}}\right) d\eta = \frac{2}{\sqrt{\xi}}\int_0^\infty e^{-p^2}(1 - e^{-\frac{\xi}{p^2}})\, dp.$$

We readily get that

$$\mathcal{L}(K)(\xi) = \frac{1}{\sqrt{\xi}}\left(2\int_0^\infty e^{-p^2}\, dp\right)(1 + o(1)) \quad \text{as } \xi \to \infty. \tag{6-15}$$

To describe the behavior of $\mathcal{L}(K)(\xi)$, for $\xi \to 0$, we first notice that

$$\int_0^{\frac{1}{\xi}} e^{-\xi\eta}\left(\frac{1 - e^{-\frac{1}{\eta}}}{\sqrt{\eta}}\right) d\eta = \int_0^\infty \frac{1 - e^{-\frac{1}{\eta}}}{\sqrt{\eta}}\, d\eta + O(\sqrt{\xi}).$$

On the other hand,

$$\int_{\frac{1}{\xi}}^\infty e^{-\xi\eta}\left(\frac{1 - e^{-\frac{1}{\eta}}}{\sqrt{\eta}}\right) d\eta = \int_{\frac{1}{\xi}}^\infty e^{-\xi\eta}\left(\frac{1 - \frac{1}{\eta} - e^{-\frac{1}{\eta}}}{\sqrt{\eta}}\right) d\eta + \int_{\frac{1}{\xi}}^\infty \frac{e^{-\xi\eta}}{\eta\sqrt{\eta}}\, d\eta = O(\sqrt{\xi}).$$

Thus we conclude that

$$\mathcal{L}(K)(\xi) = \int_0^\infty \frac{1-e^{-\frac{1}{\eta}}}{\sqrt{\eta}}\,d\eta + O(\sqrt{\xi}) \quad \text{as } \xi \to 0. \tag{6-16}$$

From (6-15) and (6-16), we conclude that

$$\frac{1}{\xi\mathcal{L}(K)(\xi)} = \begin{cases} c_1/\xi + c_2/\sqrt{\xi} + O(1) & \text{if } \xi \to 0, \\ c_3/\sqrt{\xi}(1+o(1)) & \text{if } \xi \to \infty. \end{cases}$$

Let now $G = G(t)$ be so that $\mathcal{L}(G)(\xi) = 1/(\xi\mathcal{L}(K)(\xi))$. Standard arguments on Laplace transformation imply that

$$G(t) = \begin{cases} \tilde{c}_1 + \tilde{c}_2/\sqrt{t} + O(1/t) & \text{if } t \to \infty, \\ \tilde{c}_3/\sqrt{t}(1+o(1)) & \text{if } t \to 0 \end{cases}$$

for certain constants $\tilde{c}_1$, $\tilde{c}_2$ and $\tilde{c}_3$. From (6-14), taking the anti-Laplace transform of both sides, we get

$$\begin{aligned} \tilde{\beta}(b) &= \tilde{\beta}(0) + M\int_0^b \tilde{h}(a)G(b-a)\,da \\ &= \tilde{\beta}(0) + M\tilde{c}_1\int_0^\infty \tilde{h}(a)\,da + M\tilde{c}_1\int_b^\infty \tilde{h}(a)\,da + M\int_0^b \tilde{h}(a)[G(b-a)-\tilde{c}_1]\,da. \end{aligned}$$

We select the solution to problem (6-13) so that

$$\tilde{\beta}(0) + M\tilde{c}_1\int_0^\infty \tilde{h}(a)\,da = 0.$$

In the original variables, we thus obtain an explicit solution to (6-8):

$$\beta(t) = \underbrace{\frac{\tilde{c}_1}{M}\int_t^\infty h(s)\,ds}_{:=\beta_1(t)} + \underbrace{\frac{1}{M}\int_{t_0-1}^t h(s)\Big[G\Big(\frac{t-s}{M^2}\Big)-\tilde{c}_1\Big]\,ds}_{:=\beta_2(t)}. \tag{6-17}$$

Let us now check (6-11). Since (6-10) holds, we easily get

$$\sup_{t>t_0} \mu_0^{-\frac{3}{2}}|\beta_1(t)| \lesssim M^{-1}.$$

To control the second term in (6-17), we change variable $t = M^2\bar{t}$, $s = M^2\bar{s}$, so that

$$\beta_2(t) = M\int_{\frac{t_0-1}{M^2}}^{\bar{t}} h(M^2\bar{s})[G(\bar{t}-\bar{s})-\tilde{c}_1]\,d\bar{s}.$$

Since $t_0 = M^2$ and since (6-10) holds, we get

$$|\beta_2(t)| \lesssim \frac{1}{M}\int_{1-\frac{1}{t_0}}^{\bar{t}} \frac{\mu_0^{\frac{3}{2}}(\bar{s})}{\bar{s}}[G(\bar{t}-\bar{s})-\tilde{c}_1]\,d\bar{s} \lesssim \frac{1}{M}\mu_0^{\frac{3}{2}}(\bar{t}) \lesssim M^{-1}\mu_0^{\frac{3}{2}}(t),$$

from which we get the validity of (6-11).

The assumption that $\mu_0^{-\frac{3}{2}} t[h]_{0,\sigma,[t,t+1]}$ is bounded guarantees that the function β defined in (6-17) is differentiable. Indeed, trivially one has $\beta_1'(t) = -(\tilde{c}_1/M)h(t)$. Let us write β_2 in the following way:

$$\beta_2(t) = \frac{1}{M}\int_{t_0}^{t} (h(s)-h(t))\Big[G\Big(\frac{t-s}{M^2}\Big)-\tilde{c}_1\Big]\,ds + \frac{h(t)}{M}\int_{t_0}^{t}\Big[G\Big(\frac{t-s}{M^2}\Big)-\tilde{c}_1\Big]\,ds.$$

Thus we have

$$\begin{aligned}\beta_2'(t) &= \underbrace{\frac{1}{M}\lim_{s\to t}\Big[(h(s)-h(t))\Big[G\Big(\frac{t-s}{M^2}\Big)-\tilde{c}_1\Big]\Big]}_{=0} + \frac{1}{M^3}\int_{t_0}^{t}(h(s)-h(t))G'\Big(\frac{t-s}{M^2}\Big)\,ds\\ &\quad + \underbrace{\frac{h'(t)}{M}\int_{t_0}^{t}\Big[G\Big(\frac{t-s}{M^2}\Big)-\tilde{c}_1\Big]\,ds - \frac{h'(t)}{M}\int_{t_0}^{t}\Big[G\Big(\frac{t-s}{M^2}\Big)-\tilde{c}_1\Big]\,ds}_{=0}\\ &\quad + \frac{h(t)}{M}\frac{d}{dt}\Big(\int_{t_0}^{t}\Big[G\Big(\frac{t-s}{M^2}\Big)-\tilde{c}_1\Big]\,ds\Big)\\ &= \frac{1}{M^3}\int_{t_0}^{t}(h(s)-h(t))G'\Big(\frac{t-s}{M^2}\Big)\,ds + \frac{h(t)}{M}\frac{d}{dt}\Big(\int_{t_0}^{t}\Big[G\Big(\frac{t-s}{M^2}\Big)-\tilde{c}_1\Big]\,ds\Big).\end{aligned}$$

Both the last two integrals are well-defined, as consequence of the behavior of $G(\eta)$, as $\eta\to 0$, and the assumption that $\mu_0^{-\frac{3}{2}} t[h]_{0,\sigma,[t,t+1]}$ is bounded. Since $G(\eta)\sim \eta^{-\frac{1}{2}}$ as $\eta\to 0$, direct computations give the bounds in (6-11) for $\beta'(t)$. □

7. Choice of λ: part II

This section is devoted to solving (5-1) in λ for fixed ϕ satisfying (4-1). We have the validity of the following:

Proposition 7.1. *For any ϕ satisfying (4-1), there exists $L>0$ and a unique solution $\lambda=\lambda[\phi]$ to (5-1), with*

$$\|\lambda\|_\sharp \le LM^{-1}, \tag{7-1}$$

where $M=\sqrt{t_0}$, provided the initial time t_0 in problem (2-1) is chosen large enough. Furthermore, there exists a constant $\boldsymbol{c}\in(0,1)$ such that, for any ϕ_1, ϕ_2 satisfying (4-1), we have

$$\|\lambda[\phi_1]-\lambda[\phi_2]\|_\sharp \le \boldsymbol{c}\,\|\phi_1-\phi_2\|_{\nu,a}. \tag{7-2}$$

Proof of Proposition 7.1.. Lemma 5.1 states that solving (5-1) is equivalent to solving (5-2). We write (5-2) as

$$T[\lambda](t)+\widehat{T}[\lambda](t) = (1+\mu_0\mu_0' b(t)+q_1(\lambda))^{-1}[g(t)+G[\lambda,\phi](t)], \tag{7-3}$$

where T and $\widehat{T}$ are defined in (6-2) and (6-3), while b, g and G satisfy the bounds in (5-4), (5-5) and (5-6). Here $q_1=q_1(s)$ denotes a smooth function such that $q_1(0)=0$ and $q_1'(0)\neq 0$. We observe first that

$$(1+\mu_0\mu_0' b(t)+q_1(\lambda))^{-1}[g(t)+G[\lambda,\phi](t)] = g_1(t)+G_1[\lambda,\phi](t)$$

for some new functions g_1 and G_1 that also satisfy (5-4), (5-5), and (5-6).

Thanks to the result of Proposition 6.2, solving (7-3) in λ reduces to solving the fixed-point problem

$$\lambda(t) = \mathcal{F}(\lambda)(t), \quad \mathcal{F}(\lambda) := T^{-1}(g_1 + G_1[\lambda, \phi] - \widehat{T}[\lambda]), \tag{7-4}$$

where T^{-1} is the operator introduced in Proposition 6.2.

Step 1: First we show that, for any fixed ϕ satisfying (4-1), there exists a unique fixed point $\lambda = \lambda[\phi]$ of contraction type for $\mathcal{F}$ in the set

$$B = \{\lambda \in X_\sharp : \|\lambda\|_\sharp \le LM^{-1}\}$$

for some $L > 0$ large.

In order to prove this fact, we claim that, if the initial time t_0 in problem (2-1) is large enough, there are positive constants $\bar{c}_1, \bar{c}_2 \in (0, 1)$ so that, for any $\lambda \in B$,

$$\|\widehat{T}[\lambda]\|_\flat \le \bar{c}_1 M \|\lambda\|_\sharp, \qquad \text{with } \bar{c}_1 C < 1, \tag{7-5}$$

$$\|\widehat{T}[\lambda_1] - \widehat{T}[\lambda_2]\|_\flat \le \bar{c}_2 \|\lambda_1 - \lambda_2\|_\sharp, \quad \text{with } CM^{-1}(\boldsymbol{c} + \bar{c}_2) < 1, \tag{7-6}$$

for any $\lambda_1, \lambda_2 \in B$. The constant C is the constant appearing in (6-12), while $\boldsymbol{c}$ is the one appearing in (5-5).

Assume for the moment the validity of (7-5) and (7-6). For any $\lambda \in B$, we have

$$\begin{aligned}\|\mathcal{F}(\lambda)\|_\sharp &\le CM^{-1}\|g_1 + G_1[\lambda, \phi] - \widehat{T}[\lambda]\|_\flat \\ &\le CM^{-1}(\|g_1\|_\flat + \|G_1[\lambda, \phi]\|_\flat + \|\widehat{T}[\lambda]\|_\flat) \\ &\le CM^{-1}(2c + \bar{c}_1 L) \le LM^{-1},\end{aligned}$$

provided $L > (2cC)/(1 - \bar{c}_1 C)$, where C is the constant in (6-12), c is the constant in (5-4), and $\bar{c}_1$ is the constant in (7-5), which satisfies $\bar{c}_1 C < 1$.

Let us take now $\lambda_1, \lambda_2 \in B$. We have

$$\begin{aligned}\|\mathcal{F}(\lambda_1) - \mathcal{F}(\lambda_2)\|_\sharp &= \|T^{-1}(G_1[\lambda_1, \phi] - G_1[\lambda_2, \phi]) - T^{-1}(\widehat{T}[\lambda_1] - \widehat{T}[\lambda_2])\|_\sharp \\ &\le CM^{-1}(\|G_1[\lambda_1, \phi] - G_1[\lambda_2, \phi]\|_\flat + \|\widehat{T}[\lambda_1] - \widehat{T}[\lambda_2]\|_\flat) \\ &\le CM^{-1}(c_1 + \bar{c}_2)\|\lambda_1 - \lambda_2\|_\sharp < \varepsilon\|\lambda_1 - \lambda_2\|_\sharp\end{aligned}$$

for some $\varepsilon < 1$, thanks to the choice of $\bar{c}_2$ in (7-6).

A direct application of the Banach fixed-point theorem gives the existence and uniqueness of a solution λ to (5-1), satisfying (7-1). We complete the first part of the proof of the proposition with the proofs of (7-5) and (7-6).

Proof of (7-5). Let $\lambda \in B$. From (6-2) and (6-3), we get

$$\begin{aligned}\widehat{T}[\lambda](t) &= -\int_{t_0-1}^{t}\int_{\mathbb{R}^3} \frac{\bar{\alpha}(s)}{(4\pi(t-s))^{\frac{3}{2}}} \frac{e^{-\frac{|y|^2}{4(t-s)}}}{|y|} \frac{\mu(s)}{\mu(s)+|y|} \mathbf{1}_{\{|y|<M\}}\, dz\, ds \\ &= \bar{c}\int_{t_0-1}^{t} \frac{\bar{\alpha}(s)\mu(s)}{\sqrt{t-s}} \int_0^{\frac{M}{\sqrt{t-s}}} e^{-\rho^2} \frac{\rho}{\mu+\rho}\, d\rho\, ds\end{aligned}$$

for some explicit constant $\bar{c}$. Since

$$\left|\int_0^{\frac{M}{\sqrt{t-s}}} e^{-\rho^2}\frac{\rho}{\mu+\rho}\,d\rho\right| \le c\frac{M}{\sqrt{t}}$$

for any t large, we observe that

$$|\widehat{T}[\lambda](t)| \le A\frac{M}{\sqrt{t}}\left|\int_{t_0-1}^{t}\frac{\bar{\alpha}(s)\mu(s)}{\sqrt{t-s}}\,ds\right| \tag{7-7}$$

for some fixed constant A. We claim that

$$\int_{t_0-1}^{t}\frac{\bar{\alpha}(s)\mu(s)}{\sqrt{t-s}}\,ds = \bar{\alpha}(t)\mu(t)\sqrt{t-t_0+1}\,\Pi(t), \quad t > t_0, \tag{7-8}$$

for some smooth and uniformly bounded function $\Pi(t)$. Indeed, we write, for $\beta_*(s) = \bar{\alpha}(s)\mu(s)$,

$$\int_{t_0-1}^{t}\frac{\beta_*(s)}{\sqrt{t-s}}\,ds = \int_{t_0-1}^{t}\frac{\beta_*(s)-\beta_*(t)}{\sqrt{t-s}}\,ds + 2\beta_*(t)\sqrt{t-t_0+1} = i + 2\beta_*(t)\sqrt{t-t_0+1}. \tag{7-9}$$

Using the change of variables $x = \sqrt{t-s}$,

$$i = -2\int_0^{\sqrt{t-t_0+1}}[\beta_*(t)-\beta_*(t-x^2)]\,dx = -2\beta_*(t)\int_0^{\sqrt{t-t_0+1}}\frac{[\beta_*(t)-\beta_*(t-x^2)]}{\beta_*(t)}\,dx.$$

We now observe that the function $x \to [\beta_*(t)-\beta_*(t-x^2)]/\beta_*(t)$ is uniformly bounded in $x \in [0, \sqrt{t-t_0+1}]$, since

$$\frac{[\beta_*(t)-\beta_*(t-x^2)]}{\beta_*(t)} = \begin{cases} 1-(1-x^2/t)^{-1}(1-x^2/t)^{-\frac{3}{2}\bar{\gamma}-1} & \text{for } \gamma \neq 2, \\ 1-(1-x^2/t)^{-\frac{5}{2}}[1+\log(1-x^2/t)]^3 & \text{for } \gamma = 2, \end{cases}$$

where $\bar{\gamma} = 1$ if $\gamma > 2$, and $\bar{\gamma} = \gamma - 1$ if $1 < \gamma < 2$. With this in mind, we conclude that

$$i = \beta_*(t)\sqrt{t-t_0+1}\,\Pi(t) \tag{7-10}$$

for some smooth and bounded function Π. Inserting (7-10) into (7-9), we get (7-8).

Using (7-8) in (7-7), we conclude that

$$|\widehat{T}[\lambda](t)| \le A\mu_0(t)M\|\lambda\|_{\sharp}[\mu_0^{\frac{3}{2}}(t)t^{-1}]$$

for some fixed constant A, independent of t and of M. Thus, for t large, if we choose t_0 sufficiently large, there exists a constant $c_1 \in (0, 1)$ such that

$$|\widehat{T}[\lambda](t)| \le c_1 M\|\lambda\|_{\sharp}[\mu_0^{\frac{3}{2}}(t)t^{-1}].$$

Let now consider t_1 and $t_2 \in [t, t+1]$. We write

$$\widehat{T}[\lambda](t_1) - \widehat{T}[\lambda](t_2) = \bar{c} \int_{t_0-1}^{t_1} \left[\frac{\bar{\alpha}(s)}{\sqrt{t_1 - s}} - \frac{\bar{\alpha}(s)}{\sqrt{t_2 - s}} \right] \int_0^{\frac{M}{\sqrt{t_1-s}}} e^{-\rho^2} \frac{\rho\mu}{\mu+\rho} \, d\rho \, ds$$
$$- \bar{c} \int_{t_0-1}^{t_1} \frac{\bar{\alpha}(s)}{\sqrt{t_2 - s}} \int_{\frac{M}{\sqrt{t_2-s}}}^{\frac{M}{\sqrt{t_1-s}}} e^{-\rho^2} \frac{\rho\mu}{\mu+\rho} \, d\rho \, ds$$
$$- \bar{c} \int_{t_1}^{t_2} \frac{\bar{\alpha}(s)}{\sqrt{t_2 - s}} \int_0^{\frac{M}{\sqrt{t_2-s}}} e^{-\rho^2} \frac{\rho\mu}{\mu+\rho} \, d\rho \, ds = \sum_{j=1}^{3} i_j.$$

Observe that, for $t_1, t_2 \in [t, t+1]$, for t large, we have

$$\sup_{t_1,t_2 \in [t,t+1]} \frac{|\mu(t_1) - \mu(t_2)|}{|t_1 - t_2|^\sigma} \le C\mu_0(t) \quad \sup_{t_1,t_2 \in [t,t+1]} \frac{|\Lambda(t_1) - \Lambda(t_2)|}{|t_1 - t_2|^\sigma} \le CM^{-1}\mu_0(t)(\mu_0^{\frac{3}{2}}(t)t^{-1}) \tag{7-11}$$

for some constant C. With this, we can estimate i_1 and i_5 as follows:

$$[i_j]_{0,\sigma,[t,t+1]} \le CM^{-1}\mu_0(t)(\mu_0^{\frac{3}{2}}(t)t^{-1}) \quad \text{for } j = 1, 5.$$

Straightforward computation gives

$$[i_j]_{0,\sigma,[t,t+1]} \le CM^{-1}\mu_0(t)t^{-\sigma}\|\lambda\|_\sharp(\mu_0^{\frac{3}{2}}(t)t^{-1}) \quad \text{for } j = 1, 2, 3.$$

These estimates, together with the ones we obtained before, constitute the proof of (7-5).

Proof of (7-6). Let $\lambda_1, \lambda_2 \in B$. From (6-2) and (6-3),

$$\widehat{T}[\lambda_1](t) - \widehat{T}[\lambda_2](t) = \bar{c} \int_{t_0-1}^{t} \frac{\bar{\alpha}(s)}{\sqrt{t-s}} \int_0^{\frac{M}{\sqrt{t-s}}} e^{-\rho^2} \left[\frac{\rho\mu[\lambda_1]}{\mu[\lambda_1]+\rho} - \frac{\rho\mu[\lambda_2]}{\mu[\lambda_2]+\rho} \right] d\rho \, ds.$$

Observe that

$$|(\mu[\lambda_1] - \mu[\lambda_2])(s)| \le A\mu_0(s)|\Lambda_1(s) - \Lambda_2(s)|$$
$$\le A\mu_0(s) \int_s^\infty |\lambda_1 - \lambda_2|(x)\, dx \le A\mu_0^2(s)\|\lambda_1 - \lambda_2\|_\sharp$$

for some constant A, whose value may change from one line to the other, and which is independent of t and t_0. A Taylor expansion gives

$$|\widehat{T}[\lambda_1](t) - \widehat{T}[\lambda_2](t)| \le \int_{t_0-1}^{t} \frac{|\bar{\alpha}(s)|}{\sqrt{t-s}} \int_0^{\frac{M}{\sqrt{t-s}}} e^{-\rho^2} \frac{\rho}{(\tilde{\mu}+\rho)^2} |\mu[\lambda_1](s) - \mu[\lambda_2](s)| \, d\rho \, ds$$

for some $\tilde{\mu}$ between $\mu[\lambda_1]$ and $\mu[\lambda_2]$. Thus we get

$$|\widehat{T}[\lambda_1](t) - \widehat{T}[\lambda_2](t)| \le A\mu_0^2(t)M[\mu_0^{\frac{3}{2}}(t)t^{-1}]\|\lambda_1 - \lambda_2\|_\sharp,$$

where A is a constant independent of t_0 and t. Using again (7-11), we can show that

$$|\widehat{T}[\lambda_1] - \widehat{T}[\lambda_2]|_{0,\sigma,[t,t+1]} \le A\mu_0^2(t)M[\mu_0^{\frac{3}{2}}(t)t^{-1}]\|\lambda_1 - \lambda_2\|_\sharp,$$

where A is a constant independent of t_0 and t. Choosing t_0 large enough, we can find $\bar{c}_2$ small enough so that (7-6) holds true.

Step 2: In the second part of the proof, we show the validity of (7-2). For this purpose, we fix ϕ_1 and ϕ_2 satisfying (4-1), and we let $\lambda_j = \lambda[\phi_j]$, $j = 1, 2$. If $\bar{\lambda} = \lambda_1 - \lambda_2$, then we see that $\bar{\lambda}$ solves

$$\begin{aligned}\bar{\lambda} &= T^{-1}(G_1[\lambda_1, \phi_1] - G_1[\lambda_2, \phi_2]) \\ &= T^{-1}(G_1[\bar{\lambda}_1, \phi_1] - G[\bar{\lambda}_1, \phi_2]) + T^{-1}(G_1[\lambda_1, \phi_2] - G[\lambda_2, \phi_2]).\end{aligned}$$

Thus

$$\begin{aligned}\|\bar{\lambda}\|_\sharp &\le CM^{-1}(\|G_1[\bar{\lambda}_1, \phi_1] - G[\bar{\lambda}_1, \phi_2]\|_\flat + \|G_1[\lambda_1, \phi_2] - G[\lambda_2, \phi_2]\|_\flat) \\ &\le CM^{-1}(\boldsymbol{c}\|\phi_1 - \phi_2\|_{\nu,a} + \boldsymbol{c}\|\lambda_1 - \lambda_2\|_\sharp),\end{aligned}$$

where C is the constant in (6-12), $M^2 = t_0$, $\boldsymbol{c}$ are the constants defined respectively in (5-5) and (5-6). We now observe that the proof of Lemma 5.1 also gives that the constants $\boldsymbol{c}$ in (5-5) and (5-6) can be such that $CM^{-1}\boldsymbol{c} < 1$. Thus the proof of (7-2) readily follows. □

Remark 7.2. Recall that the function $\psi = \bar{\Psi}[\psi_0]$, the solution to problem (3-6), depends smoothly on the initial condition ψ_0, provided ψ_0 belongs to a small neighborhood of 0 in the Banach space $L^\infty(\Omega)$ equipped with the norm defined in (4-21), as observed in Remark 4.2. This fact implies that also $\lambda = \lambda[\psi_0]$, the solution to (5-1), depends on ψ_0. A closer look at the definitions of $\lambda = \lambda[\psi_0]$ gives that

$$\|\lambda[\psi_0^{(1)}] - \lambda[\psi_0^{(2)}]\|_\sharp \lesssim \|e^{b|y|}[\psi_0^{(1)} - \psi_0^{(1)}]\|_{L^\infty(\mathbb{R}^3)} + \|e^{b|y|}[\nabla\psi_0^{(1)} - \nabla\psi_0^{(1)}]\|_{L^\infty(\mathbb{R}^3)}.$$

This fact will be useful in the final argument of finding ϕ, the solution to (3-13).

8. Final argument: solving (3-8)

We are constructing a global unbounded solution to problem (2-1)–(2-2) of the form (3-1)

$$u = U_2[\lambda](r, t) + \tilde{\phi}.$$

The function U_2 is defined in (2-57), while $\tilde{\phi}$ is given in (3-2). The function ψ which enters in the definition of $\tilde{\phi}$ solves the *outer problem* (3-6), and its properties are contained in Propositions 4.1 and 4.3. The parameter $\lambda = \lambda(t)$ belongs to the space $X_\sharp$, (2-12), and has been chosen to solve (5-1). The properties of this $\lambda = \lambda(t)$ are collected in Proposition 7.1. What is left is to solve the *inner problem* (3-8) in ϕ. Thanks to the choice of $\lambda = \lambda(t)$, the orthogonality condition (3-19) is satisfied, so that we can use the result of Proposition 3.1 to solve problem (3-8) in ϕ.

In other words, we want to find ϕ, with its $\|\phi\|_{\nu,a}$-bounded, solution to problem (3-8). The function $\psi = \Psi[\lambda[\phi], \phi]$ solves (3-6), while $\lambda = \lambda[\phi]$ solves (5-1).

At this point, we fix a in the definition of $\|\cdot\|_{\nu,a}$ to be equal to 1. Proposition 3.1 defines a linear operator $\phi = \mathcal{T}(h)$, where ϕ is the solution to (3-16) so that

$$\|\phi\|_{\nu,1} \le C_0 R^4 \|h\|_{\nu,3}$$

for some fixed constant C_0. We refer to (3-17) for $\|h\|_{\nu,2+a}$ and to (3-18) for $\|\phi\|_{\nu,a}$, for $a = 1$. Thus we can say that ϕ solves (3-11)–(3-13) if and only if ϕ is a fixed point for the problem

$$\phi = \mathcal{T}(\boldsymbol{H}[\phi]), \quad \text{where } \boldsymbol{H}[\phi] = H(\psi[\phi], \lambda[\phi], \phi), \tag{8-1}$$

and H is defined in (3-12). Choose the number R in the cut-off function η_R, defined in (2-59) and appearing in the ansatz (3-2), to be sufficiently large in terms of t_0, say $R^6 \mu_0^{\frac{1}{2}}(t_0) = 1$. We claim that there exists a unique ϕ solution to (8-1) in the set

$$B_1 = \{\phi : \|\phi\|_{\nu,1} \le L_1\}$$

for some $L_1 > 0$, fixed.

From (2-59) and (4-7), we see that

$$|\mu_0^{\frac{5}{2}} \mathcal{E}_{22}(\mu_0 y, t)| \lesssim \mu_0^{\frac{1}{2}} \frac{\mu_0^{\frac{3}{2}} t^{-1}}{(1+|y|^2)^2}, \qquad \left| 5 \frac{\mu_0^{\frac{1}{2}}}{(1+\Lambda)^4} w^4\left(\frac{y}{(1+\Lambda)^2}\right) \psi(\mu_0 y, t) \right| \lesssim \frac{\mu_0^2(t) t^{-1}}{(1+|y|^3)}.$$

Furthermore,

$$|B[\phi](t)| \le C R^2 \mu_0 \mu_0' \frac{\mu_0^{\frac{3}{2}} t^{-1}}{(1+|y|^{2+a})}, \qquad |B^0[\phi](t)| \le C \Lambda(t) \frac{\mu_0^{\frac{3}{2}} t^{-1}}{(1+|y|^{4+a})}.$$

In fact, one can prove that

$$\|\boldsymbol{H}[\phi]\|_{\nu,2+a} \le C_1 R^{-4}$$

for some fixed number C_1, independent from t and of t_0. This implies that, if $\phi \in B_1$, then $\mathcal{T}(\phi) \in B_1$, provided L_1 is chosen large. Furthermore, combining (2-61), the result of Proposition 4.3, and the result of Proposition 7.1, we get the existence of a number $\boldsymbol{c} \in (0,1)$ so that

$$\|\mathcal{T}[\phi_1] - \mathcal{T}[\phi_2]\|_{\nu,a} \le \boldsymbol{c} \|\phi_1 - \phi_2\|_{\nu,a}$$

for any ϕ_1 and $\phi_2 \in B_1$. We apply Banach fixed-point theorem to get the existence of a unique solution to (8-1) with $\|\cdot\|_{\nu,a}$-bounded.

This concludes the proof of the existence of the solution to problem (2-1)–(2-2), or equivalently problem (1-3)–(1-4), as predicted by Theorem 1.1. □

9. Basic linear theory for the inner problem

Let $R > 0$ be a fixed large number. This section is devoted to constructing a solution to the initial value problem

$$\phi_\tau = \Delta\phi + 5w^4\phi + h(y,\tau) \quad \text{in } B_{2R} \times (\tau_0, \infty), \qquad \phi(y, \tau_0) = e_0 Z(y) \quad \text{in } B_{2R}, \tag{9-1}$$

for any given function h with $\|h\|_{\nu,2+a} < +\infty$, not necessarily radial in the y-variable. We refer to (3-17) for the explicit definition of the $\|\cdot\|_{\nu,2+a}$ norm. The corresponding problem in dimension $n \ge 5$ has already been treated in [Cortázar et al. 2016, Section 7]. We follow the same strategy in the procedure to construct the solution to (9-1), but in dimension 3 we get a decay estimate for the solution different from the one valid for dimensions $n \ge 5$.

We recall that the operator $L_0(\phi) = \Delta\phi + 5w^4\phi$ has a 4-dimensional kernel generated by the bounded functions Z_0 defined in (2-7) and also by

$$Z_i(y) = \frac{\partial w}{\partial y_i}, \quad i = 1, 2, 3. \tag{9-2}$$

In the class of radially symmetric functions, the only element in the kernel of L_0 is Z_0. To describe our construction, we consider an orthonormal basis ϑ_m, $m = 0, 1, \dots,$ in $L^2(S^2)$ of spherical harmonics, namely eigenfunctions of the problem

$$\Delta_{S^2}\vartheta_m + \lambda_m \vartheta_m = 0 \quad \text{in } S^2$$

so that $0 = \lambda_0 < \lambda_1 = \cdots = \lambda_3 = 2 < \lambda_4 \le \cdots$. Let $h(\cdot, \tau) \in L^2(B_{2R})$ for any $\tau \in [\tau_0, \infty)$. We decompose it into the form

$$h(y,\tau) = \sum_{j=0}^{\infty} h_j(r,\tau)\vartheta_j\left(\frac{y}{r}\right), \quad r = |y|, \quad h_j(r,\tau) = \int_{S^2} h(r\theta,\tau)\vartheta_j(\theta)\, d\theta.$$

In addition, we write $h = h^0 + h^1 + h^\perp$, where

$$h^0 = h_0(r,\tau), \quad h^1 = \sum_{j=1}^{3} h_j(r,\tau)\vartheta_j, \quad h^\perp = \sum_{j=4}^{\infty} h_j(r,\tau)\vartheta_j.$$

Observe that $h^1 = h^\perp = 0$ if h is radially symmetric in the y-variable. Consider also the analogous decomposition for ϕ into $\phi = \phi^0 + \phi^1 + \phi^\perp$. We build the solution ϕ of problem (9-1) by doing so separately for the pairs (ϕ^0, h^0), (ϕ^1, h^1) and $(\phi^\perp, h^\perp)$.

Our main result in this section is the following proposition.

Proposition 9.1. *Let ν, a be given positive numbers with $0 < a < 2$. Then, for all sufficiently large $R > 0$ and any $h = h(y,\tau)$ with $\|h\|_{\nu,2+a} < +\infty$ that satisfies for all $j = 0, 1, \dots, 3$*

$$\int_{B_{2R}} h(y,\tau) Z_j(y)\, dy = 0 \quad \textit{for all } \tau \in (\tau_0, \infty), \tag{9-3}$$

there exist $\phi = \phi[h]$ and $e_0 = e_0[h]$ which solve problem (9-1). They define linear operators of h that satisfy the estimates

$$|\phi(y,\tau)| \lesssim \tau^{-\nu}\left[\frac{R^{4-a}}{1+|y|^3}\|h^0\|_{\nu,2+a} + \frac{R^{4-a}}{1+|y|^4}\|h^1\|_{\nu,2+a} + \frac{\|h\|_{\nu,2+a}}{1+|y|^a}\right], \tag{9-4}$$

$$|\nabla_y\phi(y,\tau)| \lesssim \tau^{-\nu}\left[\frac{R^{4-a}}{1+|y|^4}\|h^0\|_{\nu,2+a} + \frac{R^{4-a}}{1+|y|^5}\|h^1\|_{\nu,2+a} + \frac{\|h\|_{\nu,2+a}}{1+|y|^{a+1}}\right], \tag{9-5}$$

and

$$|e_0[h]| \lesssim \|h\|_{\nu,2+a}. \tag{9-6}$$

Proposition 3.1 is a direct consequence of Proposition 9.1. Indeed, if h is radially symmetric in the y-variable, (9-3) is automatically satisfied for $j = 1, \dots, 3$, and $h \equiv h^0$.

The result contained in Proposition 9.1 follows from the next proposition, which refers to the problem

$$\phi_\tau = \Delta\phi + 5w(y)^4\phi + h(y,\tau) - c(\tau)Z \quad \text{in } B_{2R}\times(\tau_0,\infty), \qquad \phi(y,\tau_0) = 0 \quad \text{in } B_{2R}. \tag{9-7}$$

Proposition 9.2. *Let ν, a be given positive numbers with $0 < a < 2$. Then, for all sufficiently large $R > 0$ and any h with $\|h\|_{\nu,2+a} < +\infty$ and satisfying the orthogonality conditions (3-19), there exist $\phi = \phi[h]$ and $c = c[h]$ which solve problem (9-7), and define linear operators of h. The function $\phi[h]$ satisfies estimate (9-4), (9-5) and for some $\Gamma > 0$*

$$\left| c(\tau) - \int_{B_{2R}} hZ \right| \lesssim \tau^{-\nu}\left[R^{2-a}\left\| h - Z\int_{B_{2R}} hZ \right\|_{\nu,2+a} + e^{-\Gamma R}\|h\|_{\nu,2+\alpha}\right]. \tag{9-8}$$

Assuming the validity of Proposition 9.2, we proceed with:

Proof of Proposition 9.1. Let ϕ_1 be the solution of problem (9-7) predicted by Proposition 9.2. Let us write

$$\phi(y,\tau) = \phi_1(y,\tau) + e(\tau)Z(y) \tag{9-9}$$

for some $e \in C^1([\tau_0,\infty))$. We find

$$\partial_\tau\phi = \Delta\phi + 5w^4\phi + h(y,\tau) + [e'(\tau) - \lambda_0 e(\tau) - c(\tau)]Z(y).$$

We choose $e(\tau)$ to be the unique bounded solution of the equation

$$e'(\tau) - \lambda_0 e(\tau) = c(\tau), \qquad \tau \in (\tau_0,\infty),$$

which is explicitly given by

$$e(\tau) = \int_\tau^\infty \exp(\sqrt{\lambda_0}(\tau - s))c(s)\,ds.$$

The function e depends linearly on h. Besides, we clearly have from (9-8), $|e(\tau)| \lesssim \tau^{-\nu}\|h\|_{\nu,2+a}$, and thus, from the fact that ϕ_1 satisfies estimates (9-4), (9-5), so does ϕ given by (9-9). Thus ϕ satisfies problem (9-1) with initial condition $\phi(y,\tau_0) = e(\tau_0)Z(y)$. □

The rest of the section is devoted to the following:

Proof of Proposition 9.2. The proof is divided in two steps. In the first step, we construct a solution to (9-7) which has value zero on the boundary ∂B_{2R}, at any time τ, for a right-hand side h not necessarily satisfying the orthogonality conditions (9-3). In the second step, we make use of this construction to solve (9-7), for a right-hand side satisfying (9-3), and to obtain estimates (9-4), (9-5) and (9-6).

Step 1: We claim that for all sufficiently large $R > 0$ and any H with $\|H\|_{\nu,a} < +\infty$ there exist $\phi = \phi(y,\tau)$ and $c = c(\tau)$ which solve problem

$$\begin{aligned} &\phi_\tau = \Delta\phi + 5w^4\phi + H(y,\tau) - c(\tau)Z(y) \quad \text{in } B_{2R}\times(\tau_0,\infty),\\ &\phi = 0 \quad \text{on } \partial B_{2R}\times(\tau_0,\infty), \qquad \phi(\cdot,\tau_0) = 0 \quad \text{in } B_{2R}. \end{aligned} \tag{9-10}$$

The functions ϕ and c are linear operators of h and satisfy the estimates

$$(1+|y|)|\nabla\phi(y,\tau)|+|\phi(y,\tau)|\lesssim\tau^{-\nu}\left[\frac{R^{4-a}\|H^0\|_{\nu,a}}{1+|y|}+\frac{R^{4-a}\|H^1\|_{\nu,a}}{1+|y|^2}+R^2\frac{\|H\|_{\nu,a}}{1+|y|^a}\right] \tag{9-11}$$

and for some $\Gamma>0$

$$\left|c(\tau)-\int_{B_{2R}}HZ\right|\lesssim\tau^{-\nu}\left[R^2\left\|H-Z\int_{B_{2R}}HZ\right\|_{\nu,a}+e^{-\Gamma R}\|H\|_{\nu,a}\right]. \tag{9-12}$$

We construct the solution ϕ mode by mode, considering first mode 0, then modes $1,2,3$ and finally modes greater or equal to 4. For each mode, we get the corresponding estimates.

Construction at mode 0. Consider problem (9-10) for a right-hand side $H=H_0(r,\tau)$ radially symmetric. Let $\eta(s)$ be the smooth cut-off function in (1-9), and consider $\eta_\ell(y)=\eta(|y|-\ell)$ for a large but fixed number ℓ independent of R. By standard parabolic theory, there exists a unique solution $\phi_*[\bar{h}_0]$ to

$$\begin{aligned}\phi_\tau=\Delta\phi+5w(r)^4(1-\eta_\ell)\phi+\overline{H}_0(y,\tau)\quad&\text{in } B_{2R}\times(\tau_0,\infty),\\ \phi=0\quad\text{on } \partial B_{2R}\times(\tau_0,\infty),\qquad \phi(\,\cdot\,,\tau_0)=0\quad&\text{in } B_{2R},\end{aligned} \tag{9-13}$$

where

$$\overline{H}_0=H_0-c_0(\tau)Z,\quad c_0(\tau)=\int_{B_{2R}}H_0(y,\tau)Z(y)\,dy.$$

The function $\phi_*[\bar{h}_0]$ is radial and satisfies the bound

$$|\phi_*[\overline{H}_0]|\lesssim\tau^{-\nu}R^{2-a}\|H\|_{\nu,a}.$$

This can be proven with the use of a special supersolution arguing as in Lemma 7.3 in [Cortázar et al. 2016]. Setting $\phi=\phi_*[\overline{H}_0]+\tilde{\phi}$ and $c(\tau)=c_0(\tau)+\tilde{c}(\tau)$, problem (9-10) gets reduced to

$$\begin{aligned}\tilde{\phi}_\tau=\Delta\tilde{\phi}+5w(r)^4\tilde{\phi}+\widetilde{H}_0(r,\tau)-\tilde{c}(\tau)Z\quad&\text{in } B_{2R}\times(\tau_0,\infty),\\ \tilde{\phi}=0\quad\text{on } \partial B_{2R}\times(\tau_0,\infty),\qquad \tilde{\phi}(\,\cdot\,,\tau_0)=0\quad&\text{in } B_{2R}.\end{aligned} \tag{9-14}$$

where $\widetilde{H}_0=5w^4\eta_\ell\phi_*[\overline{H}_0]$. Observe that $\widetilde{H}_0$ is radial, compactly supported and with size controlled by that of $\overline{H}_0$. In particular we have that for any $m>0$

$$|\widetilde{H}_0(r,\tau)|\lesssim\frac{\tau^{-\nu}}{1+r^m}\Big[\sup_{\tau>\tau_0}\tau^\nu\|\phi_*[\overline{H}_0](\,\cdot\,,\tau)\|_{L^\infty}\Big]\lesssim\frac{\tau^{-\nu}}{1+r^m}R^{2-a}\|H\|_{\nu,a}. \tag{9-15}$$

We shall next solve problem (9-14) under the additional orthogonality constraint

$$\int_{B_{2R}}\tilde{\phi}(\,\cdot\,,\tau)Z=0\quad\text{for all } \tau\in(\tau_0,\infty). \tag{9-16}$$

Problem (9-14)–(9-16) is equivalent to solving just (9-14) for $\tilde{c}$ given by the explicit linear functional $\tilde{c}:=\tilde{c}[\tilde{\phi},\widetilde{H}_0]$ determined by the relation

$$\tilde{c}(\tau)\int_{B_{2R}}Z^2=\int_{B_{2R}}\widetilde{H}_0(\,\cdot\,,\tau)Z+\int_{\partial B_{2R}}\partial_r\tilde{\phi}(\,\cdot\,,\tau)Z. \tag{9-17}$$

If the function $\tilde{c} = \tilde{c}(\tau)$ defined by (9-17) were independent of ϕ, standard linear parabolic theory would give the existence of a unique solution. On the other hand, a close look at (9-17) shows that the dependence of $\tilde{c} = \tilde{c}(\tau)$ on ϕ is small in an L^∞-$C^{1+\alpha,\frac{1+\alpha}{2}}$ setting, since $Z(R) = O(e^{-\Gamma R})$ for some $\Gamma > 0$. A contraction argument applies to yield existence of a unique solution to (9-14)–(9-16) defined at all times. To get the estimates, we assume smoothness of the data so that integrations by parts and differentiations can be carried over, and then argue by approximations. Testing (9-14)–(9-16) against $\tilde{\phi}$ and integrating in space, we obtain the relation

$$\partial_\tau \int_{B_{2R}} \tilde{\phi}^2 + Q(\tilde{\phi}, \tilde{\phi}) = \int_{B_{2R}} g\tilde{\phi}, \quad g = \tilde{H}_0 - \tilde{c}(\tau) Z_0,$$

where Q is the quadratic form defined by

$$Q(\phi, \phi) := \int [|\nabla \phi|^2 - 5w^4 |\phi|^2]. \tag{9-18}$$

Since dimension is 3, there exists $\beta > 0$ such that, for any ϕ with $\int \phi Z = 0$, the following inequality holds:

$$Q(\phi, \phi) \geq \frac{\beta}{R^2} \int \phi^2.$$

The proof of this inequality is a slight modification of the proof for the corresponding inequality in dimensions $n \geq 5$ that can be found in [Cortázar et al. 2016, Lemma 7.2], considering that $\int_{B_R} Z_0^2 = O(R)$ as $R \to \infty$, when dimension is 3. Thus we have, for some $\beta' > 0$,

$$\partial_\tau \int_{B_{2R}} \tilde{\phi}^2 + \frac{\beta'}{R^2} \int_{B_{2R}} \tilde{\phi}^2 \lesssim R^2 \int_{B_{2R}} g^2. \tag{9-19}$$

We observe that from (9-17) and (9-15) for $m = 0$ we get

$$|\tilde{c}(\tau)| \leq \tau^{-\nu} K, \quad K := \big[\sup_{\tau > \tau_0} \tau^\nu \|\phi_*[\overline{H}_0](\cdot, \tau)\|_{L^\infty}\big] + e^{-\Gamma R}\big[\sup_{\tau > \tau_0} \tau^\nu \|\nabla \phi_*[\overline{H}_0](\cdot, \tau)\|_{L^\infty}\big].$$

Besides, using again estimate (9-15) for a sufficiently large m, we get

$$\int_{B_{2R}} g^2 \lesssim \tau^{-2\nu} K^2.$$

Using that $\tilde{\phi}(\cdot, \tau_0) = 0$ and Gronwall's inequality, we readily get from (9-19) the L^2-estimate

$$\|\tilde{\phi}(\cdot, \tau)\|_{L^2(B_{2R})} \lesssim \tau^{-\nu} R^2 K \tag{9-20}$$

for all $\tau > \tau_0$. Now, using standard parabolic estimates in the equation satisfied by $\tilde{\phi}$ we obtain then that on any large fixed radius $\ell > 0$,

$$\|\tilde{\phi}(\cdot, \tau)\|_{L^\infty(B_M)} \lesssim \tau^{-\nu} R^2 K \quad \text{for all } \tau > \tau_0.$$

Since the right-hand side has a fast decay at infinity and taking into account that we are in dimension 3, outside B_ℓ we can dominate the solution by a barrier of the order $\tau^{-\nu}|y|^{-1}$. As a conclusion, also using local parabolic estimates for the gradient, we find that

$$(1+|y|)|\nabla_y\tilde{\phi}(y,\tau)|+|\tilde{\phi}(y,\tau)| \lesssim \tau^{-\nu}\frac{R^2}{1+|y|}\Big[\sup_{\tau>\tau_0}\tau^{\nu}\|\phi_*[\overline{H}_0](\cdot,\tau)\|_{L^\infty}\Big]. \tag{9-21}$$

It clearly follows from this estimate and inequality (9-15) that the function

$$\phi_0[h_0] := \tilde{\phi}+\phi_*[\overline{H}_0] \tag{9-22}$$

solves problem (9-10) for $H = H_0$ and satisfies

$$(1+|y|)|\nabla_y\phi_0(y,\tau)|+|\phi_0(y,\tau)| \lesssim \tau^{-\nu}\frac{R^{4-a}}{1+|y|}\|H\|_{\nu,a}.$$

Finally, from (9-17) we see that

$$c(\tau)=\int_{B_{2R}} HZ+\int_{B_{2R}} 5w^4\eta_\ell\phi_*[\overline{H}_0]Z+O(e^{-\Gamma R})\|H\|_{\nu,a}.$$

From here we find the validity of the estimate

$$\left|c(\tau)-\int_{B_{2R}} H_0Z\right| \lesssim \tau^{-\nu}\left[R^2\left\|H_0-Z\int_{B_{2R}} H_0Z\right\|_{\nu,a}+e^{-\Gamma R}\|H_0\|_{\nu,a}\right].$$

Hence estimates (9-11) and (9-12) hold. The construction of the solution at mode 0 is concluded.

Construction at modes 1 *to* 3. Here we consider the case $H = H^1$, where $H^1(y,\tau)=\sum_{j=1}^3 H_j(r,\tau)\vartheta_j$. The function

$$\phi^1[H^1] := \sum_{j=1}^{n}\phi_j(r,\tau)\vartheta_j \tag{9-23}$$

solves the initial-boundary value problem

$$\begin{aligned} \phi_\tau &= \Delta\phi+5w^4\phi+H^1(y,\tau) \quad \text{in } B_{2R}\times(\tau_0,\infty),\\ \phi &= 0 \quad \text{on } \partial B_{2R}\times(\tau_0,\infty), \qquad \phi(\cdot,\tau_0)=0 \quad \text{in } B_{2R}, \end{aligned} \tag{9-24}$$

if the function $\phi_j(r,\tau)$ solves

$$\begin{aligned} \partial_\tau\phi_j &= \mathcal{L}_1[\phi_j]+H_j(r,\tau) \quad \text{in } (0,2R)\times(\tau_0,\infty),\\ \partial_r\phi_j(0,\tau) &= 0=\phi_j(R,\tau) \quad \text{for all } \tau\in(\tau_0,\infty), \qquad \phi_j(r,\tau_0)=0 \quad \text{for all } r\in(0,R), \end{aligned} \tag{9-25}$$

where

$$\mathcal{L}_1[\phi_j] := \partial_{rr}\phi_j+2\frac{\partial_r\phi_j}{r}-2\frac{\phi_j}{r^2}+5w^4\phi_j. \tag{9-26}$$

Let us consider the solution of the stationary problem $\mathcal{L}_1[\phi]+(1+r)^{-a}=0$ given by the variation of parameters formula

$$\bar{\phi}(r)=Z(r)\int_r^{2R}\frac{1}{\rho^2Z(\rho)^2}\int_0^{\rho}(1+s)^{-a}Z(s)s^2\,ds,$$

where $Z(r)=w_r(r)$. Since $w_r(r)\sim r^{-2}$ for large r, we find the estimate $|\bar{\phi}(r)|\lesssim(R^{4-a})/(1+r^2)$. Then, provided that τ_0 was chosen sufficiently large, the function $2\|H_j\|_{\nu,a}\tau^{-\nu}\bar{\phi}(r)$ is a positive supersolution of problem (9-25) and thus we find $|\phi_j(r,\tau)| \lesssim \tau^{-\nu}((R^{4-a})/(1+r^2))\|H_j\|_{\nu,a}$. Hence $\phi^1[H^1]$ given by (9-23) satisfies

$$|\phi^1[H^1](y,\tau)| \lesssim \frac{R^{4-a}}{1+|y|^2}\|H^1\|_{\nu,a}.$$

A corresponding estimate for the gradient follows.

Construction at higher modes. We consider now the case of higher modes,

$$\begin{aligned} &\phi_\tau = \Delta\phi + 5w^4\phi + H^\perp \quad \text{in } B_{2R}\times(\tau_0,\infty),\\ &\phi = 0 \quad \text{on } \partial B_{2R}\times(\tau_0,\infty), \qquad \phi(\,\cdot\,,\tau_0) = 0 \quad \text{in } B_{2R}, \end{aligned} \tag{9-27}$$

where $H = H^\perp = \sum_{j=4}^\infty H_j(r)\Theta_j$ whose solution has the form $\phi^\perp = \sum_{j=4}^\infty \phi_j(r,\tau)\Theta_j$. Given the quadratic form in (9-18), for $\phi^\perp \in H_0^1(B_{2R})$,

$$\int_{B_{2R}} \frac{|\phi^\perp|^2}{r^2} \lesssim Q(\phi^\perp,\phi^\perp). \tag{9-28}$$

The proof of this fact is elementary. The interested reader can find it in [Cortázar et al. 2016]. Let $\phi_*[H^\perp]$ be the solution to

$$\begin{aligned} &\phi_\tau = \Delta\phi + 5w(r)^4(1-\eta_\ell)\phi + \overline{H}^\perp(y,\tau) \quad \text{in } B_{2R}\times(\tau_0,\infty),\\ &\phi = 0 \quad \text{on } \partial B_{2R}\times(\tau_0,\infty), \qquad \phi(\,\cdot\,,0) = 0 \quad \text{in } B_{2R}, \end{aligned}$$

where $\overline{H}^\perp = H^\perp - c^\perp Z$, and $c^\perp = \int_{B_{2R}} H^\perp Z$. By writing $\phi = \phi_*[H^\perp] + \tilde\phi$, problem (9-27) reduces to solving

$$\begin{aligned} &\tilde\phi_\tau = \Delta\tilde\phi + 5w(y)^4\tilde\phi + \tilde H \quad \text{in } B_{2R}\times(\tau_0,\infty),\\ &\tilde\phi = 0 \quad \text{on } \partial B_{2R}\times(\tau_0,\infty), \qquad \tilde\phi(\,\cdot\,,\tau_0) = 0 \quad \text{in } B_{2R}, \end{aligned} \tag{9-29}$$

where $\tilde H = 5w(y)^4\eta_\ell\phi_*[H^\perp]$ for a sufficiently large ℓ. Arguing as in (9-19) we now get

$$\partial_\tau \int_{B_{2R}} \tilde\phi^2 + c\int_{B_{2R}} \frac{|\tilde\phi|^2}{|y|^2} \lesssim \int_{B_{2R}} |y|^2|\tilde H|^2. \tag{9-30}$$

Similarly to (9-20) we get

$$\| |y|^{-1}\tilde\phi(\,\cdot\,,\tau)\|_{L^2(B_{2R})} \lesssim \tau^{-\nu}R^{2-a}\|H\|_{\nu,a}. \tag{9-31}$$

From elliptic estimates we then get

$$\|\tilde\phi(\,\cdot\,,\tau)\|_{L^\infty(B_{2R})} \lesssim \tau^{-\nu}R^{2-a}\|H^\perp\|_{\nu,a} \quad \text{for all } \tau > \tau_0,$$

so that with the aid of a barrier we obtain

$$|\tilde\phi(y,\tau)| \lesssim \tau^{-\nu}R^{2-a}\|H^\perp\|_{\nu,a}(1+|y|)^{-1}.$$

It follows that the function

$$\phi^\perp[H^\perp] := \tilde\phi + \phi_*[H^\perp] \tag{9-32}$$

satisfies

$$|\phi^\perp[H^\perp](y,\tau)| \lesssim \tau^{-\nu} R^2[(1+|y|)^{-1} + (1+|y|)^{-a}]\|H^\perp\|_{\nu,a} \quad \text{in } B_{2R}.$$

Similar estimates for the gradient follow.

Conclusion. Let

$$\phi[h] := \phi^0[h^0] + \phi^1[h^1] + \phi^\perp[h^\perp]$$

for the functions defined in (9-22), (9-23), (9-32). By construction, $\phi[h]$ solves (9-10). It defines a linear operator of h and satisfies (9-11). The proof of Step 1 is concluded.

Step 2: To complete the proof of Proposition 9.2, we decompose the right-hand side h in (9-7) in modes, $h = h^0 + h^1 + h^\perp$ as before, and define separately associated solutions of (9-7) in a decomposition $\phi = \phi^0 + \phi^1 + \phi^\perp$.

Construction at mode 0. For a bounded radial $h = h(|y|)$ defined in B_{2R} with $\int_{B_{2R}} hZ_0 = 0$, let $\tilde h$ designate the extension of h as zero outside B_{2R}. The equation

$$\Delta H + 5w^4(y)H + \tilde h(|y|) = 0 \quad \text{in } \mathbb{R}^3, \qquad H(y) \to 0 \quad \text{as } |y| \to \infty,$$

has a solution $H =: L_0^{-1}[h]$ represented by the variation of parameters formula

$$H(r) = \tilde Z(r) \int_r^\infty \tilde h(s) Z_0(s) s^2\, ds + Z_0(r) \int_r^\infty \tilde h(s) \tilde Z(s) s^2\, ds, \tag{9-33}$$

where $\tilde Z(r)$ is a suitable second radial solution of $L_0[\tilde Z] = 0$, linearly independent with Z_0. The mode-0 function $h_0 = h_0(|y|,\tau)$ is defined in B_{2R} and satisfies $\|h_0\|_{\nu,2+a} < +\infty$ and $\int_{B_{2R}} h_0 Z_0 = 0$ for all τ. Then $H_0 := L_0^{-1}[h_0(\,\cdot\,,\tau)]$ satisfies the estimate

$$|H_0(r,\tau)| \lesssim \frac{\tau^{-\nu}}{(1+r)^a} \|h_0\|_{\nu,2+a}.$$

Let $\Phi_0[h_0]$ be the radial solution in B_{3R} to

$$\begin{aligned} \Phi_\tau &= \Delta\Phi + 5w^4(y)\Phi + H_0(|y|,\tau) - c_0(\tau) Z \quad \text{in } B_{3R} \times (\tau_0,\infty), \\ \Phi &= 0 \quad \text{on } \partial B_{3R} \times (\tau_0,\infty), \qquad \Phi(\,\cdot\,,\tau_0) = 0 \quad \text{in } B_{3R} \end{aligned} \tag{9-34}$$

that we discussed in Step 1. The solution $\Phi_0[h_0]$ defines a linear operator of h_0 and satisfies the estimates

$$|\Phi_0(y,\tau)| \lesssim \frac{\tau^{-\nu} R^{4-a}}{(1+|y|)} \|H_0\|_{\nu,a}, \tag{9-35}$$

where for some $\Gamma > 0$

$$\left| c_0(\tau) - \int_{B_{2R}} H_0 Z \right| \lesssim \tau^{-\nu} \left[R^2 \left\| H_0 - Z \int_{B_{2R}} H_0 Z \right\|_{\nu,a} + e^{-\Gamma R} \|H_0\|_{\nu,a} \right]. \tag{9-36}$$

Since $L_0[Z] = \lambda_0 Z$,

$$\lambda_0 \int_{B_{2R}} H_0 Z = \int_{B_{2R}} H_0 L_0[Z] = \int_{B_{2R}} L_0[H_0] Z + \int_{\partial B_{2R}} (Z \partial_\nu H_0 - H_0 \partial_\nu Z),$$

and hence

$$\int_{B_{2R}} H_0 Z = \lambda_0^{-1} \int_{B_{2R}} h_0 Z + O(e^{-\Gamma R})\tau^{-\nu} \|h_0\|_{\nu,2+a}.$$

Also, from the definition of the operator L_0^{-1} we see that $Z = \lambda_0 L_0^{-1}[Z]$. Thus

$$\begin{aligned}\left\| H_0 - Z \int_{B_{2R}} H_0 Z \right\|_{\nu,a} &= \left\| L_0^{-1}\left[h_0 - \lambda_0 Z \int_{B_{2R}} H_0 Z \right]\right\|_{\nu,a} \\ &\lesssim \left\| h_0 - Z \int_{B_{2R}} h_0 Z \right\|_{\nu,2+a} + e^{-\Gamma R} \|h_0\|_{\nu,2+a}.\end{aligned}$$

Next, we discuss estimates on the first and second derivatives of Φ_0. Let us fix now a vector e with $|e| = 1$, a large number $\rho > 0$ with $\rho \le 2R$ and a number $\tau_1 \ge \tau_0$. Consider the change of variables

$$\Phi_\rho(z,t) := \Phi_0(\rho e + \rho z, \tau_1 + \rho^2 t), \quad H_\rho(z,t) := \rho^2[H_0(\rho e + \rho z, \tau_1 + \rho^2 t) - c_0(\tau_1 + \rho^2 t) Z(\rho e + \rho z)].$$

Then $\Phi_\rho(z,t)$ satisfies an equation of the form

$$\partial_t \Phi_\rho = \Delta_z \Phi_\rho + B_\rho(z,t)\Phi_\rho + H_\rho(z,t) \quad \text{in } B_1(0) \times (0,2),$$

where $B_\rho = O(\rho^{-2})$ uniformly in $B_2(0) \times (0,\infty)$. Standard parabolic estimates yield that for any $0 < \alpha < 1$

$$\|\nabla_z \Phi_\rho\|_{L^\infty(B_{1/2}(0)\times(1,2))} \lesssim \|\Phi_\rho\|_{L^\infty(B_1(0)\times(0,2))} + \|H_\rho\|_{L^\infty(B_1(0)\times(0,2))}.$$

Moreover

$$\|H_\rho\|_{L^\infty(B_1(0)\times(0,2))} \lesssim \rho^{2-a}\tau_1^{-\nu}\|H_0\|_{\nu,a}, \quad \|\Phi_\rho\|_{L^\infty(B_1(0)\times(0,2))} \lesssim \tau_1^{-1} K(\rho),$$

where

$$K(\rho) = \frac{R^{2-a}}{1+\rho} R^2 \|h^0\|_{\nu,2+a}. \tag{9-37}$$

This yields in particular that

$$\rho|\nabla_y \Phi(\rho e, \tau_1 + \rho^2)| = |\nabla\tilde{\phi}(0,1)| \lesssim \tau_1^{-\nu} K(\rho).$$

Hence if we choose $\tau_0 \ge R^2$, we get that for any $\tau > 2\tau_0$ and $|y| \le 3R$

$$(1+|y|)|\nabla_y \Phi(y,\tau)| \lesssim \tau^{-\nu} K(|y|). \tag{9-38}$$

We obtain that these bounds are as well valid for $\tau < 2\tau_0$ by the use of similar parabolic estimates up to the initial time (with condition 0).

Now, we observe that the function H_0 is of class C^1 in the variable y and $\|\nabla_y H_0\|_{\nu,1+a} \le \|h^0\|_{\nu,2+a}$. It follows that we have the estimate

$$(1+|y|^2)|D_y^2 \Phi(y,\tau)| \lesssim \tau^{-\nu} K(|y|)$$

for all $\tau > \tau_0$, $|y| \le 2R$, where K is the function in (9-37). The proof follows simply by differentiating the equation satisfied by Φ, rescaling in the same way we did to get the gradient estimate, and applying the bound already proven for $\nabla_y \Phi$. Thus we have in B_{2R}

$$(1+|y|^2)|D^2\Phi(y,\tau)| + (1+|y|)|\nabla\Phi(y,\tau)| + |\Phi(y,\tau)| \lesssim \tau^{-\nu}\|h^0\|_{\nu,2+a}\frac{R^{4-a}}{1+|y|}.$$

This yields in particular

$$|L_0[\Phi](\,\cdot\,,\tau)| \lesssim \tau^{-\nu}\|h^0\|_{\nu,2+a}\frac{R^{4-a}}{1+|y|^3} \quad \text{in } B_{2R}.$$

We define

$$\phi^0[h_0] := L_0[\Phi]\big|_{B_{2R}}.$$

Then $\phi^0[h_0]$ solves problem (9-7) with

$$c(\tau) := \lambda_0 c_0(\tau), \tag{9-39}$$

$\phi^0[h_0]$ satisfies the estimate

$$|\phi^0[h_0](y,\tau)| \lesssim \tau^{-\nu}\|h_0\|_{\nu,2+a}\frac{R^{4-a}}{1+|y|^3} \quad \text{in } B_{2R}, \tag{9-40}$$

and from (9-36), estimate (9-8) holds too.

Construction for modes 1 *to* 3. We consider now $h^1(y,\tau) = \sum_{j=1}^3 h_j(r,\tau)\vartheta_j$, with $\|h^1\|_{\nu,2+a} < +\infty$, that satisfies, for all $i = 1,\dots,3$, $\int_{B_{2R}} h^1 Z_i = 0$ for all $\tau \in (\tau_0,\infty)$. We will show that there is a solution

$$\phi^1[h^1] = \sum_{j=1}^3 \phi_j(r,\tau)\vartheta_j\left(\frac{y}{r}\right)$$

to problem (9-7) for $h = h^1$ which defines a linear operator of h^1 and satisfies the estimate

$$|\phi^1(y,\tau)| \lesssim \frac{R^4}{1+|y|^4}R^{-a}\|h\|_{\nu,2+a}. \tag{9-41}$$

Let us fix $1 \le j \le 3$. For a function $h = h_j(r)\vartheta_j(y/r)$ defined in B_{2R}, we let $H = L_0^{-1}[h] := H_j(r)\vartheta_j(y/r)$ be the solution of the equation

$$\Delta H + pU^{p-1}H + \tilde{h}_j\vartheta_j = 0 \quad \text{in } \mathbb{R}^n, \qquad H(y) \to 0 \quad \text{as } |y| \to \infty,$$

where $\tilde{h}_j$ designates the extension of h_j as zero outside B_{2R}, represented by the variation of parameters formula

$$H_j(r) = w_r(r)\int_r^{2R}\frac{1}{\rho^{n-1}w_r(\rho)^2}\int_\rho^\infty \tilde{h}_j(s)w_r(s)s^{n-1}\,ds.$$

If we consider a function $h^j = h_j(r,\tau)\vartheta_j$ defined in B_{2R} with $\|h^j\|_{2+a,\nu} < +\infty$ and $\int_{B_{2R}} h^j Z_j = 0$ for all τ, then $H_j = L_0^{-1}[h^j(\,\cdot\,,\tau)]$ satisfies the estimate $\|H_j\|_{\nu,a} \lesssim \|h_j\|_{\nu,2+a}$. Let us consider the boundary value problem in B_{3R}

$$\begin{aligned} \Phi_\tau &= \Delta\Phi + pU(y)^{p-1}\Phi + H_j(r)\vartheta_j(y) \quad \text{in } B_{3R}\times(\tau_0,\infty), \\ \Phi &= 0 \quad \text{on } \partial B_{3R}\times(\tau_0,\infty), \qquad \Phi(\,\cdot\,,\tau_0) = 0 \quad \text{in } B_{3R}. \end{aligned} \tag{9-42}$$

As consequence of Step 1, we find a solution $\Phi_j[h]$ to this problem, which defines a linear operator of h_j and satisfies the estimates

$$|\Phi_j(y,\tau)| \lesssim \frac{\tau^{-\nu} R^{3-a}}{1+|y|^2} R^1 \|h_j\|_{\nu,2+a}. \tag{9-43}$$

Arguing by scaling and parabolic estimates, we find as in the construction for mode 0,

$$|L[\Phi_j](\,\cdot\,,\tau)| \lesssim \tau^{-\nu} \|h\|_{\nu,2+a} \frac{R^{4-a}}{1+|y|^4} \quad \text{in } B_{2R}.$$

We define $\phi_j[h_j] := L[\Phi_j]\big|_{B_{2R}}$ and $\phi^1[h^1] := \sum_{j=1}^{3} \phi_j[h_j]\vartheta_j$. This function solves (9-7) for $h = h^1$ and satisfies

$$|\phi^1[h^1](y,\tau)| \lesssim \tau^{-\nu} \|h_j\|_{2+a,\nu} \frac{R^{4-a}}{1+|y|^4} \quad \text{in } B_{2R}. \tag{9-44}$$

Construction at higher modes. In order to deal with the higher modes, for $h = h^{\perp} = \sum_{j=4}^{\infty} h_j(r)\Theta_j$ we let $\phi^{\perp}[h^{\perp}]$ be just the unique solution of the problem

$$\begin{gathered} \phi_\tau = \Delta\phi + pU(y)^{p-1}\phi + h^{\perp} \quad \text{in } B_{2R} \times (\tau_0,\infty), \\ \phi = 0 \quad \text{on } \partial B_{2R} \times (\tau_0,\infty), \qquad \phi(\,\cdot\,,\tau_0) = 0 \quad \text{in } B_{2R}, \end{gathered} \tag{9-45}$$

which is estimated as

$$|\phi^{\perp}[h^{\perp}](y,\tau)| \lesssim \tau^{-\nu} \frac{\|h^{\perp}\|_{\nu,2+a}}{1+|y|^a} \quad \text{in } B_{2R}. \tag{9-46}$$

We just let

$$\phi[h] := \phi^0[h^0] + \phi^1[h^1] + \phi^{\perp}[h^{\perp}]$$

be the function constructed above. According to estimates (9-40) and (9-46) we find that this function solves problem (9-7) for $c(\tau)$ given by (9-17), with bounds (9-4), (9-5), (9-8) as required. ☐

10. Nonradially symmetric case

In this section, we discuss the existence of solutions for problem (2-1) when the initial condition is not radially symmetric, and we discuss the codimension-1 stability. Let $\bar{v}_0$ be a positive, uniformly bounded smooth function, not radially symmetric and define

$$v_0(x) = \frac{\bar{v}_0(x)}{|x|^\kappa}, \quad \text{with } \kappa > \max\left\{\frac{\gamma+3}{2}, \gamma\right\}. \tag{10-1}$$

We construct a solution to the initial value problem

$$\begin{cases} u_t = \Delta u + u^5 & \text{in } \mathbb{R}^3 \times (t_0,\infty), \\ u(x,t_0) = u_0(|x|) + v_0(x), \end{cases} \tag{10-2}$$

where u_0 is radial and satisfies the decay condition (2-2), while v_0 is a nonradial function of the form (10-1).

Since the strategy of the proof is similar to the one already performed in detail for $\bar{v}_0(x) \equiv 0$, we shall indicate the changes in the argument that are required when the initial condition is not radially symmetric.

We start with a slightly different first approximation. Let $p = p(t) : [t_0, \infty) \to \mathbb{R}^3$ be a smooth function so that

$$p(t_0) = \mathbf{0}, \quad p(t) = \int_{t_0}^{t} P(s)\, ds,$$

where P satisfies

$$\|P\|_{\diamond} := \sup_{t>t_0} \mu_0(t)^{-\frac{1}{2}} t^{\kappa-1} [\|P(s)\|_{\infty,[t,t+1]} + [P]_{0,\sigma,[t,t+1]}] \le \ell, \tag{10-3}$$

with σ the number fixed in (2-11), and ℓ a positive fixed number. Observe that, under these assumptions, and the bound on κ in (10-1), we have $|p(t)|/\mu_0(t) \to 0$ as $t \to \infty$. Define

$$U[\lambda, P](x,t) = \hat{U}_2(x,t) + U_3(x,t), \quad \hat{U}_2(x,t) := U_2(|x - p(t)|, t), \tag{10-4}$$

where U_2 is given by (2-57) and

$$U_3(x,t) = \left(1 - \eta\left(\frac{|x|}{t}\right)\right) v_0(x). \tag{10-5}$$

If we call $\mathcal{E}[\lambda, P](x,t) := \Delta U + U^5 - U_t$, we can write

$$\mathcal{E}[\lambda, P](x,t) = \mathcal{E}_2[\lambda](|x-p|,t) - \nabla U_2(|x-p|,t) \cdot \dot{p}(t) + \underbrace{\Delta U_3 - \frac{\partial U_3}{\partial t} + (\hat{U}_2 + U_3)^5 - (\hat{U}_2)^5}_{:=\mathcal{E}_3}.$$

Define

$$\bar{\mathcal{E}}(x,t) = \mathcal{E}_{21}(|x-p|,t) + \left(1 - \eta_R\left(\frac{|x|}{R\mu_0}\right)\right)[\mathcal{E}_{22}(|x-p|,t) - \nabla U_2(|x-p|,t) \cdot \dot{p}(t)], \tag{10-6}$$

where η_R is defined in (2-59). We have

$$|\bar{\mathcal{E}}(x,t)| \le C \mu_0^{\frac{1}{2}} t^{-\frac{3}{2}} h_0\left(\frac{|x|}{\sqrt{t}}\right), \quad |\mathcal{E}_3(x,t)| \le C t^{-\kappa+\frac{1}{2}} h_0\left(\frac{|x|}{\sqrt{t}}\right). \tag{10-7}$$

A solution to (10-2) does exist and has the form

$$u = U[\lambda, P](r,t) + \tilde{\phi}, \quad t > t_0, \tag{10-8}$$

where U is defined in (10-4), while $\tilde{\phi}(x,t)$ is given as in (3-2),

$$\tilde{\phi}(x,t) = \psi(x,t) + \phi^{\text{in}}(x,t), \quad \text{where } \phi^{\text{in}}(x,t) := \eta_R(x,t)\hat{\phi}(x,t),$$

and $\hat{\phi}(x,t) := \mu_0^{-\frac{1}{2}} \phi(x/\mu_0, t)$. For any $\psi_0 \in C^2(\mathbb{R}^3)$ so that

$$|y||\psi_0(y)| + |y||\nabla\psi_0(y)| \le t_0^{-a} e^{-b|y|} \tag{10-9}$$

for some positive constants a and b, the function ψ is the solution to

$$\begin{aligned} \partial_t \psi &= \Delta\psi + V\psi + [2\nabla\eta_R \nabla_x \hat{\phi} + \hat{\phi}(\Delta_x - \partial_t)\eta_R] + N[\lambda](\tilde{\phi}) + \bar{\mathcal{E}}_+ \mathcal{E}_3 \quad \text{in } \mathbb{R}^3 \times [t_0, \infty), \\ \psi(x, t_0) &= \psi_0, \end{aligned} \tag{10-10}$$

where V is defined as in (3-5) with U instead of U_2, and $N(\tilde{\phi}) = (U+\tilde{\phi})^5 - U^5 - 5U^4\tilde{\phi}$. This solution ψ can be described as

$$\psi(x,t) = \psi_r(x,t) + \psi_{nr}(x,t), \tag{10-11}$$

where ψ_r is a radial function in $|x-p(t)|$, for any t, and

$$|\psi_r(x,t)| \le C\mu_0^{\frac{1}{2}} t^{-\frac{1}{2}} \varphi_0\left(\frac{|x|}{\sqrt{t}}\right), \quad |\psi_{nr}(x,t)| \le C t^{-\eta+\frac{3}{2}} \varphi_0\left(\frac{|x|}{\sqrt{t}}\right). \tag{10-12}$$

We refer to (4-2) for the definition of φ_0.

On the other hand, the function $\hat{\phi}$ satisfies

$$\partial_t \hat{\phi} = \Delta\hat{\phi} + 5w_\mu^4 \hat{\phi} + 5w_\mu^4 \psi + \mathcal{E}_{22}(|x-p(t)|,t) - \nabla U_2(|x-p(t)|,t)\cdot \dot{p}(t) \quad \text{in } B_{2R\mu_0}(0)\times[t_0,\infty),$$

with $\hat{\phi}(x,t_0) = \mu_0^{-\frac{1}{2}}(t_0) e_0 Z(x/\mu_0(t_0))$. In terms of ϕ, this equation becomes

$$\begin{aligned} \mu_0^2 \partial_t \phi &= \Delta_y \phi + 5w^4\phi + f(y,t) \quad \text{in } B_{2R}(0)\times[t_0,\infty), \\ \phi(y,t_0) &= e_0 Z(y), \end{aligned} \tag{10-13}$$

where

$$\begin{aligned} f(y,t) = \mu_0^{\frac{5}{2}} \mathcal{E}_{22}(|\mu_0 y - p(t)|,t) - \nabla U_2(|\mu_0 y - p(t)|,t)\cdot \dot{p}(t) \\ + 5\frac{\mu_0^{\frac{1}{2}}}{(1+\Lambda)^4} w^4\left(\frac{y}{(1+\Lambda)^2}\right)\psi(\mu_0 y,t) + B[\phi] + B^0[\phi]. \end{aligned}$$

In the above expression, ψ is the solution to (10-10), while B and B^0 are defined respectively in (3-9) and (3-10). The solution ϕ exists in the class of functions with $\|\cdot\|_{\nu,a}$-norm bounded, see (4-1), as consequence of Proposition 9.1, and a contraction-type argument, provided the parameter functions λ and P can be chosen so that

$$\int_{B_R} f(y,t) Z_j(y)\, dy = 0 \quad \text{for all } t > t_0, \quad j = 0,1,\dots,n. \tag{10-14}$$

The system of $(n+1)$ nonlinear, nonlocal equations in λ and P is solvable for λ and P satisfying (2-11) and (10-3). Indeed, (10-14), for $j=0$, can be treated as we did for (5-1) in Sections 5, 6, 7. On the other hand, when $j = 1,\dots,n$, (10-14) are perturbations of

$$\dot{p}(t) = \mu_0^{\frac{1}{2}} t^{-\kappa+1} \bar{u}$$

for some fixed vector $\bar{u} \in \mathbb{R}^3$. Thus it can be solved for parameters $p(t) = \int_{t_0}^t P(s)\, ds$ satisfying (10-3). This concludes the proof of existence of a positive global solution to (10-2).

Next we discuss the codimension-1 stability. Let us observe that the construction of ϕ and e_0, the solution to (10-13), is possible for any initial condition ψ_0 to the outer problem (10-10). We have the validity of Lipschitz dependence of $\phi = \phi[\psi_0]$, and $e_0 = e_0[\psi_0]$ in the C^1-topology described in (10-9). As a consequence of the implicit function theorem, the maps $\phi[\psi_0]$ and $e_0[\psi_0]$ depend in the C^1-sense on ψ_0 in our C^1-topology (10-9), thanks to the corresponding dependence for ψ, λ and p.

Let us consider the following map defined in a small neighborhood of 0 in $X = C^1(\bar\Omega)$:

$$F(\psi_0) = \psi_0 - (e_0[\psi_0] - e_0[t_0])Z_0$$

so that $F[0] = 0$, F is differentiable, and

$$D_{\psi_0} F(0)[h] = h - \langle D_{\psi_0} e_0[0], h\rangle Z_0, \quad h \in X.$$

We have a solution which blows-up as $t \to +\infty$ provided that

$$u(\cdot, t_0) = u^*(\cdot, t_0) - e_0[0]Z_0 + g, \tag{10-15}$$

where u^* is the solution corresponding to $\psi_0 = 0$, and $g = F[\psi_0]$ for any small ψ_0.

The vector space of the functionals in X given by $D_{\psi_0} e_0[0]$ has dimension 1. We write $W := \mathrm{Ker}(D_{\psi_0} e_0[0])$ is a space with codimension 1. Indeed, we can find a nonzero function u such that

$$X = W \oplus \langle u\rangle.$$

We consider the operator in a neighborhood of 0 in X given by

$$G(w + \alpha u) = \alpha u + F(w), \quad \alpha_j \in \mathbb{R}, \quad w \in W.$$

Then G is of class C^1 near the origin, $G(0) = 0$ and $D_{\psi_0} G(0)[h] = h$. By the local inverse theorem, G defines a local C^1 diffeomorphism onto a neighborhood of the origin. For all small g we can find smooth functions $\alpha(g)$, $w(g)$ with

$$\alpha(g)u + F(w(g)) = g.$$

Thus the set $\mathcal{M}$ of functions $F[w]$, $w \in W$, can be described in a neighborhood of 0 exactly as those $g \in X$ such that

$$\alpha(g) = 0.$$

This says precisely that $\mathcal{M}$ is locally a codimension-1 C^1-manifold such that if g in (10-15) is selected there, then the desired phenomenon takes place. □

Appendix A

Proof of Lemma 2.2. We denote by $y_2(s)$ the solution to (2-17) with $\lim_{s\to\infty} s^{2\nu} y_2(s) = 1$, and by $y_1(s)$ another solution, linearly independent from y_2, defined explicitly by

$$y_1(s) = c y_2(s) \int_s^\infty \frac{e^{-\frac{z^2}{4}}}{y_2(z)^2 z^2}\, dz \tag{A-1}$$

for some positive constant c we fix later. The function $y_1(s)$ decays fast at infinity, since

$$y_1(s) = c_1 e^{-\frac{s^2}{4}} s^{4\nu-3}(1 + o(s^{-1})),$$

as $s \to \infty$, for some positive constant c_1, as a direct consequence from (A-1). The function $y_2(s)$ is definite for any $s \in (0, \infty)$, and it is positive. Indeed, we first observe that the operator L_ν satisfies the maximum principle. This is consequence of the fact that the positive function $g_0(s) = e^{-\frac{s^2}{4}}/s$, which solves $L_1(g_0) = 0$, satisfies $L_\nu(g_0) < 0$ in $(0,\infty)$. With this is mind, we define $\bar{g}_0(s) = \int_s^\infty e^{-\frac{z^2}{4}}/z^2\,dz$. This is a positive function which satisfies $L_\nu(\bar{g}_0) = \nu\bar{g}_0 > 0$ in $(0,\infty)$. Thus $\bar{g}_0$ is a subsolution. Moreover, it is easy to see that $\bar{g}_0(R) < y_2(R)$ for any R large enough. A standard application of the maximum principle thus gives that y_2 is positive in $(0,\infty)$.

We now claim that $\lim_{s\to 0^+} s y_1(s)$ exists and it is positive. Write $y_1(s) = \phi(s^2/4)$, $x = s^2/4$, from which we get

$$x\phi'' + \left(\tfrac{3}{2} + x\right)\phi' + \nu\phi = 0, \quad x \in (0,\infty).$$

Performing the further change of variables $\phi(x) = e^{-x}\varphi(x)$, we get that φ satisfies

$$x\varphi'' + \left(\tfrac{3}{2} - x\right)\varphi' - \left(\tfrac{3}{2} - \nu\right)\varphi = 0, \quad x \in (0,\infty). \tag{A-2}$$

In [Filippas et al. 2000, Appendix A], it is proven that (A-2) admits polynomial solutions if and only if $\frac{3}{2} - \nu = -k$, $k = 0, 1, 2, \dots$. Since $\frac{1}{2} < \nu < 1$, this never happens; thus φ cannot be bounded, as $x \to 0^+$. On the other hand, the behavior of the solutions to (A-2), as $x \to 0^+$, is determined by $x\varphi'' + \frac{3}{2}\varphi' = 0$, which implies that the solutions to (A-2) are bounded around $x = 0$, or they behave like $x^{-\frac{1}{2}}$ as $x \to 0^+$. Combining all the above information, for a proper choice of the constant c in (A-1), we get

$$y_1(s) = \frac{1}{s}(1 + o(1)) \quad \text{as } s \to 0.$$

To understand further the behavior of y_1 around $s = 0$, we write $s y_1(s) = f(s)$, so that

$$f'' + \tfrac{1}{2}sf' + \left(\nu - \tfrac{1}{2}\right)f = 0, \quad s \in (0,\infty). \tag{A-3}$$

Integrating (A-3) between 0 and ∞, and using the fast decay of y_1 to 0 as $s \to \infty$, we compute

$$f'(0) = (\nu - 1)\int_0^\infty f(s)\,ds < 0, \quad f''(0) = \tfrac{1}{2} - \nu. \tag{A-4}$$

With this information, we get the estimates (2-18) and (2-20) for $y_1(s)$.

Since the Wronskian associated to problem (2-17) is given by a multiple of $e^{-\frac{s^2}{4}}/s^2$, we conclude that, since y_1 is unbounded as $s \to 0^+$, we have $y_2(s)$ is bounded as $s \to 0^+$. □

Lemma A.1. *Let $h = h(s)$ be a smooth function defined for $s \geq 0$ so that*

$$h(s) = \begin{cases} 1/s & \text{for } s \to 0, \\ 1/s^3 & \text{for } s \to \infty. \end{cases}$$

Then there exists a solution to

$$\partial_t \psi = \Delta\psi + t^{-\beta} h\left(\frac{r}{\sqrt{t}}\right), \tag{A-5}$$

of the form

$$\psi(r,t) = t^{-\beta+1}\varphi\left(\frac{r}{\sqrt{t}}\right), \quad \text{with } \varphi(s) = \begin{cases} s & \text{for } s \to 0, \\ 1/s^3 & \text{for } s \to \infty. \end{cases} \tag{A-6}$$

Proof. We look for a solution to (A-5) of the form $\psi(r,t) = t^{-(\beta-1)}\varphi(r/\sqrt{t})$. Thus φ satisfies

$$\varphi'' + \left(\frac{2}{s} + \frac{s}{2}\right)\varphi' + (\beta - 1)\varphi + h(s) = 0.$$

We look for a solution of the above equation of the form

$$\varphi(s) = z(s)y_1(s),$$

where y_1 solves $y_1'' + (2/s + s/2)y_1' + (\beta - 1)y_1 = 0$, and

$$y_1(s) \sim \begin{cases} 1/s & \text{as } s \to 0, \\ e^{-\frac{s^2}{4}} s^{4(\beta-1)-3} & \text{as } s \to \infty. \end{cases}$$

The existence of y_1 is consequence of Lemma 2.2. A direct computation gives

$$z(s) = -\int_0^s \frac{e^{-\frac{\eta^2}{4}}}{y_1(\eta)^2\eta^2}\left(\int_0^\eta h(x)y_1(x)x^2 e^{\frac{x^2}{4}}\,dx\right)d\eta.$$

One can easily see that

$$z(s) \sim \begin{cases} s^2 & \text{as } s \to 0, \\ e^{\frac{s^2}{4}} s^{-4(\beta-1)} & \text{as } s \to \infty. \end{cases}$$

This fact gives (A-6), and concludes the proof of the lemma. □

Proof of (5-10). For $x \in B_{2R}$, we shall prove

$$\phi_0(\mu_0 x, t) - \phi_0(0,t) = \alpha(t)|\mu_0 x|^\sigma \Pi(t)\Theta(|x|) \tag{A-7}$$

for some $\sigma \in (0,1)$. Here $\Pi = \Pi(t)$ denotes a smooth and bounded function of t, and Θ a smooth and bounded function of x.

We have

$$\begin{aligned}\phi_0(\mu_0 x, t) - \phi_0(0,t) &= \int_{t_0}^t \frac{1}{(4\pi(t-s))^{\frac{3}{2}}} \int_{\mathbb{R}^3} [e^{-\frac{|x-y|^2}{4(t-s)}} - e^{\frac{-|y|^2}{4(t-s)}}]\frac{\alpha(s)}{|y|}\mathbf{1}_{\{r<M\}}\,dy\,ds \\ &= \frac{1}{2}\int_{t_0}^t\int \frac{\beta'(s)}{(t-s)^{\frac{1}{2}}}[e^{-|z-\frac{\mu_0 x}{2\sqrt{t-s}}|^2} - e^{-|z|^2}]\frac{1}{|z|}\mathbf{1}_{\{|z|<\frac{M}{2\sqrt{t-s}}\}}\,dy\,ds \\ &= I + II,\end{aligned}$$

where

$$I = \int_{t_0}^{t-(\frac{\mu_0 x}{2m})}\int \frac{\beta'(s)}{(t-s)^{\frac{1}{2}}}[e^{-|z-\frac{\mu_0 x}{2\sqrt{t-s}}|^2} - e^{-|z|^2}]\frac{1}{|z|}\mathbf{1}_{\{|z|<\frac{M}{2\sqrt{t-s}}\}}\,dy\,ds.$$

We start estimating II. We observe that, if $t - (\mu_0 x/(2m)) < s < t$, then $\mu_0|x|/(2\sqrt{t-s}) > m$. We write

$$II = II_1 + II_2 + II_3,$$

where

$$II_j = \int_{t-(\frac{\mu_0 x}{2m})}^{t} \int_{D_j} \frac{\beta'(s)}{(t-s)^{\frac{1}{2}}} [e^{-|z-\frac{\mu_0 x}{2\sqrt{t-s}}|^2} - e^{-|z|^2}] \frac{1}{|z|} \mathbf{1}_{\{|z|<\frac{M}{2\sqrt{t-s}}\}} \, dy \, ds,$$

with

$$D_1 = \left\{ z : \left| z - \frac{\mu_0 x}{2\sqrt{t-s}} \right| < \frac{1}{4} \frac{\mu_0 |y|}{2\sqrt{t-s}} \right\}, \quad D_2 = \left\{ z : |z| < \frac{1}{4} \frac{\mu_0 |y|}{2\sqrt{t-s}} \right\},$$

and D_3 is the complement of the two above regions.

We start estimating II_1. We see that

$$\int_{D_1} e^{-|z-\frac{\mu_0 x}{2\sqrt{t-s}}|^2} \frac{1}{|z|} \mathbf{1}_{\{|z|<\frac{M}{2\sqrt{t-s}}\}} \, dy = \int e^{-|\bar{z}|} \frac{1}{\left| \bar{z} + \frac{\mu_0 x}{2\sqrt{t-s}} \right|} \, d\bar{z} = c \frac{2\sqrt{t-s}}{\mu_0 |x|}$$

for some constant c, as a direct application of dominated convergence theorem. Thus

$$\int_{t-(\frac{\mu_0 x}{2m})}^{t} \int_{D_1} e^{-|z-\frac{\mu_0 x}{2\sqrt{t-s}}|^2} \frac{1}{|z|} \mathbf{1}_{\{|z|<\frac{M}{2\sqrt{t-s}}\}} \, dy \, ds = \frac{2c}{\mu_0 |x|} \int_{t-(\frac{\mu_0 x}{2m})}^{t} \sqrt{t-s} \, ds = c' (\mu_0 |x|)^{\frac{1}{2}}.$$

On the other hand, for any z in D_1, one has

$$|z| > \frac{1}{4} \frac{\mu_0 |x|}{2\sqrt{t-s}},$$

and hence we can bound

$$\left| \int_{D_1} e^{-|z|^2} \frac{1}{|z|} \, dz \right| \le c \left[\frac{\sqrt{t-s}}{\mu_0 |x|} \right]^{\sigma}$$

for any $\sigma > 0$. We take $\sigma > 1$, so that

$$\left| \int_{t_0}^{t-(\frac{\mu_0 x}{2m})} \int_{D_1} e^{-|z|^2} \frac{1}{|z|} \mathbf{1}_{\{|z|<\frac{M}{2\sqrt{t-s}}\}} \, dy \, ds \right| \le \frac{1}{(\mu_0 |x|)^{\sigma}} \left| \int_{t_0}^{t-(\frac{\mu_0 x}{2m})} (t-s)^{\frac{\sigma}{2}-\frac{1}{2}} \, ds \right| \le c' \mu_0 |x|.$$

Thus we conclude that

$$|II_1| \lesssim \beta'(t) (\mu_0 |x|)^{\frac{1}{2}}.$$

Arguing in a similar way, one finds the same type of estimate for II_2. In the third region D_3, we have

$$|z| > \frac{1}{4} \frac{\mu_0 |x|}{2\sqrt{t-2}}, \quad |z - \frac{\mu_0 x}{2\sqrt{t-s}}| > \frac{1}{4} \frac{\mu_0 |x|}{2\sqrt{t-s}},$$

so that again one gets the estimate

$$|II_3| \lesssim \beta'(t) \mu_0 |x|.$$

Let us now consider the interval of time $t_0 < s < t - (\mu_0 |x| / (2m\sqrt{t-s}))^2$, that is, region where one has $\mu_0 |x| / (2\sqrt{t-s}) < m$. We have the decomposition

$$I = III + IV,$$

where

$$III = \int_{t_0}^{t-1} \int \frac{\beta'(s)}{(t-s)^{\frac{1}{2}}} [e^{-|z-\frac{\mu_0 x}{2\sqrt{t-s}}|^2} - e^{-|z|^2}] \frac{1}{|z|} \mathbf{1}_{\{|z|<\frac{M}{2\sqrt{t-s}}\}} \, dy \, ds.$$

We start with IV, where we expand in Taylor:

$$IV = \int_{t-1}^{t-(\frac{\mu_0|x|}{2m\sqrt{t-s}})^2} \int \frac{\beta'(s)}{(t-s)^{\frac{1}{2}}} [e^{-|z-\frac{\mu_0 x}{2\sqrt{t-s}}|^2} - e^{-|z|^2}] \frac{1}{|z|} \mathbf{1}_{\{|z|<\frac{M}{2\sqrt{t-s}}\}} \, dy \, ds$$
$$= \beta'(t) \int_{t-1}^{t-(\frac{\mu_0|x|}{2m\sqrt{t-s}})^2} \frac{\mu_0|x|}{t-s} \left(\int \frac{e^{-|z|^2}}{|z|} \, dz \right) ds = \beta'(t) \log\left(\frac{1}{t} \left(t - \left(\frac{\mu_0|x|}{2m\sqrt{t-s}} \right)^2 \right) \right) \mu_0|x|$$
$$= \beta'(t)\mu_0|x|[\log(\mu_0|x|)] = \beta'(t)(\mu_0|x|)^\sigma$$

for some positive $\sigma < 1$. Finally, we consider III. Again, after a Taylor expansion, we have

$$III = \mu_0|x| \int_{t_0}^{t-1} \frac{\beta'(s)}{(t-s)} \, ds = \mu_0|x| \int_{t_0}^{t-1} \frac{\beta'(s)}{t-s} \, ds.$$

Collecting the previous estimates, we conclude with the validity of (A-7). □

Appendix B

Proof of Lemma 2.3. Throughout the proof of the lemma, we denote by $q_i = q_i(s)$, for any integer i, a smooth real function with the property that

$$\frac{d}{(ds)^j} q_i(0) = 0 \quad \text{for } j < i \qquad \text{and} \qquad \frac{d}{(ds)^i} q_i(0) \neq 0.$$

With $\Theta = \Theta(r)$ we intend a smooth function of the space variable, which is uniformly bounded. Also, $\Pi = \Pi(t)$ stands for a smooth function of the time variable, which is uniformly bounded in $t \in (0, \infty)$. The explicit expressions of these functions change from line to line, and also within the same line.

Let $R_0 = r_0\sqrt{t}$. A simple computation gives the explicit expression of the error $\mathcal{E}_1$ in (2-40):

$$\mathcal{E}_1(r,t) = \mathcal{E}^1_{\text{in}} \eta\left(\frac{r}{R_0}\right) + \mathcal{E}^1_{\text{out}} \left(1 - \eta\left(\frac{r}{R_0}\right)\right)$$
$$+ \underbrace{R_0^{-2}(u_{\text{in}} - u_{\text{out}}) \Delta\eta\left(\frac{r}{R_0}\right) + 2R_0^{-1} \nabla(u_{\text{in}} - u_{\text{out}}) \cdot \nabla\eta\left(\frac{r}{R_0}\right)}_{:=\bar{\mathcal{E}}_1} + \underbrace{(u_{\text{in}} - u_{\text{out}}) \frac{R_0'}{R_0^2} \eta'\left(\frac{r}{R_0}\right)}_{:=\hat{\mathcal{E}}_1}, \tag{B-1}$$

where

$$\mathcal{E}^1_{\text{in}} = \Delta u_{\text{in}} + u^5_{\text{in}} - \partial_t u_{\text{in}} \quad \text{and} \quad \mathcal{E}^1_{\text{out}} = \Delta u_{\text{out}} + u^5_{\text{out}} - \partial_t u_{\text{out}}. \tag{B-2}$$

We start analyzing $\mathcal{E}^1_{\text{in}}$, getting

$$\mathcal{E}^1_{\text{in}}(r,t) = \mu_0'[\Delta\psi_1 + 5w^4_\mu \psi_1] - \mu' \frac{\partial w_\mu}{\partial \mu} + (w_\mu + \mu_0'\psi_1)^5 - w^4_\mu - 5w^4_\mu \mu_0' \psi_1 - \mu_0'' \psi_1 - \mu_0' \mu' \frac{\partial \psi_1}{\partial \mu}$$
$$= (\mu' - \mu_0')\mu^{-\frac{3}{2}} Z_0\left(\frac{r}{\mu}\right) + [(w_\mu + \mu_0'\psi_1)^5 - w^5_\mu - 5w^4_\mu \mu_0' \psi_1] - \mu_0'' \psi_1 - \mu_0' \mu' \frac{\partial \psi_1}{\partial \mu}. \tag{B-3}$$

Now we write

$$(\mu'-\mu_0')\mu^{-\frac{3}{2}}Z_0\left(\frac{r}{\mu}\right)=[2(\mu^{\frac{1}{2}}-\mu_0^{\frac{1}{2}})'+(\mu^{\frac{1}{2}}-\mu_0^{\frac{1}{2}})\mu_0^{-1}\mu_0']\mu^{-1}Z_0\left(\frac{r}{\mu}\right)-\frac{(\mu^{\frac{1}{2}}-\mu_0^{\frac{1}{2}})^2}{\mu^{\frac{1}{2}}}\mu_0^{-1}\mu_0'\mu^{-1}Z_0\left(\frac{r}{\mu}\right).$$

Taking into account that

$$Z_0(s)=\frac{3^{\frac{1}{4}}}{2}\frac{1}{s}+O\left(\frac{1}{s^3}\right)$$

as $s\to\infty$, it is convenient to write

$$\begin{aligned}&[2(\mu^{\frac{1}{2}}-\mu_0^{\frac{1}{2}})'+(\mu^{\frac{1}{2}}-\mu_0^{\frac{1}{2}})\mu_0^{-1}\mu_0']\mu^{-1}Z_0\left(\frac{r}{\mu}\right)\\ &\qquad=\frac{\alpha(t)}{\mu+r}+[2(\mu^{\frac{1}{2}}-\mu_0^{\frac{1}{2}})'+(\mu^{\frac{1}{2}}-\mu_0^{\frac{1}{2}})\mu_0^{-1}\mu_0']\mu^{-1}\left[Z_0\left(\frac{r}{\mu}\right)-\frac{3^{\frac{1}{4}}}{2}\frac{\mu}{\mu+r}\right],\end{aligned}$$

where α is defined in (2-41). We decompose (B-3) as

$$\mathcal{E}_{\text{in}}^1(r,t)=\frac{\alpha(t)}{\mu+r}+\bar{\mathcal{E}}_{\text{in}}^1(r,t),\tag{B-4}$$

where $\bar{\mathcal{E}}_{\text{in}}^1$ is explicitly given by

$$\begin{aligned}\bar{\mathcal{E}}_{\text{in}}^1(r,t)=&-\frac{(\mu^{\frac{1}{2}}-\mu_0^{\frac{1}{2}})^2}{\mu^{\frac{1}{2}}}\mu_0^{-1}\mu_0'\mu^{-1}Z_0\left(\frac{r}{\mu}\right)-\mu_0''\psi_1+[(w_\mu+\mu_0'\psi_1)^5-w_\mu^5-5w_\mu^4\mu_0'\psi_1]\\ &+[2(\mu^{\frac{1}{2}}-\mu_0^{\frac{1}{2}})'+(\mu^{\frac{1}{2}}-\mu_0^{\frac{1}{2}})\mu_0^{-1}\mu_0']\mu^{-1}\left[Z_0\left(\frac{r}{\mu}\right)-\frac{3^{\frac{1}{4}}}{2}\frac{\mu}{\mu+r}\right]-\mu_0'\mu'\frac{\partial\psi_1}{\partial\mu}=\sum_{j=1}^{5}e_j.\end{aligned}\tag{B-5}$$

We observe now that $(e_1+e_2+e_3)\eta(r/R_0)$ can be described as sum of functions of the form

$$\frac{\mu_0^{\frac{1}{2}}t^{-2}R_0^2}{\mu_0+r}q_0(\Lambda)\Pi(t)\Theta(r),\quad \frac{\mu_0^{-\frac{1}{2}}t^{-1}}{\mu_0+r}q_2(\Lambda)\Pi(t)\Theta(r),\tag{B-6}$$

where q_0 is a smooth function with $q(0)\neq 0$, while q_2 is a smooth function with $q_2(0)=q_2'(0)=0$, and $q_2''(0)\neq 0$. On the other hand, we see that

$$e_4=\frac{\alpha(t)\mu_0^2}{\mu_0^3+r^3}\Pi(t)\Theta(r),\quad e_5=\frac{\mu_0^{\frac{1}{2}}t^{-1}}{\mu_0+r}[R_0^2\Lambda'+R_0^2t^{-1}q_1(\Lambda)]\Pi(t)\Theta(r),\tag{B-7}$$

where q_1 is a smooth function with $q_1(0)=0,\ q_1'(0)\neq 0$. Under assumption (2-11) and combining (B-4), (B-6) and (B-7) we find that

$$|\bar{\mathcal{E}}_{\text{in}}^1\eta|_{\infty,B(x,1)\times[t,t+1]}\lesssim\mu_0^{\frac{1}{2}}t^{-\frac{3}{2}}h_0\left(\frac{r}{\sqrt{t}}\right),\quad r=|x|.\tag{B-8}$$

Since (2-41), we observe that

$$\left|\frac{\alpha(t)}{\mu+r}\left(1-\eta\left(\frac{r}{R_0}\right)\right)\right|\lesssim\mu_0^{\frac{3}{2}}t^{-\frac{3}{2}}h_0\left(\frac{r}{\sqrt{t}}\right),\quad r=|x|.$$

Let us fix λ_1 and λ_2 satisfying (2-11). We write, for some $\bar\lambda = s\lambda_1 + (1-s)\lambda_2,\ s \in (0,1)$,

$$(\bar{\mathcal{E}}^1_{\text{in}}[\lambda_1] - \bar{\mathcal{E}}^1_{\text{in}}[\lambda_2])\eta\left(\frac{r}{R_0}\right) = (D_\lambda \bar{\mathcal{E}}^1_{\text{in}}[\bar\lambda][\lambda_1-\lambda_2])\eta\left(\frac{r}{R_0}\right), \quad \text{with } D_\lambda \bar{\mathcal{E}}^1_{\text{in}}[\bar\lambda] = \sum_{j=1}^5 (D_\lambda e_j)[\bar\lambda],$$

where the e_j are defined in (B-5). Let us consider e_1. We have

$$(D_\lambda e_1)[\bar\lambda] = 2\mu_0(1+\Lambda)D_\mu(e_1)[\bar\lambda].$$

Direct computation gives

$$|D_\mu(e_1)[\bar\lambda](r,t)| \lesssim \frac{\mu_0^{-\frac{1}{2}} t^{-1}}{\mu_0 + r} q_0(\bar\lambda)\Pi(t)\Theta(r).$$

We combine the above estimates to get

$$\begin{aligned}
|e_1[\lambda_1] - e_1[\lambda_2]|\eta\left(\frac{r}{R_0}\right) &\le \mu_0 \frac{\mu_0^{-\frac{1}{2}} t^{-1}}{\mu_0 + r} |\lambda_1 - \lambda_2| \eta\left(\frac{r}{R_0}\right) \\
&\le C(\mu_0(t)t^{-1})\mu_0^{\frac{3}{2}}(t)t^{-\frac{3}{2}} h_0\left(\frac{r}{\sqrt{t}}\right)\|\lambda_1-\lambda_2\|_\sharp \\
&\le C(\mu_0(t_0)t_0^{-1})\mu_0^{\frac{3}{2}}(t)t^{-\frac{3}{2}} h_0\left(\frac{r}{\sqrt{t}}\right)\|\lambda_1-\lambda_2\|_\sharp .
\end{aligned}$$

Choosing t_0 large if necessary, we get $C(\mu_0(t_0)t_0^{-1}) < 1$. Similar estimates can be obtained for the other terms $e_2, \dots, e_5$. Thus we get

$$|(\bar{\mathcal{E}}^1_{\text{in}}[\lambda_1] - \bar{\mathcal{E}}^1_{\text{in}}[\lambda_2])\chi|_{\infty, B(x,1)\times[t,t+1]} \le c_1^o \mu_0^{\frac{1}{2}} t^{-\frac{3}{2}} h_0\left(\frac{r}{\sqrt{t}}\right)\|\lambda_1-\lambda_2\|_\sharp$$

for some constant c_1^o which can be made arbitrarily small, if t_0 is chosen large. Also, we have

$$[\bar{\mathcal{E}}^1_{\text{in}}[\lambda_1] - \bar{\mathcal{E}}^1_{\text{in}}[\lambda_2]]_{0,\sigma,[t,t+1]} \le c_1^o \mu_0^{\frac{1}{2}} t^{-\frac{3}{2}} h_0\left(\frac{r}{\sqrt{t}}\right)[\lambda_1-\lambda_2]_{0,\sigma,[t,t+1]}.$$

Let us now describe $\mathcal{E}^1_{\text{out}}$. A first observation is that, for any value of γ, we immediately see that $\mathcal{E}^1_{\text{out}}$ does not depend on λ. On the other hand, if $1 < \gamma \le 2$ the expression for $\mathcal{E}^1_{\text{out}}$ becomes

$$\mathcal{E}^1_{\text{out}}(r,t) = u^5_{\text{out}},$$

so that we directly get

$$\left|\mathcal{E}^1_{\text{out}}\left(1-\chi\left(\frac{r}{R_0}\right)\right)\right| \le C\frac{\mu_0^{\frac{5}{2}}}{r^5}\mathbf{1}_{\{r>R_0^{-1}\}}. \tag{B-9}$$

Let us consider now $\gamma > 2$. In this case, the expression of $\mathcal{E}^1_{\text{out}}$ is a bit more involved

$$\begin{aligned}
\mathcal{E}^1_{\text{out}}(r,t) = \eta\left(\frac{r}{t}\right)(u^1_{\text{out}})^5 + \left(1-\eta\left(\frac{r}{t}\right)\right)A\left[\frac{\gamma(\gamma-1)}{r^{\gamma+2}} + \frac{A^4}{r^{5\gamma}}\right] \\
+ \underbrace{t^{-2}(u^1_{\text{out}}-u^2_{\text{out}})\Delta\eta\left(\frac{r}{t}\right) + 2t^{-}\nabla(u^1_{\text{out}}-u^2_{\text{out}})\cdot\nabla\eta\left(\frac{r}{t}\right)}_{:=\bar{\mathcal{E}}_1^{\text{out}}} + \underbrace{(u^1_{\text{out}}-u^2_{\text{out}})t^{-2}\eta'\left(\frac{r}{t}\right)}_{:=\hat{\mathcal{E}}_1^{\text{out}}}.
\end{aligned} \tag{B-10}$$

A close analysis of each one of the terms appearing in (B-10) gives

$$\left|\mathcal{E}^1_{\text{out}}\left(1-\eta\left(\frac{r}{R_0}\right)\right)\right| \le C\left\{\frac{t^{-(\gamma-1)}}{r^3}\mathbf{1}_{\{r>t\}} + \frac{t^{-2}\mu_0^{\frac12}}{r}\mathbf{1}_{\{t<r<2t\}} + \frac{t^{-\frac52}}{r^5}\mathbf{1}_{\{r_0\sqrt{t}<r<t\}}\right\}. \tag{B-11}$$

From (B-9)–(B-10) and (B-11), we obtain

$$\left|\mathcal{E}^1_{\text{out}}\left(1-\chi\left(\frac{r}{R_0}\right)\right)\right| \lesssim \begin{cases}\mu_0^{\frac12}t^{-\frac32}h_0(r/\sqrt{t}) & \text{if } 1<\gamma\le 2,\\ t^{-2}h_0(r/\sqrt{t}) & \text{if } \gamma>2.\end{cases}$$

Going back to (B-1), we are left with the description of $\bar{\mathcal{E}}_1 = \bar{\mathcal{E}}_1[\lambda]$ and $\hat{\mathcal{E}}_1[\lambda]$. Directly we check

$$|\bar{\mathcal{E}}_1(r,t)|, |\hat{\mathcal{E}}_1(r,t)| \le CR_0^{-2}\frac{\mu_0^{\frac12}}{r}\mathbf{1}_{\{R_0<r<2R_0\}} \tag{B-12}$$

for some positive constant C. This gives right away

$$|\bar{\mathcal{E}}_1 + \hat{\mathcal{E}}_1| \lesssim \mu_0^{\frac12}t^{-\frac32}h_0\left(\frac{r}{\sqrt{t}}\right).$$

Let us fix λ_1 and λ_2 satisfying (2-11). We write, for some $\bar{\lambda} = s\lambda_1 + (1-s)\lambda_2,\ s\in(0,1)$,

$$\bar{\mathcal{E}}_1[\lambda_1](r,t) - \bar{\mathcal{E}}_1[\lambda_2](r,t) = D_\lambda\bar{\mathcal{E}}_1[\bar{\lambda}][\lambda_1-\lambda_2](r,t),$$

where

$$D_\lambda\bar{\mathcal{E}}_1[\bar{\lambda}] = R_0^{-2}(\partial_\lambda u_{\text{in}}[\bar{\lambda}])\Delta\eta\left(\frac{r}{R_0}\right) + 2R_0^{-1}\nabla((\partial_\lambda u_{\text{in}})[\bar{\lambda}])\cdot\nabla\eta\left(\frac{r}{R_0}\right).$$

Since in the region we are considering

$$\partial_\lambda u_{\text{in}}[\bar{\lambda}] = 2\mu_0(1+\Lambda)(\partial_\mu u_{\text{in}})[\bar{\lambda}], \quad |(\partial_\mu u_{\text{in}})| \le c\frac{\mu_0^{-\frac12}}{r},$$

we have

$$\begin{aligned}|\bar{\mathcal{E}}_1[\lambda_1](r,t) - \bar{\mathcal{E}}_1[\lambda_2]|_{\infty,B(x,1)\times[t,t+1]} &\le (\mu_0(t_0)t_0^{-1})\mu_0^{\frac12}t^{-\frac32}h_0\left(\frac{r}{\sqrt{t}}\right)\|\lambda_1-\lambda_2\|_\sharp\\ &\le c_1\mu_0^{\frac12}t^{-\frac32}h_0\left(\frac{r}{\sqrt{t}}\right)\|\lambda_1-\lambda_2\|_\sharp\end{aligned}$$

for some constant $c_1\in(0,1)$, provided t_0 is large enough. Furthermore, we also have, for any $t>t_0$,

$$[\bar{\mathcal{E}}_1[\lambda_1] - \bar{\mathcal{E}}_1[\lambda_2]]_{0,\sigma,[t,t+1]} \le c_1\mu_0^{\frac12}t^{-\frac32}h_0\left(\frac{r}{\sqrt{t}}\right)([\lambda_1-\lambda_2]_{0,\sigma,[t,t+1]}),$$

with again $c_1\in(0,1)$. Collecting all the previous estimates, we get the proof of the lemma. □

Remark B.1. From the proof of the result, we also get that the constants $\boldsymbol{c}$ in (2-50) and (2-51) can be made as small as one needs, provided that the initial time t_0 is chosen large enough.

Appendix C

Proof of Lemma 2.4. Under the assumptions (2-11) on λ, we get that, for any $r>0$ and $t>t_0$,

$$|\mathcal{E}_{2,1}(r,t)|+[\mathcal{E}_{2,1}(r,t)]_{0,\sigma,[t,t+1]}\lesssim \mu_0^{\frac12}t^{-\frac32}h_0\Big(\frac{r}{\sqrt t}\Big), \tag{C-1}$$

where h_0 is given by (2-44), and also estimates similar to (2-50) and (2-51) for $\partial_\lambda\mathcal{E}_{2,1}$. These estimates follow from (2-49)–(2-50), (2-41) and from

$$\left|\frac{\alpha(t)}{\mu+r}\Big(\eta\Big(\frac{r}{R_0}\Big)-\mathbf{1}_{\{r<2M\}}\Big)\right|\le|\alpha(t)|t^{-\frac12}h_0\Big(\frac{r}{\sqrt t}\Big).$$

Here we use again $R_0=r_0\sqrt t$. Furthermore, in the region where $\eta(r/R_0)-\mathbf{1}_{\{r<2M\}}\neq0$, the above function is regular enough to have

$$\left[\frac{\alpha(t)}{\mu+r}\Big(\eta\Big(\frac{r}{R_0}\Big)-\mathbf{1}_{\{r<2M\}}\Big)\right]_{0,\sigma,B(x,1)\times[t,t+1]}\le|\alpha(t)|t^{-\frac12}h_0\Big(\frac{r}{\sqrt t}\Big),\qquad r=|x|.$$

Using (2-43), we get (C-1). Let us consider now $\mathcal{E}_{22}(1-\eta_R)(r,t)$. We claim that

$$\|\mathcal{E}_{22}(1-\eta_R)(r,t)\|_*\le c_2. \tag{C-2}$$

Given $d>1$, define

$$h_*(s)=\begin{cases}1/s & \text{for } s\to0,\\ 1/s^d & \text{for } s\to\infty.\end{cases}$$

Arguing as in the proof of Lemma A.1, we get the existence of ψ_* so that

$$\partial_t\psi_*=\Delta\psi_*+\mu_0^{\frac12}t^{-\frac32}h_*\Big(\frac{r}{\sqrt t}\Big),\quad \text{with } \psi_*(r,t)=\mu_0^{\frac12}t^{-\frac12}\varphi_*\Big(\frac{r}{\sqrt t}\Big),\quad \varphi(s)=\begin{cases}s & \text{for } s\to0,\\ 1/s^d & \text{for } s\to\infty.\end{cases}$$

Comparing the above equation and the equation satisfied by ϕ_0, and using the maximum principle, we obtain that, in the region where $(1-\eta_R)\neq0$,

$$|\phi_0(x,t)|\le\|\lambda\|_\sharp\mu_0^{\frac12}t^{-\frac12}\varphi_*\Big(\frac{r}{\sqrt t}\Big). \tag{C-3}$$

We proceed now with the estimate of $(1-\eta_R)\mathcal{E}_{22}$. A Taylor expansion gives the existence of $s^*\in(0,1)$, so that

$$\mathcal{E}_{22}(r,t)=5(U_1+s^*\phi_0)^4\phi_0.$$

Let $\overline{M}$ be a large fixed number. From (2-38) and (2-13), we see that, if $r<\overline{M}\sqrt t$,

$$|(1-\eta_R)\mathcal{E}_{22}|\lesssim w_\mu^4\phi_0\lesssim R^{-2}\mu_0^{\frac12}t^{-\frac32}h_0\Big(\frac{r}{\sqrt t}\Big).$$

On the other hand, thanks to (C-3) we see that, for $r>\overline{M}\sqrt t$, we get

$$|(1-\eta_R)\mathcal{E}_{22}|\lesssim(\phi_0)^5\lesssim\mu_0^{\frac52}t^{-\frac52}h_0\Big(\frac{r}{\sqrt t}\Big).$$

Thus we get the L^∞ bound in estimate (C-2). The control on the Hölder norm contained in (2-61) and (2-62) follows arguing as in the proof of (2-50)–(2-51) in the proof of Lemma 2.3, and from the assumption on λ in (2-11). We leave the details to the reader. □

Acknowledgements

We are indebted to Marek Fila for introducing this problem to us and for many useful discussions. Del Pino was supported by a UK Royal Society Research Professorship and Grant PAI AFB-170001, Chile. Musso was partly supported by FONDECYT grant 1160135, Chile. The research of Wei is partially supported by the NSERC of Canada.

References

[Aubin 1976] T. Aubin, "Problèmes isopérimétriques et espaces de Sobolev", *J. Differential Geom.* **11**:4 (1976), 573–598. MR Zbl

[Caffarelli et al. 1989] L. A. Caffarelli, B. Gidas, and J. Spruck, "Asymptotic symmetry and local behavior of semilinear elliptic equations with critical Sobolev growth", *Comm. Pure Appl. Math.* **42**:3 (1989), 271–297. MR Zbl

[Caputo 1967] M. Caputo, "Linear models of dissipation whose Q is almost frequency independent, II", *Geophys. J. R. Astr. Soc.* **13**:5 (1967), 529–539.

[Ciraolo et al. 2018] G. Ciraolo, A. Figalli, and F. Maggi, "A quantitative analysis of metrics on $\mathbb{R}^n$ with almost constant positive scalar curvature, with applications to fast diffusion flows", *Int. Math. Res. Not.* **2018**:21 (2018), 6780–6797. MR Zbl

[Collot et al. 2017] C. Collot, F. Merle, and P. Raphaël, "Dynamics near the ground state for the energy critical nonlinear heat equation in large dimensions", *Comm. Math. Phys.* **352**:1 (2017), 215–285. MR Zbl

[Cortázar et al. 2016] C. Cortázar, M. del Pino, and M. Musso, "Green's function and infinite-time bubbling in the critical nonlinear heat equation", preprint, 2016. To appear in *J. Eur. Math. Soc.* arXiv

[Daskalopoulos et al. 2018] P. Daskalopoulos, M. del Pino, and N. Sesum, "Type II ancient compact solutions to the Yamabe flow", *J. Reine Angew. Math.* **738** (2018), 1–71. MR Zbl

[Davila et al. 2017] J. Davila, M. del Pino, and J. Wei, "Singularity formation for the two-dimensional harmonic map flow into S^2", preprint, 2017. arXiv

[Donninger and Krieger 2013] R. Donninger and J. Krieger, "Nonscattering solutions and blowup at infinity for the critical wave equation", *Math. Ann.* **357**:1 (2013), 89–163. MR Zbl

[Duyckaerts et al. 2012] T. Duyckaerts, C. Kenig, and F. Merle, "Profiles of bounded radial solutions of the focusing, energy-critical wave equation", *Geom. Funct. Anal.* **22**:3 (2012), 639–698. MR Zbl

[Duyckaerts et al. 2013] T. Duyckaerts, C. Kenig, and F. Merle, "Classification of radial solutions of the focusing, energy-critical wave equation", *Cambridge J. Math.* **1**:1 (2013), 75–144. MR Zbl

[Duyckaerts et al. 2016a] T. Duyckaerts, C. Kenig, and F. Merle, "Concentration-compactness and universal profiles for the non-radial energy critical wave equation", *Nonlinear Anal.* **138** (2016), 44–82. MR Zbl

[Duyckaerts et al. 2016b] T. Duyckaerts, C. Kenig, and F. Merle, "Solutions of the focusing nonradial critical wave equation with the compactness property", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **15** (2016), 731–808. MR Zbl

[Fila and King 2012] M. Fila and J. R. King, "Grow up and slow decay in the critical Sobolev case", *Netw. Heterog. Media* **7**:4 (2012), 661–671. MR Zbl

[Filippas et al. 2000] S. Filippas, M. A. Herrero, and J. J. L. Velázquez, "Fast blow-up mechanisms for sign-changing solutions of a semilinear parabolic equation with critical nonlinearity", *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* **456**:2004 (2000), 2957–2982. MR Zbl

[Fujita 1966] H. Fujita, "On the blowing up of solutions of the Cauchy problem for $u_t = \Delta u + u^{1+\alpha}$", *J. Fac. Sci. Univ. Tokyo Sect. I* **13** (1966), 109–124. MR Zbl

[Ghoul and Masmoudi 2016] T.-E. Ghoul and N. Masmoudi, "Stability of infinite time blow up for the Patlak–Keller–Segel system", preprint, 2016. arXiv

[Gui et al. 1992] C. Gui, W.-M. Ni, and X. Wang, "On the stability and instability of positive steady states of a semilinear heat equation in $\mathbb{R}^n$", *Comm. Pure Appl. Math.* **45**:9 (1992), 1153–1181. MR Zbl

[Krieger et al. 2009] J. Krieger, W. Schlag, and D. Tataru, "Slow blow-up solutions for the $H^1(\mathbb{R}^3)$ critical focusing semilinear wave equation", *Duke Math. J.* **147**:1 (2009), 1–53. MR Zbl

[Krieger et al. 2015] J. Krieger, K. Nakanishi, and W. Schlag, "Center-stable manifold of the ground state in the energy space for the critical wave equation", *Math. Ann.* **361**:1-2 (2015), 1–50. MR Zbl

[Matano and Merle 2004] H. Matano and F. Merle, "On nonexistence of type II blowup for a supercritical nonlinear heat equation", *Comm. Pure Appl. Math.* **57**:11 (2004), 1494–1541. MR Zbl

[Matano and Merle 2009] H. Matano and F. Merle, "Classification of type I and type II behaviors for a supercritical nonlinear heat equation", *J. Funct. Anal.* **256**:4 (2009), 992–1064. MR Zbl

[Matano and Merle 2011] H. Matano and F. Merle, "Threshold and generic type I behaviors for a supercritical nonlinear heat equation", *J. Funct. Anal.* **261**:3 (2011), 716–748. MR Zbl

[Merle et al. 2013] F. Merle, P. Raphaël, and I. Rodnianski, "Blowup dynamics for smooth data equivariant solutions to the critical Schrödinger map problem", *Invent. Math.* **193**:2 (2013), 249–365. MR Zbl

[Ortoleva and Perelman 2013] C. Ortoleva and G. Perelman, "Nondispersive vanishing and blow up at infinity for the energy critical nonlinear Schrödinger equation in $\mathbb{R}^3$", *Algebra i Analiz* **25**:2 (2013), 162–192. In Russian; translated in *St. Petersburg Math. J.* **25**:2 (2014), 271–294. MR Zbl

[Poláčik and Quittner 2006] P. Poláčik and P. Quittner, "A Liouville-type theorem and the decay of radial solutions of a semilinear heat equation", *Nonlinear Anal.* **64**:8 (2006), 1679–1689. MR Zbl

[Poláčik and Yanagida 2003] P. Poláčik and E. Yanagida, "On bounded and unbounded global solutions of a supercritical semilinear heat equation", *Math. Ann.* **327**:4 (2003), 745–771. MR Zbl

[Poláčik and Yanagida 2014] P. Poláčik and E. Yanagida, "Global unbounded solutions of the Fujita equation in the intermediate range", *Math. Ann.* **360**:1-2 (2014), 255–266. MR Zbl

[Poláčik et al. 2007] P. Poláčik, P. Quittner, and P. Souplet, "Singularity and decay estimates in superlinear problems via Liouville-type theorems, II: Parabolic equations", *Indiana Univ. Math. J.* **56**:2 (2007), 879–908. MR Zbl

[Quittner and Souplet 2007] P. Quittner and P. Souplet, *Superlinear parabolic problems: blow-up, global existence and steady states*, Birkhäuser, Basel, 2007. MR Zbl

[Raphaël and Rodnianski 2012] P. Raphaël and I. Rodnianski, "Stable blow up dynamics for the critical co-rotational wave maps and equivariant Yang–Mills problems", *Publ. Math. Inst. Hautes Études Sci.* **115** (2012), 1–122. MR Zbl

[Raphaël and Schweyer 2013] P. Raphaël and R. Schweyer, "Stable blowup dynamics for the 1-corotational energy critical harmonic heat flow", *Comm. Pure Appl. Math.* **66**:3 (2013), 414–480. MR Zbl

[Rey 1990] O. Rey, "The role of the Green's function in a nonlinear elliptic equation involving the critical Sobolev exponent", *J. Funct. Anal.* **89**:1 (1990), 1–52. MR Zbl

[Schweyer 2012] R. Schweyer, "Type II blow-up for the four dimensional energy critical semi linear [sic] heat equation", *J. Funct. Anal.* **263**:12 (2012), 3922–3983. MR Zbl

Received 22 Apr 2018. Accepted 29 Dec 2018.

MANUEL DEL PINO: m.delpino@bath.ac.uk
Department of Mathematical Sciences, University of Bath, Bath, United Kingdom

MONICA MUSSO: m.musso@bath.ac.uk
Department of Mathematical Sciences, University of Bath, Bath, United Kingdom

JUNCHENG WEI: jcwei@math.ubc.ca
Department of Mathematics, University of British Columbia, Vancouver, BC, Canada

ANALYSIS AND PDE
Vol. 13, No. 1, 2020
dx.doi.org/10.2140/apde.2020.13.275

msp

A WELL-POSEDNESS RESULT FOR VISCOUS COMPRESSIBLE FLUIDS WITH ONLY BOUNDED DENSITY

RAPHAËL DANCHIN, FRANCESCO FANELLI AND MARIUS PAICU

We are concerned with the existence and uniqueness of solutions with only bounded density for the barotropic compressible Navier–Stokes equations. Assuming that the initial velocity has slightly subcritical regularity and that the initial density is a small perturbation (in the L^∞ norm) of a positive constant, we prove the existence of local-in-time solutions. In the case where the density takes two constant values across a smooth interface (or, more generally, has striated regularity with respect to some nondegenerate family of vector fields), we get uniqueness. This latter result supplements the work by D. Hoff (*Comm. Pure Appl. Math.* **55**:11 (2002), 1365–1407) with a uniqueness statement, and is valid in any dimension $d \geq 2$ and for general pressure laws.

Introduction	275
1. Main results	278
2. Tools	282
3. An existence statement for almost critical data and only bounded density	290
4. Tangential regularity and uniqueness	296
Appendix: Harmonic analysis estimates	308
Acknowledgements	315
References	315

Introduction

We are concerned with the multidimensional barotropic compressible Navier–Stokes system in the whole space:

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho u) = 0, \\ \partial_t(\rho u) + \operatorname{div}(\rho u \otimes u) - \mu \Delta u - \lambda \nabla \operatorname{div} u + \nabla P(\rho) = 0. \end{cases} \tag{0-1}$$

Here $\rho = \rho(t,x)$ and $u = u(t,x)$, with $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$ and $d \geq 1$, denote the density and velocity of the fluid, respectively. The pressure P is a given function of ρ. We shall take that function locally in $W^{1,\infty}$ in all that follows, and assume (with no loss of generality) that it vanishes at some constant reference density $\bar\rho > 0$. The (constant) viscosity coefficients μ and λ satisfy

$$\mu > 0 \quad \text{and} \quad \nu := \lambda + \mu > 0, \tag{0-2}$$

MSC2010: primary 35Q35; secondary 35B65, 76N10, 35B30, 35A02.
Keywords: compressible Navier–Stokes equations, bounded density, maximal regularity, tangential regularity, Lagrangian formulation.

which ensures ellipticity of the operator

$$\mathcal{L} := -\mu\Delta - \lambda\nabla\mathrm{div}. \tag{0-3}$$

We supplement system (0-1) with the initial conditions

$$\rho|_{t=0} = \rho_0 \quad \text{and} \quad u|_{t=0} = u_0. \tag{0-4}$$

A number of recent works have been dedicated to the study of solutions with discontinuous density (so-called "shock data") for models of viscous compressible fluids. Even though the situation is by now quite well understood for $d = 1$ (see, e.g., [Hoff 1991]), the multidimensional case is far from being completely elucidated. In this direction, we mention works of D. Hoff [1995; 2002], who greatly contributed with the construction of "intermediate solutions" allowing for discontinuity of the density, in between the weak solutions of P.-L. Lions [1998] and the classical ones of, e.g., J. Nash [1962].

In the two-dimensional case, Hoff [2002] succeeded in getting very accurate information on the propagation of density discontinuities across suitably smooth curves (as predicted by the Rankine–Hugoniot condition), under the assumption that the pressure is a linear function of ρ (see Theorem 1.2 therein). In particular, he proved that those curves are convected by the flow and keep their initial regularity *even though the gradient of the velocity is not continuous*. The result was strongly based on the observation that, for such solutions, the "effective viscous flux" $F := \nu\,\mathrm{div}\,u - P(\rho)$ is continuous, although singularities persist in $\mathrm{div}\,u$ and $P(\rho)$ separately.

The present paper aims at completing the aforementioned works in several directions.

First, we want to supplement them with a uniqueness result. Indeed, Hoff [2002] constructed solutions (that are global in time under some smallness assumption) and pointed out very accurate qualitative properties for the geometric structure of singularities, but did not address uniqueness. That latter issue has been considered afterward in [Hoff 2006], but only for *linear* pressure laws. In fact, the main uniqueness theorem therein requires either the pressure law to be linear (as opposed to the standard isentropic assumption $P(\rho) = a\rho^\gamma$ with $\gamma > 1$) or some Lebesgue-type information on $\nabla\rho$ (thus precluding us from considering jumps across interfaces). To the best of our knowledge, exhibiting an appropriate functional framework for uniqueness *without imposing a special structure for the solutions* has remained an open question for nonlinear pressure laws and discontinuous densities.

Our second goal is to extend Hoff's works concerning discontinuity across interfaces and uniqueness to any dimension and to more general pressure laws and density singularities.

Finding conditions on the initial data that ensure that ∇u belongs to $L^1([0, T]; L^\infty)$ for some $T > 0$ is the key to our two goals. That latter property will be achieved by combining parabolic maximal regularity estimates with *tangential* (or striated) *regularity* techniques that are borrowed from the work by J.-Y. Chemin [1991].

In order to introduce the reader to our use of maximal regularity, let us consider the slightly simpler situation of a fluid fulfilling the inhomogeneous incompressible Navier–Stokes equations:

$$\begin{cases} \partial_t\rho + \mathrm{div}(\rho u) = 0, \\ \partial_t(\rho u) + \mathrm{div}(\rho u \otimes u) - \mu\Delta u + \nabla\Pi = 0, \\ \mathrm{div}\,u = 0. \end{cases} \tag{0-5}$$

We have in mind the case where the initial density is given by

$$\rho_0 = c_1 \mathbb{1}_D + c_2 \mathbb{1}_{D^c} \tag{0-6}$$

for some positive constants c_1, c_2, with D being a smooth bounded domain of $\mathbb{R}^d$ (above, $\mathbb{1}_A$ designates the characteristic function of a set A).

Several recent works have been devoted to proposing conditions on u_0 allowing for solving (0-5) either locally or globally, and uniquely. The first result in that direction has been obtained by the first author and P. B. Mucha [Danchin and Mucha 2012]. It is based on endpoint maximal regularity and requires the jump $c_1 - c_2$ to be small enough. However, that approach requires the use of *multiplier spaces* to handle the low regularity of the density, and is very unlikely to be extendable to the compressible setting, owing to the pressure term that now depends on ρ.

Therefore, we shall rather take advantage of the approach that has been proposed recently in [Huang et al. 2013] by the third author together with J. Huang and P. Zhang to investigate (0-5) with only bounded density. Indeed, it is based on the standard parabolic maximal regularity and requires only very elementary tools like Hölder inequality and Sobolev embedding. In order to present the main steps, assume for simplicity that the reference density $\bar\rho$ is 1. Then, setting $\varrho := \rho - 1$, system (0-5) can be rewritten as

$$\begin{cases} \partial_t \varrho + u\cdot\nabla\varrho = 0,\\ \partial_t u - \mu\Delta u + \nabla\Pi = -\varrho\partial_t u - (1+\varrho)u\cdot\nabla u,\\ \operatorname{div} u = 0. \end{cases}$$

From basic maximal regularity estimates (recalled in Section 2B below), we have[1] for all $1 < p, r < \infty$,

$$\|u\|_{L^\infty_T(\dot B^{2-2/r}_{p,r})} + \|(\partial_t u, \mu\nabla^2 u)\|_{L^r_T(L^p)} \le C(\|u_0\|_{\dot B^{2-2/r}_{p,r}} + \|\varrho\partial_t u\|_{L^r_T(L^p)} + \|(1+\varrho)u\cdot\nabla u\|_{L^r_T(L^p)}).$$

It is obvious that the second term of the right-hand side may be absorbed by the left-hand side if the nonhomogeneity ϱ is small enough for the L^∞ norm. As for the last term, it may be absorbed either for a short time if the velocity is large, or for all times if the velocity is small, and the norm $L^r(L^p)$ is *scaling-invariant* for the incompressible Navier–Stokes equations, that is to say, satisfies

$$\frac{2}{r} + \frac{d}{p} = 3. \tag{0-7}$$

Those simple observations are the keys to the proof of global existence for (0-5) in [Huang et al. 2013]. As regards uniqueness, owing to the hyperbolic part of the system (viz. the first equation of (0-5)), we need (at least) a $L^1_T(\mathrm{Lip})$ control on the velocity. Note that if one combines the above control in $L^r_T(L^p)$ for $\nabla^2 u$ with the corresponding critical Sobolev embedding, then we miss that information by a little in the critical regularity setting, as having (0-7) and $r > 1$ implies that $p < d$; nonetheless, it turns out that, if working in a slightly subcritical framework (that is, $2/r + d/p < 3$), one can get existence and uniqueness together for any initial density $\rho_0 \in L^\infty$ that is close enough to some positive constant (see [Danchin and Zhang 2014; Huang et al. 2013] for more details).

[1]Throughout the paper we agree that, if E is a Banach space, $r \in [1, \infty]$ and $T > 0$, then $L^r_T(E)$ designates the space $L^r([0, T]; E)$ and $\|\cdot\|_{L^r_T(E)}$ the corresponding norm; when $T = \infty$, we use the notation $L^r(E)$.

In the particular case where ρ_0 is given by (0-6), once the Lipschitz control of the transport field is available, it is possible to propagate the Lipschitz regularity of the domain D, as it is just advected by the (Lipschitz continuous) flow of the velocity field. Based on that observation, further developments and more accurate information on the evolution of the boundary of D have been obtained very recently by X. Liao and P. Zhang [2016; 2019], and by the first author with X. Zhang [Danchin and Zhang 2017] and with Mucha [Danchin and Mucha 2017]. In most of those works, a key ingredient is the propagation of tangential regularity, in the spirit of the seminal work [Chemin 1991; 1993] dedicated to the vortex patch problem for the incompressible Euler equations. We refer also to papers [Gamblin and Saint Raymond 1995; Danchin 1997; 1999; Hmidi 2005] for extensions of the results of [Chemin 1991; 1993] to higher dimensions and to viscous homogeneous fluids, the case of nonhomogeneous inviscid flows being treated in [Fanelli 2012].

The rest of the paper is devoted to obtaining similar results for the compressible Navier–Stokes equations (0-1) and is structured as follows. In the next section, we present our main results and give some insight to the proofs. Then, in Section 2, we recall the definition of Besov spaces and introduce the tools for achieving our results: Littlewood–Paley decomposition, maximal regularity and estimates involving striated regularity. Section 3 is devoted to the proof of our main existence theorem for general discontinuous densities, while the next section concerns the propagation of striated regularity and uniqueness. Some technical results that are based on harmonic analysis are postponed until the Appendix.

1. Main results

In order to evaluate our chances of getting the same results for (0-1) as for (0-5) after suitable adaptation of the method described above, let us rewrite (0-1) in terms of (ϱ, u). We get, just denoting by P (instead of $P(1+\varrho)$) the pressure term, the system

$$\begin{cases}\partial_t\varrho + u\cdot\nabla\varrho + (\varrho+1)\,\mathrm{div}\,u = 0,\\ \partial_t u - \mu\Delta u - \lambda\nabla\,\mathrm{div}\,u = -\varrho\partial_t u - (1+\varrho)u\cdot\nabla u - \nabla P.\end{cases} \tag{1-1}$$

As the so-called Lamé system (that is, the left-hand side of the second equation) enjoys the same maximal regularity properties as the heat equation, one can handle the terms $\varrho\partial_t u$ and $(1+\varrho)u\cdot\nabla u$ in a suitable $L^r(L^p)$ framework exactly as in the incompressible situation. However, by this method, bounding ∇P requires $\nabla\varrho$ to be in some Lebesgue space, a condition that we want to avoid. In fact, the coupling between the density and the velocity equations is stronger than for (0-5) so that the two equations of (1-1) should not be considered separately. For that reason, it is much more difficult to prove a well-posedness result for rough densities here than in the incompressible case, and the presence of the “out-of-scaling” term ∇P precludes us from achieving global existence (even for small data) by simple arguments. A standard way to weaken the coupling between the two equations of (1-1) (that has been used by Hoff [1995; 2002] and, more recently, by B. Haspot [2011] or by the first author with L. He [Danchin and He 2016]) is to reformulate the system in terms of a “modified velocity field”, in the same spirit as the effective viscous flux mentioned in the Introduction: we set

$$w := u + \nabla(-\nu\Delta)^{-1}P, \quad \text{with } \nu := \lambda+\mu.$$

The modified velocity w absorbs the main part of ∇P, as may be observed when writing out the equation for w:

$$(1+\varrho)\partial_t w - \mu\Delta w - \lambda\nabla \operatorname{div} w + (1+\varrho)u\cdot\nabla u + (1+\varrho)(-\nu\Delta)^{-1}\nabla\partial_t P = 0. \tag{1-2}$$

As, by virtue of the mass equation, we have

$$\partial_t P = (P-\rho P')\operatorname{div} u - \operatorname{div}(Pu),$$

the last term of (1-2) is indeed of lower order and can be bounded in $L^r_T(L^p)$ whenever ϱ is bounded and belongs to some suitable Lebesgue space.

As we shall see in the present paper, working with w and ϱ rather than with the original unknowns u and ρ proves to be efficient if one is concerned with existence (and possibly uniqueness) results for (0-1) with only bounded density. In fact, we shall implement the maximal regularity estimates on the equation fulfilled by w, and bound ϱ by means of the standard a priori estimates in Lebesgue spaces for the transport equation. For technical reasons, however, it will be wise to replace $(-\nu\Delta)^{-1}$ by its nonhomogeneous version $(\mathrm{Id}-\nu\Delta)^{-1}$, which is much less singular.

That strategy will enable us to prove the following local-in-time result of existence for (0-1) supplemented with a rough initial density.[2]

Theorem 1.1. *Let $d \geq 1$. Let the pair (p, r) satisfy*

$$d < p < \infty \quad \text{and} \quad 1 < r < \frac{2p}{2p-d}, \tag{1-3}$$

and define the pair of indices (r_0, r_1) by the relations

$$\frac{1}{r_0} = \frac{1}{r} - 1 + \frac{d}{2p} \quad \text{and} \quad \frac{1}{r_1} = \frac{1}{r} - \frac{1}{2}. \tag{1-4}$$

Let the initial density ρ_0 and velocity u_0 satisfy

$$\varrho_0 := \rho_0 - 1 \quad \text{in } (L^p\cap L^\infty)(\mathbb{R}^d),$$
$$w_0 := u_0 - v_0 \quad \text{in } \dot B^{2-2/r}_{p,r}, \quad \text{with } v_0 := -\nabla(\mathrm{Id}-\nu\Delta)^{-1}(P(\rho_0)).$$

There exist $\varepsilon > 0$ and a time $T > 0$ such that, if

$$\|\varrho_0\|_{L^\infty} \leq \varepsilon, \tag{1-5}$$

then there exists a solution (ρ, u) to system (0-1)–(0-4) on $[0,T]\times\mathbb{R}^d$ with $\varrho := \rho - 1$ satisfying

$$\|\varrho\|_{L^\infty([0,T]\times\mathbb{R}^d)} \leq 4\varepsilon \quad \text{and} \quad \varrho \in \mathcal{C}([0,T]; L^q) \quad \text{for all } p \leq q < \infty,$$

and $u = v + w$ with $v := -\nabla(\mathrm{Id}-\nu\Delta)^{-1}(P(\rho))$ in $\mathcal{C}([0,T]; W^{1,q}(\mathbb{R}^d))$ for all $p \leq q < \infty$, and

$$w \in \mathcal{C}([0,T]; \dot B^{2-2/r}_{p,r}) \cap L^{r_0}_T(L^\infty), \quad \nabla w \in L^{r_1}_T(L^p) \quad \text{and} \quad \partial_t w, \nabla^2 w \in L^r_T(L^p).$$

That solution is unique if $d = 1$.

[2] The reader is referred to Section 2 below for the definition of the homogeneous Besov spaces $\dot B^s_{p,r}$.

If, in addition, u_0 and ϱ_0 are in L^2, and $\inf P' > 0$ *on* $[1-4\varepsilon, 1+4\varepsilon]$, *then the energy balance*

$$\tfrac{1}{2}\|\sqrt{\rho(t)}u(t)\|_{L^2}^2 + \|\Pi(\rho(t))\|_{L^1} + \mu\|\nabla u\|_{L_t^2(L^2)}^2 + \lambda\|\operatorname{div} u\|_{L_t^2(L^2)}^2 = \tfrac{1}{2}\|\sqrt{\rho_0}u_0\|_{L^2}^2 + \|\Pi(\rho_0)\|_{L^1} \quad (1\text{-}6)$$

holds true for all $t \in [0, T]$, *where the function* $\Pi = \Pi(z)$ *is defined by the conditions* $\Pi(1) = \Pi'(1) = 0$ *and* $\Pi''(z) = P'(z)/z$.

Combining the above statement with Sobolev embeddings ensures that ∇w and $\operatorname{div} u$ are in $L_T^1(L^\infty)$. However, because the operator $\nabla^2(\mathrm{Id} - \nu\Delta)^{-1}$ does not quite map L^∞ into itself (except if $d = 1$ of course), there is no guarantee that the constructed velocity field u has gradient in $L_T^1(L^\infty)$. This seems to be the minimal requirement in order to get uniqueness of solutions (see, e.g., [Hoff 2002; Danchin 2014]). Keeping the model case (0-6) in mind, the question is whether adding up geometric hypotheses, like interfaces or tangential regularity, ensures that property and, hopefully, uniqueness.

Motivated by the pioneering work [Chemin 1991; 1993], we shall assume that the initial density has some "striated regularity" along a nondegenerate family of vector fields. To be more specific, we have to introduce more notation and give some definitions. Before doing that, let us underline that propagating tangential regularity for compressible flows means facing new difficulties compared to the incompressible case, due to the fact that $\operatorname{div} u$ does not vanish anymore.

First of all, for any p in $]d, \infty]$, we denote by $\mathbb{L}^{\infty,p}$ the space of all continuous and bounded functions with gradient in $L^p(\mathbb{R}^d)$. Now, for a given vector field Y in $\mathbb{L}^{\infty,p}$, we are interested in the regularity of a function f along Y, i.e., in the quantity

$$\partial_Y f := \sum_{j=1}^{d} Y^j \partial_j f.$$

This expression is well-defined if f is smooth enough, in which case we have the identity

$$\partial_Y f = \operatorname{div}(fY) - f \operatorname{div} Y. \quad (1\text{-}7)$$

If f is only bounded (which, typically, will be the case if f is the density given by (0-6)), the above right-hand side makes sense for any vector field Y in $\mathbb{L}^{\infty,p}$, while $\partial_Y f$ has no meaning. Then we take the right-hand side of (1-7) as a definition of $\partial_Y f$.

In order to define striated regularity, fix a family $\mathcal{X} = (X_\lambda)_{1\le\lambda\le m}$ of m vector fields with components in $\mathbb{L}^{\infty,p}$ and suppose that it is *nondegenerate*, in the sense that

$$I(\mathcal{X}) := \inf_{x\in\mathbb{R}^d} \sup_{\Lambda\in\Lambda_{d-1}^m} \left|\overset{d-1}{\wedge} X_\Lambda(x)\right|^{1/(d-1)} > 0.$$

Here $\Lambda \in \Lambda_{d-1}^m$ means that $\Lambda = (\lambda_1, \dots, \lambda_{d-1})$, with $\lambda_i \in \{1, \dots, m\}$ for all i and $\lambda_i < \lambda_j$ for $i < j$, while the symbol $\overset{d-1}{\wedge} X_\Lambda$ stands for the unique element of $\mathbb{R}^d$ such that

$$\text{for all } Y \in \mathbb{R}^d, \quad \left(\overset{d-1}{\wedge} X_\Lambda\right)\cdot Y = \det(X_{\lambda_1}, \dots X_{\lambda_{d-1}}, Y).$$

Then we set

$$\|X_\lambda\|_{\mathbb{L}^{\infty,p}} := \|X_\lambda\|_{L^\infty} + \|\nabla X_\lambda\|_{L^p} \quad \text{and} \quad \|\!|\mathcal{X}|\!\|_{\mathbb{L}^{\infty,p}} := \sup_{\lambda\in\Lambda} \|X_\lambda\|_{\mathbb{L}^{\infty,p}}.$$

More generally, whenever E is a normed space, we use the notation $\|\!|\mathcal{X}|\!\|_E := \sup_{\lambda\in\Lambda} \|X_\lambda\|_E$.

Definition 1.2. Take a vector field $Y \in \mathbb{L}^{\infty,p}$ for some $p \in]d, \infty]$. Given a function $f \in L^\infty$, we say that f belongs to $\mathbb{L}^p_Y$ if $\operatorname{div}(fY) \in L^p(\mathbb{R}^d)$.

If $\mathcal{X} = (X_\lambda)_{1\le\lambda\le m}$ is a nondegenerate family of vector fields in $\mathbb{L}^{\infty,p}$ then we set

$$\mathbb{L}^p_{\mathcal{X}} := \bigcap_{1\le\lambda\le m} \mathbb{L}^p_{X_\lambda} \quad \text{and} \quad \|f\|_{\mathbb{L}^p_{\mathcal{X}}} := \frac{1}{I(\mathcal{X})}\big(\|f\|_{L^\infty}|||\mathcal{X}|||_{\mathbb{L}^{\infty,p}} + |||\operatorname{div}(f\mathcal{X})|||_{L^p}\big).$$

The main motivation for Definition 1.2 is Proposition 2.12 below, which states that, if ϱ is bounded and if, in addition, $\varrho \in \mathbb{L}^p_{\mathcal{X}}$ for some nondegenerate family $\mathcal{X}$ of vector fields in $\mathbb{L}^{\infty,p}$ with $d < p < \infty$, then $(\eta\mathrm{Id} - \Delta)^{-1}\nabla^2 f(\varrho)$ is in L^∞ for all $\eta > 0$ and smooth enough function f. That fundamental property will enable us to consider data like (0-6) in a functional framework that ensures persistence of interface regularity and uniqueness together, as stated just below.

Theorem 1.3. *Let $d \ge 1$ and the pair (p, r) fulfill conditions (1-3). Consider initial data (ρ_0, u_0) satisfying the same assumptions as in Theorem 1.1. Assume in addition that there exists a nondegenerate family $\mathcal{X}_0 = (X_{0,\lambda})_{1\le\lambda\le m}$ of vector fields in $\mathbb{L}^{\infty,p}$ such that ρ_0 belongs to $\mathbb{L}^p_{\mathcal{X}_0}$.*

Then, there exists a time $T > 0$ and a unique solution (ρ, u) to system (0-1)–(0-4) on $[0, T] \times \mathbb{R}^d$ such that $\varrho := \rho - 1$, $v := -\nabla(\mathrm{Id} - \nu\Delta)^{-1}(P(\rho))$ and $w := u - v$ satisfy the same properties as in Theorem 1.1. Furthermore, ∇u belongs to $L^1([0, T]; L^\infty(\mathbb{R}^d))$ and u has a flow ψ_u with bounded gradient that is the unique solution of

$$\psi_u(t, x) = x + \int_0^t u(\tau, \psi_u(\tau, x))\, d\tau \quad \text{for all } (t, x) \in [0, T] \times \mathbb{R}^d. \tag{1-8}$$

Finally, if we define $X_{t,\lambda}$ by the formula $X_{t,\lambda}(x) := \partial_{X_{0,\lambda}}\psi_u(t, \psi_u^{-1}(t, x))$ then, for all times $t \in [0, T]$, the family $\mathcal{X}_t := (X_{t,\lambda})_{1\le\lambda\le m}$ remains nondegenerate and in the space $\mathbb{L}^{\infty,p}$, and the density $\rho(t)$ belongs to $\mathbb{L}^p_{\mathcal{X}_t}$.

Let us make some comments on the proof of that second main result. As pointed out above, the striated regularity hypothesis ensures that $(\mathrm{Id} - \nu\Delta)^{-1}\nabla^2 P$ is bounded. Since we know from Theorem 1.1 that ∇w is in $L^1_T(L^\infty)$, one can conclude that also ∇u belongs to $L^1_T(L^\infty)$. From this property and the remark that, for all $\lambda \in \Lambda$, one has

$$\partial_t X_\lambda + u \cdot \nabla X_\lambda = \partial_{X_\lambda} u \quad \text{and} \quad \partial_t \operatorname{div}(\rho X_\lambda) + \operatorname{div}(\operatorname{div}(\rho X_\lambda)u) = 0,$$

standard estimates for the transport equation will enable us to propagate the tangential regularity.

As regards the proof of uniqueness, we adopt the viewpoint in [Hoff 2006]: *solutions with minimal regularity are best if compared in a Lagrangian framework; that is, we compare the instantaneous states of corresponding fluid particles in two different solutions rather than the states of different fluid particles instantaneously occupying the same point of space-time.* However, the proof that is proposed therein relies on stability estimates in the negative Sobolev space $\dot H^{-1}$ for the density; at some point, it is crucial that $\rho \in \dot H^{-1}$ implies $P(\rho)$ is in $\dot H^{-1}$, too, a property that obviously fails when the pressure law is nonlinear.

In the present paper, adopting the Lagrangian viewpoint will enable us to avoid (for general pressure laws) the loss of one derivative due to the hyperbolic part of system (0-1). As a matter of fact, we shall establish stability estimates for the (Lagrangian) velocity field directly in the energy space, and the presence of variable coefficients owing to the initial density variations, either in front of the time derivative or in the elliptic part of the evolution operator, will be harmless.

Let us finally state an important application of Theorem 1.3.

Corollary 1.4. *Assume that ρ_0 is given by (0-6) for some bounded domain D of class $W^{2,p}$ with $d < p < \infty$.*

Then, there exists $\varepsilon > 0$ such that, if $|c_1 - c_2| \le \varepsilon$, then for any initial velocity field u_0 satisfying the conditions of Theorem 1.1, there exists $T > 0$ such that system (0-1)–(0-4) admits a unique solution (ρ, u). Furthermore, the density at time t has a jump discontinuity along the interface of the domain D_t transported by the flow of u, and ∂D_t keeps its $W^{2,p}$ regularity.

We end this section with a list of possible extensions/improvements of our paper.

(1) All our results may be readily adapted to the case of periodic boundary conditions; indeed, our techniques rely on Fourier analysis and thus hold true for functions defined on the torus.

(2) We expect a similar existence statement if the fluid domain is a bounded open set Ω with (say) $\mathcal{C}^2$ boundary, and the system is supplemented with homogeneous Dirichlet boundary conditions for the velocity. Indeed, in that setting, the Besov spaces may be defined by real interpolation from the domain of the Lamé (or, equivalently the heat) operator and the maximal regularity estimates remain the same. The reader may refer to [Danchin 2010] for an example of solving (0-1) in that setting, in the case of density in $W^{1,p}$ for some $p > d$.

Concerning the propagation of tangential regularity, the situation where the reference family of vector fields does not degenerate at the boundary should be tractable with few changes (this is a matter of adapting the work of N. Depauw [1999] to our system). This means that one can consider initial densities like (0-6) provided the boundary of D does not meet that of Ω.

(3) To keep the paper a reasonable size, we refrained from considering the global existence issue for rough densities. We plan to address that interesting question in the near future.

2. Tools

Here we introduce the main tools for our analysis. First of all, we recall basic facts about Littlewood–Paley theory and Besov spaces. The next subsection is devoted to maximal regularity results. Finally, in Section 2C we present key inequalities involving striated regularity.

2A. ***Littlewood–Paley theory and Besov spaces.*** We here briefly present Littlewood–Paley theory, as it will come into play for proving our main result. We refer, e.g., to Chapter 2 of [Bahouri et al. 2011] for more details. For simplicity of exposition, we focus on the $\mathbb{R}^d$ case; however, the whole construction can be adapted to the d-dimensional torus $\mathbb{T}^d$.

First of all, let us introduce the so-called "Littlewood–Paley decomposition". It is based on a nonhomogeneous dyadic partition of unity with respect to the Fourier variable: fix a smooth radial function χ supported in the ball $B(0,2)$, equal to 1 in a neighborhood of $B(0,1)$ and such that $r \mapsto \chi(re)$ is nonincreasing over $\mathbb{R}_+$ for all unitary vectors $e \in \mathbb{R}^d$. Set

$$\varphi(\xi) = \chi(\xi) - \chi(2\xi) \quad \text{and} \quad \varphi_j(\xi) := \varphi(2^{-j}\xi) \quad \text{for all } j \geq 0.$$

The nonhomogeneous dyadic blocks $(\Delta_j)_{j\in\mathbb{Z}}$ are defined by[3]

$$\Delta_j := 0 \quad \text{if } j \leq -2, \qquad \Delta_{-1} := \chi(D) \qquad \text{and} \qquad \Delta_j := \varphi(2^{-j}D) \quad \text{if } j \geq 0.$$

We also introduce the following low-frequency cut-off operator:

$$S_j := \chi(2^{-j}D) = \sum_{k \leq j-1} \Delta_k \quad \text{for } j \geq 0, \qquad \text{and} \qquad S_j = 0 \quad \text{for } j < 0. \tag{2-1}$$

It is well known that for any $u \in \mathcal{S}'$, one has the equality

$$u = \sum_{j \geq -1} \Delta_j u \quad \text{in } \mathcal{S}'.$$

Sometimes, we shall alternatively use the spectral cut-offs $\dot{\Delta}_j$ and $\dot{S}_j$ that are defined by

$$\dot{\Delta}_j := \varphi(2^{-j}D) \quad \text{and} \quad \dot{S}_j = \chi(2^{-j}D) \quad \text{for all } j \in \mathbb{Z}.$$

Note that we have

$$u = \sum_{j\in\mathbb{Z}} \dot{\Delta}_j u \tag{2-2}$$

up to polynomials only, which makes decomposition (2-2) unwieldy. A way to have equality in (2-2) in the sense of tempered distributions is to restrict oneself to elements u of the set $\mathcal{S}'_h$ of tempered distributions such that

$$\lim_{j\to-\infty} \|\dot{S}_j u\|_{L^\infty} = 0.$$

It is now time to introduce Besov spaces.

Definition 2.1. Let $s \in \mathbb{R}$ and $1 \leq p, r \leq \infty$:

(i) The *nonhomogeneous Besov space* $B^s_{p,r}$ is the set of tempered distributions u for which

$$\|u\|_{B^s_{p,r}} := \|(2^{js}\|\Delta_j u\|_{L^p})_{j\geq -1}\|_{\ell^r} < \infty.$$

(ii) The *homogeneous Besov space* $\dot{B}^s_{p,r}$ is the subset of distributions u in $\mathcal{S}'_h$ such that

$$\|u\|_{\dot{B}^s_{p,r}} := \|(2^{js}\|\dot{\Delta}_j u\|_{L^p})_{j\in\mathbb{Z}}\|_{\ell^r} < \infty.$$

It is well known that $B^s_{2,2}$ coincides with H^s (with equivalent norms) and that nonhomogeneous (resp. homogeneous) Besov spaces are interpolation spaces between Sobolev spaces $W^{k,p}$ (resp. $\dot{W}^{k,p}$). Furthermore, for all $p \in]1, \infty[$, one has the following continuous embeddings (see the proof in [Bahouri et al. 2011, Chapter 2]):

$$\dot{B}^0_{p,\min(p,2)} \hookrightarrow L^p \hookrightarrow \dot{B}^0_{p,\max(p,2)} \quad \text{and} \quad B^0_{p,\min(p,2)} \hookrightarrow L^p \hookrightarrow B^0_{p,\max(p,2)}.$$

We shall also often use the embeddings that are stated in the following proposition.

[3]Throughout $f(D)$ stands for the pseudodifferential operator $u \mapsto \mathcal{F}^{-1}(f\mathcal{F}u)$.

Proposition 2.2. *Let $1 \le p_1 \le p_2 \le \infty$. The space $B^{s_1}_{p_1,r_1}(\mathbb{R}^d)$ is continuously embedded in the space $B^{s_2}_{p_2,r_2}(\mathbb{R}^d)$ if*

$$s_2 < s_1 - d\Big(\frac{1}{p_1} - \frac{1}{p_2}\Big) \qquad \text{or} \qquad s_2 = s_1 - d\Big(\frac{1}{p_1} - \frac{1}{p_2}\Big) \quad \text{and} \quad r_1 \le r_2.$$

The space $\dot B^{s_1}_{p_1,r_1}(\mathbb{R}^d)$ is continuously embedded in the space $\dot B^{s_2}_{p_2,r_2}(\mathbb{R}^d)$ if

$$s_2 = s_1 - d\Big(\frac{1}{p_1} - \frac{1}{p_2}\Big) \quad \text{and} \quad r_1 \le r_2.$$

Finally, we shall need the following continuity result.

Lemma 2.3. *There exists a constant C, depending only on d, such that for all $p \in [1,\infty]$ we have*

$$\|\Delta(\mathrm{Id}-\Delta)^{-1} f\|_{L^p} \le C\|f\|_{L^p}.$$

Proof. It suffices to notice that

$$\Delta(\mathrm{Id}-\Delta)^{-1} f = (\mathrm{Id}-\Delta)^{-1} f - f$$

and that $(\mathrm{Id}-\Delta)^{-1}$ maps L^p to $B^2_{p,\infty}$ (see Proposition 2.78 of [Bahouri et al. 2011]) and hence to L^p, with a constant independent of p. □

Corollary 2.4. *Let u solve the elliptic equation $(\mathrm{Id}-\Delta)u = f$ in $\mathbb{R}^d$, with $f \in L^p$ for some $p \in [1,\infty]$. Then $u \in W^{1,p}(\mathbb{R}^d)$, with $\Delta u \in L^p$, and one has the estimate*

$$\|u\|_{W^{1,p}} + \|\Delta u\|_{L^p} \le C\|f\|_{L^p}$$

for some positive constant C depending just on d.

If $1 < p < \infty$, then $u \in W^{2,p}$ and we have

$$\|u\|_{W^{2,p}} \le C\|f\|_{L^p}.$$

Proof. From Lemma 2.3, we gather that $\Delta u \in L^p$; hence $u = \Delta u + f$ belongs to L^p too. This relation in particular implies the control $\|u\|_{L^p} \le C\|f\|_{L^p}$. At this point, the control of the gradient of u in L^p follows, e.g., from Gagliardo–Nirenberg inequalities (or a decomposition into low and high frequencies). Finally, in the case $1 < p < \infty$, having $\Delta u \in L^p$ implies $\nabla^2 u \in L^p$ by Calderón–Zygmund theory. □

2B. *Maximal regularity and propagation of L^p norms.* In this subsection we recall some results about maximal regularity for the heat equation, and then extend them to the elliptic operator $\mathcal{L}$ defined in (0-3). Those results will be essentially the key to Theorem 1.1, namely existence of solutions in an L^p setting.

2B1. *The case of the heat kernel.* Here we focus on maximal regularity results for the heat semigroup, as they will lead to similar ones for the Lamé semigroup generated by $-\mathcal{L}$ (see Section 2B2 below). We first look at the propagation of regularity for the initial datum. Our starting point is the proposition below, which corresponds to Theorem 2.34 of [Bahouri et al. 2011].

Proposition 2.5. *Let $s > 0$ and $(p,r) \in [1,\infty]^2$. A constant C exists such that*

$$C^{-1}\|z\|_{\dot B^{-s}_{p,r}} \le \big\| \|t^{s/2} e^{t\Delta} z\|_{L^p} \big\|_{L^r(\mathbb{R}_+;dt/t)} \le C\|z\|_{\dot B^{-s}_{p,r}}.$$

Thus we deduce that, for all $r \in [1,\infty[$, having z in $\dot B^{-2/r}_{p,r}(\mathbb{R}^d)$ is equivalent to the condition $e^{t\Delta}z \in L^r(\mathbb{R}_+; L^p(\mathbb{R}^d))$. In particular, taking $z = \Delta u_0$ and assuming that u_0 has *critical regularity* $\dot B^{-1+d/p}_{p,r}(\mathbb{R}^d)$ for some $p \in]1,\infty[$ and r fulfilling

$$2-\frac{2}{r}=\frac{d}{p}-1 \quad \text{with } 1<r<\infty, \tag{2-3}$$

Proposition 2.5 combined with the classical L^p theory for the Laplace operator implies that $\nabla^2 e^{t\Delta}u_0$ is in $L^r(\mathbb{R}_+; L^p(\mathbb{R}^d))$.

Note however that (2-3) gives the constraint $d/3 < p \le d$, which is too restrictive for our scope: we will need $p > d$ in order to guarantee that ∇u is in $L^1([0,T]; L^\infty(\mathbb{R}^d))$ for some $T > 0$ (see Section 3 for more details). This fact will preclude us from working in the critical regularity setting.

Before going on, let us introduce more notation: throughout this section, we will use an index j to designate the regularity of Lebesgue exponents $(p_j, r_j) \in [1,\infty]^2$ pertaining to the term $\nabla^j h$.

According to Proposition 2.5, if u_0 is in $\dot B^{s_2}_{p_2,r_2}$ with $s_2 = 2-2/r_2$ and $1 < p_2, r_2 < \infty$, then

$$\nabla^2 e^{t\Delta}u_0 \in L^{r_2}(\mathbb{R}_+; L^{p_2}(\mathbb{R}^d)). \tag{2-4}$$

Furthermore, we have

$$e^{t\Delta}u_0 \in \mathcal{C}_b(\mathbb{R}_+; \dot B^{s_2}_{p_2,r_2}) \tag{2-5}$$

since

$$\|e^{t\Delta}u_0\|_{\dot B^{2-2/r_2}_{p_2,r_2}} \sim \|\nabla^2 e^{t\Delta}u_0\|_{\dot B^{-2/r_2}_{p_2,r_2}} \sim \left(\int_{\mathbb{R}_+} \|e^{\tau\Delta}(\Delta e^{t\Delta}u_0)\|^{r_2}_{L^{p_2}}\,d\tau\right)^{1/r_2}$$

and, using the fact that the heat semigroup is contractive on L^{p_2},

$$\int_{\mathbb{R}_+} \|e^{\tau\Delta}(\Delta e^{t\Delta}u_0)\|^{r_2}_{L^{p_2}}\,d\tau \le \int_{\mathbb{R}_+} \|e^{\tau\Delta}\Delta u_0\|^{r_2}_{L^{p_2}}\,d\tau \le C\|\Delta u_0\|^{r_2}_{\dot B^{-2/r_2}_{p_2,r_2}}.$$

Time continuity in (2-5) just follows from the fact that $\mathcal{S}$ is densely embedded in L^{p_2}.

Next, by the embedding properties of Proposition 2.2, we have, if $p_1 \ge p_2$ and $r_1 \ge r_2$,

$$\nabla u_0 \in \dot B^{s_1}_{p_1,r_1} \quad \text{with } s_1 = 1-\frac{2}{r_2}-d\Big(\frac{1}{p_2}-\frac{1}{p_1}\Big).$$

In order to be in the position to apply Proposition 2.5 so as to get that $\nabla e^{t\Delta}u_0 \in L^{r_1}(\mathbb{R}_+; L^{p_1}(\mathbb{R}^d))$, we need to have in addition $s_1 < 0$, that is to say,

$$\frac{2}{r_2}+\frac{d}{p_2}-\frac{d}{p_1} > 1. \tag{2-6}$$

Then, defining r_1 by

$$\frac{2}{r_2}+\frac{d}{p_2} = 1+\frac{2}{r_1}+\frac{d}{p_1}, \tag{2-7}$$

we get

$$\nabla e^{t\Delta}u_0 \in L^{r_1}(\mathbb{R}_+; L^{p_1}(\mathbb{R}^d)). \tag{2-8}$$

Finally, let us consider $e^{t\Delta}u_0$. By critical embedding, we have, if $p_0 \geq p_2$ and $r_0 \geq r_2$,

$$u_0 \in \dot{B}^{s_0}_{p_0,r_0}, \quad \text{with } s_0 = 2 - \frac{2}{r_2} - d\Big(\frac{1}{p_2} - \frac{1}{p_0}\Big).$$

If we want to resort again to Proposition 2.5, we need to have in addition that

$$s_0 = 2 - \frac{2}{r_2} - \frac{d}{p_1} < 0. \tag{2-9}$$

Under that condition, choosing r_0 so that $s_0 = -2/r_0$, that is to say, such that

$$\frac{2}{r_2} + \frac{d}{p_2} = 2 + \frac{2}{r_0} + \frac{d}{p_0}, \tag{2-10}$$

we end up with

$$e^{t\Delta}u_0 \in L^{r_0}(\mathbb{R}_+; L^{p_0}(\mathbb{R}^d)). \tag{2-11}$$

Let us next consider the propagation of regularity for the forcing term in the heat equation. We start by presenting the standard maximal $L^r(L^p)$ regularity for the heat semigroup (see the proof in [Lemarié-Rieusset 2002, Lemma 7.3] for instance).

Lemma 2.6. *Let us define the operator $\mathcal{A}_2$ by the formula*

$$\mathcal{A}_2 : f \mapsto \int_0^t \nabla^2 e^{(t-s)\Delta} f(s,\cdot)\,ds.$$

Then $\mathcal{A}_2$ is bounded from $L^{r_2}(]0,T[; L^{p_2}(\mathbb{R}^d))$ to $L^{r_2}(]0,T[; L^{p_2}(\mathbb{R}^d))$ for every $T \in]0,\infty]$ and $1 < p_2, r_2 < \infty$. Moreover, there holds

$$\|\mathcal{A}_2 f\|_{L^{r_2}_T(L^{p_2})} \leq C\|f\|_{L^{r_2}_T(L^{p_2})}.$$

As for the propagation of regularity for the first derivatives, we have the following statement.

Lemma 2.7. *Assume that the Lebesgue exponents p_1 and p_2 fulfill $0 \leq 1/p_2 - 1/p_1 < 1/d$ and that $1 < r_2 < r_1 < \infty$ are related by (2-7). Let us define the operator $\mathcal{A}_1$ by*

$$\mathcal{A}_1 : f \mapsto \int_0^t \nabla e^{(t-s)\Delta} f(s,\cdot)\,ds.$$

Then $\mathcal{A}_1$ is bounded from $L^{r_2}(]0,T[; L^{p_2}(\mathbb{R}^d))$ to $L^{r_1}(]0,T[; L^{p_1}(\mathbb{R}^d))$ for every $T \in]0,\infty[$, and there holds

$$\|\mathcal{A}_1 f\|_{L^{r_1}_T(L^{p_1})} \leq C\|f\|_{L^{r_2}_T(L^{p_2})}$$

for a suitable constant $C > 0$ depending only on the space dimension $d \geq 1$ and on p_1, r_1, p_2, r_2.

Proof. We use the fact that for all $0 \leq s \leq t \leq T$

$$\begin{aligned}\nabla e^{(t-s)\Delta} f(s,x) &= \frac{\sqrt{\pi}}{(4\pi(t-s))^{(d+1)/2}} \int_{\mathbb{R}^d} \frac{(x-y)}{2\sqrt{(t-s)}} \exp\Big(-\frac{|x-y|^2}{4(t-s)}\Big) f(s,y)\,dy \\ &:= \frac{\sqrt{\pi}}{(4\pi(t-s))^{(d+1)/2}} K_1\Big(\frac{\cdot}{\sqrt{4(t-s)}}\Big) *_x f(s,\cdot).\end{aligned}$$

Applying Young's inequality in the space variables yields

$$\|\nabla e^{(t-s)\Delta}(\mathbb{1}_{[0,T]}f)(s,\cdot)\|_{L^{p_1}} \le C(t-s)^{-(d+1)/2}\left\|K_1\left(\frac{\cdot}{\sqrt{4\pi(t-s)}}\right)\right\|_{L^{m_1}}\|\mathbb{1}_{[0,T]}f\|_{L^{p_2}}$$
$$\le C(t-s)^{-\beta}\|\mathbb{1}_{[0,T]}f\|_{L^{p_2}},$$

where we have defined m_1 and β by

$$\frac{1}{p_1}+1=\frac{1}{m_1}+\frac{1}{p_2} \quad \text{and} \quad \beta=\frac{d+1}{2}-\frac{d}{2m_1}=\frac{1}{2}+\frac{d}{2p_2}-\frac{d}{2p_1}.$$

Note that the conditions in the lemma ensure that $\beta\in\left[\frac{1}{2},1\right[$. At this point, we apply the Hardy–Littlewood–Sobolev inequality (see, e.g., Theorem 1.7 of [Bahouri et al. 2011]) with respect to time: since r_1 and r_2 satisfy $1/r_1+1=1/r_2+\beta$ by hypothesis (2-7), we immediately get the claimed inequality. The lemma is thus proved. □

We now state integrability properties concerning f itself, without taking any derivative.

Lemma 2.8. *Assume that the Lebesgue exponents p_0 and p_2 fulfill $0\le 1/p_2-1/p_0<2/d$ and that $1<r_2<r_0<\infty$ are related by (2-10). Define the operator $\mathcal{A}_0$ by the formula*

$$\mathcal{A}_0 : f\mapsto \int_0^t e^{(t-s)\Delta}f(s,\cdot)\,ds.$$

Let $s_2:=2-2/r_2$. Then $\mathcal{A}_0$ is bounded from $L^{r_2}(]0,T[;L^{p_2})$ to $L^{r_0}(]0,T[;L^{p_0})\cap\mathcal{C}([0,T];\dot B^{s_2}_{p_2,r_2})$ for every $T\in]0,\infty]$, and there holds

$$\|\mathcal{A}_0 f\|_{L^\infty_T(\dot B^{s_2}_{p_2,r_2})}+\|\mathcal{A}_0 f\|_{L^{r_0}_T(L^{p_0})}\le C\|f\|_{L^{r_2}_T(L^{p_2})}$$

for a suitable constant $C>0$ depending on the space dimension $d\ge 1$ and on p_0, r_0, p_2, r_2.

Proof. The proof of the continuity in $L^{r_0}(]0,T[;L^{p_0})$ goes as in Lemma 2.7: We start by writing

$$e^{(t-s)\Delta}f(s,x)=\frac{1}{(4\pi(t-s))^{d/2}}\int_{\mathbb{R}^d}\exp\left(-\frac{|x-y|^2}{4(t-s)}\right)f(s,y)\,dy$$
$$:=\frac{1}{(4\pi(t-s))^{d/2}}K_0\left(\frac{\cdot}{\sqrt{4(t-s)}}\right)*_x f(s,\cdot).$$

Then, we apply Young's inequality in the space variables and get, for every $s>0$ fixed,

$$\|e^{(t-s)\Delta}(\mathbb{1}_{[0,T]}f)(s,\cdot)\|_{L^{p_0}}\le C(t-s)^{-d/2}\left\|K_0\left(\frac{\cdot}{\sqrt{4\pi(t-s)}}\right)\right\|_{L^{m_0}}\|\mathbb{1}_{[0,T]}f\|_{L^{p_2}}$$
$$\le C(t-s)^{-\gamma}\|\mathbb{1}_{[0,T]}f\|_{L^{p_2}},$$

where, exactly as before, we have defined m_0 and γ by the relations

$$1+\frac{1}{p_0}=\frac{1}{m_0}+\frac{1}{p_2} \quad\text{and}\quad \gamma=\frac{d}{2p_2}-\frac{d}{2p_0}.$$

Our assumptions ensure that $\gamma\in]0,1[$ and one may apply the Hardy–Littlewood–Sobolev inequality with respect to time. Since r_0 and r_2 satisfy $1/r_0+1=1/r_2+\gamma$ by hypothesis (2-10), we immediately get that $\mathcal{A}_0$ is bounded from $L^{r_2}_T(L^{p_2})$ to $L^{r_0}_T(L^{p_0})$.

The second part of the statement is classical. Arguing by density, it suffices to establish that

$$\|\mathcal{A}_0 f(T)\|_{\dot B^{s_2}_{p_2,r_2}} \le C\|f\|_{L^{r_2}_T(L^{p_2})}.$$

To this end, we write, using Proposition 2.5 and an obvious change of variable, that

$$\begin{aligned}\|\mathcal{A}_0 f(T)\|_{\dot B^{s_2}_{p_2,r_2}} &\sim \left(\int_0^\infty \|\Delta e^{t\Delta}\mathcal{A}_0 f(T)\|^{r_2}_{L^{p_2}}\,dt\right)^{1/r_2}\\ &\sim \left(\int_0^\infty \left\|\int_0^T \Delta e^{(t+T-\tau)\Delta} f(\tau)\,d\tau\right\|^{r_2}_{L^{p_2}} dt\right)^{1/r_2}\\ &\le C\left(\int_T^\infty \left\|\int_0^{t'} \Delta e^{(t'-\tau)\Delta}(f\mathbb{1}_{[0,T]})(\tau)\,d\tau\right\|^{r_2}_{L^{p_2}} dt\right)^{1/r_2}.\end{aligned}$$

Then using Lemma 2.6 to bound the right-hand side yields the claimed inequality. □

2B2. *Maximal regularity results for the operator $\mathcal{L}$.* Here we want to extend the results of the previous subsection to the elliptic operator $\mathcal{L} = -\mu\Delta - \lambda\nabla\,\mathrm{div}$, under condition (0-2): for suitable initial datum h_0 and external force f, let us consider the equation

$$\begin{cases}\partial_t h + \mathcal{L}h = f,\\ h_{|t=0} = h_0.\end{cases} \tag{2-12}$$

The following statement will be a key ingredient in the proof of our existence result.

Proposition 2.9. *Let $((p_j, r_j))_{j=0,1,2}$ satisfy*

$$1 < p_2, r_2 < \infty, \quad r_2 < r_0, \quad r_2 < r_1, \quad p_0 \ge p_2, \quad p_1 \ge p_2,$$

and the relations (2-7) and (2-10). Let h_0 be in $\dot B^{s_2}_{p_2,r_2}$ with $s_2 := 2-2/r_2$, and let f be in $L^{r_2}_{\mathrm{loc}}(\mathbb{R}_+; L^{p_2}(\mathbb{R}^d))$. Let $(\mu,\lambda)\in\mathbb{R}^2$ satisfy condition (0-2).

Then, for all $T > 0$, system (2-12) has a unique solution h in $\mathcal{C}([0,T]; \dot B^{s_2}_{p_2,r_2}) \cap L^{r_0}([0,T]; L^{p_0})$, with $\nabla h \in L^{r_1}([0,T]; L^{p_1})$ and $\partial_t h, \nabla^2 h \in L^{r_2}([0,T]; L^{p_2})$. Moreover, there exists a constant $C_0 > 0$ (depending just on μ, λ, d, p_0, p_1, p_2 and r_2) such that the following estimate holds true:

$$\|h\|_{L^\infty_T(\dot B^{s_2}_{p_2,r_2})} + \|h\|_{L^{r_0}_T(L^{p_0})} + \|\nabla h\|_{L^{r_1}_T(L^{p_1})} + \|(\partial_t h, \nabla^2 h)\|_{L^{r_2}_T(L^{p_2})} \le C_0(\|h_0\|_{\dot B^{s_2}_{p_2,r_2}} + \|f\|_{L^{r_2}_T(L^{p_2})}). \tag{2-13}$$

Proof. Let us write down the Helmholtz decomposition of the vector field h: denoting by $\mathbb{P}$ the Leray projector onto the space of divergence-free vector fields and by $\mathbb{Q}$ the projector onto the space of irrotational vector fields, we have $h = \mathbb{P}h + \mathbb{Q}h$. Recall that we have in Fourier variables

$$\mathcal{F}(\mathbb{Q}h)(\xi) = \frac{1}{|\xi|^2}(\xi\cdot\hat h(\xi))\xi.$$

Hence $\mathbb{P}$ and $\mathbb{Q}$ are linear combinations of composition of Riesz transforms and thus act continuously on L^p for all $1 < p < \infty$.

Now, applying those two operators to system (2-12), we discover that $\mathbb{P}h$ and $\mathbb{Q}h$ satisfy the heat equations

$$(\partial_t - \mu\Delta)\mathbb{P}h = \mathbb{P}f \quad\text{and}\quad (\partial_t - \nu\Delta)\mathbb{Q}h = \mathbb{Q}f, \quad \text{with } \nu := \mu+\lambda,$$

with initial data $\mathbb{P}h_0$ and $\mathbb{Q}h_0$, respectively. Therefore, denoting by $\mathbb{A}$ the operator $\mathbb{P}$ or $\mathbb{Q}$, and by α either μ or ν, Duhamel's formula gives us

$$\mathbb{A}h(t) = e^{\alpha t\Delta}\mathbb{A}h_0 + \int_0^t e^{\alpha(t-s)\Delta}\mathbb{A}f(s)\,ds. \tag{2-14}$$

For the term containing the initial datum, we apply (2-4), (2-5), (2-8) and (2-11), while for the source term we apply Lemmas 2.6, 2.7 and 2.8. Therefore, we conclude by continuity of the operators $\mathbb{P}$ and $\mathbb{Q}$ over L^p for $1 < p < \infty$, and over $\dot{B}^s_{q,r}$ for all $s \in \mathbb{R}$ and all $(q,r) \in]1,\infty[^2$. □

2C. *Tangential regularity.* We establish here fundamental stationary estimates about propagation of *striated* (or *tangential*) *regularity* in the L^p setting.

Before starting the presentation, let us recall that the classical result on pseudodifferential operators of order zero ensures that, if $g \in L^\infty$, then $\Delta^{-1}\partial_i\partial_j g \in \mathrm{BMO}$. The main result of this subsection states that, if g has suitable tangential regularity properties (similar to those exhibited in [Chemin 1991] for the vortex patches problem), then $\Delta^{-1}\partial_i\partial_j g$ is in L^∞. Our starting point is an adaptation of Lemma 5.1 in [Paicu and Zhang 2017] enabling us to handle the operator $\nabla(\eta\mathrm{Id}-\Delta)^{-1}$, valid in any dimension and for nonzero divergence vector fields.

Before stating that lemma, the proof of which is postponed in the Appendix, let us introduce, for $m \in \mathbb{R}$, the class S^m of symbols of order m, that is, the space of $\mathcal{C}^\infty(\mathbb{R}^d)$ functions σ such that for all $\alpha \in \mathbb{N}^d$ there exists $C_\alpha > 0$ satisfying, for all $\xi \in \mathbb{R}^d$,

$$|\psi^\alpha\sigma(\xi)| \le C_\alpha(1+|\xi|)^{m-|\alpha|}.$$

Lemma 2.10. *Let $1 < p < \infty$. Consider a vector field X in $\mathbb{L}^{\infty,p}$ and a function $g \in L^\infty$ such that $\partial_X g \in L^p$. Let σ be a smooth Fourier multiplier in the class S^{-1}. Then, for any fixed $0 < s < 1$, the following estimate holds true*:

$$\|\psi_X\sigma(D)g\|_{B^s_{p,\infty}} \le C(\|\psi_X g\|_{L^p} + \|\nabla X\|_{L^p}\|g\|_{L^\infty}).$$

Since the nonhomogeneous Besov space $B^s_{p,\infty}$ is embedded in L^∞ whenever $s > d/p$, Lemma 2.10 implies the following fundamental result.

Corollary 2.11. *Assume that $d < p < \infty$, and consider a vector field X in $\mathbb{L}^{\infty,p}$ and a function $g \in L^\infty$ such that $\partial_X g \in L^p$. Then, for any Fourier multiplier σ in the class S^{-1}, there exists a constant $C > 0$ such that*

$$\|\psi_X\sigma(D)g\|_{L^\infty} \le C(\|\psi_X g\|_{L^p} + \|\nabla X\|_{L^p}\|g\|_{L^\infty}).$$

From the previous results, we immediately obtain the following fundamental stationary estimate.

Proposition 2.12. *Fix $p \in]d,\infty[$ and an integer $m \ge d-1$, and take a nondegenerate family $\mathcal{X} = (X_\lambda)_{1\le\lambda\le m}$ of vector fields belonging to $\mathbb{L}^{\infty,p}$. Let $g \in L^\infty(\mathbb{R}^d)$ be such that $g \in \mathbb{L}^p_{\mathcal{X}}$.*

Then for all $\eta > 0$, one has the property

$$(\eta\mathrm{Id}-\Delta)^{-1}\nabla^2 g \in L^\infty(\mathbb{R}^d).$$

Moreover, there exists a positive constant C such that the following estimates hold true:

$$\|\nabla^2(\eta\mathrm{Id}-\Delta)^{-1}g\|_{L^\infty}\le C\Bigg(\bigg(1+\frac{|\!|\!|\mathcal{X}|\!|\!|_{L^\infty}^{4d-5}|\!|\!|\nabla\mathcal{X}|\!|\!|_{L^p}}{(I(\mathcal{X}))^{4d-4}}\bigg)\|g\|_{L^\infty}+\frac{|\!|\!|\mathcal{X}|\!|\!|_{L^\infty}^{4d-5}}{(I(\mathcal{X}))^{4d-4}}|\!|\!|\partial_{\mathcal{X}}g|\!|\!|_{L^p}\Bigg).$$

Proof. Fix some $\Lambda:=(\lambda_1,\dots,\lambda_{d-1})\in\Lambda^m_{d-1}$, and consider the set U_Λ of those $x\in\mathbb{R}^d$ such that one has $\big(\overset{d-1}{\wedge}X_\Lambda\big)(x)\ge(I(\mathcal{X}))^{d-1}$. By Lemma 3.2 of [Danchin 1999], there exists a family of functions $b^{k\ell}_{ij}$, where $(i,j,k)\in\{1,\dots,d\}^3$ and $\ell\in\{1,\dots,d-1\}$, which are homogeneous of degree $4d-5$ with respect to the coefficients of $(X_{\lambda_1},\dots,X_{\lambda_{d-1}})$ and such that the following identity holds true on U_Λ for all $\xi\in\mathbb{R}^d$:

$$\xi_i\xi_j=\frac{\big(\overset{d-1}{\wedge}X_\Lambda(x)\big)^i\big(\overset{d-1}{\wedge}X_\Lambda(x)\big)^j}{\big|\overset{d-1}{\wedge}X_\Lambda(x)\big|^2}|\xi|^2+\frac{1}{\big|\overset{d-1}{\wedge}X_\Lambda(x)\big|^4}\sum_{k,\ell}b^{k\ell}_{ij}\xi_k(X_{\lambda_\ell}(x)\cdot\xi).$$

Then, we multiply both sides by $(\eta+|\xi|^2)^{-1}\hat g(\xi)$ and take the inverse Fourier transform at x:

$$(\eta\mathrm{Id}-\Delta)^{-1}\partial_i\partial_jg=\frac{\big(\overset{d-1}{\wedge}X_\Lambda\big)^i\big(\overset{d-1}{\wedge}X_\Lambda\big)^j}{\big|\overset{d-1}{\wedge}X_\Lambda\big|^2}\Delta(\eta\mathrm{Id}-\Delta)^{-1}g+\frac{1}{\big|\overset{d-1}{\wedge}X_\Lambda\big|^4}\sum_{k,\ell}b^{k\ell}_{ij}\partial_{X_{\lambda_\ell}}(\partial_k(\eta\mathrm{Id}-\Delta)^{-1}g).$$

Hence, thanks also to Lemma 2.3, for all $\Lambda\in\Lambda^{d-1}$, we deduce the following bound on the set U_Λ:

$$\|(\eta\mathrm{Id}-\Delta)^{-1}\partial_i\partial_jg\|_{L^\infty(U_\Lambda)}\le\|g\|_{L^\infty}+\frac{C}{(I(\mathcal{X}))^{4d-4}}\sum_\ell\|X_\Lambda\|_{L^\infty}^{4d-5}\|\partial_{X_{\lambda_\ell}}(\nabla(\eta\mathrm{Id}-\Delta)^{-1}g)\|_{L^\infty}.\tag{2-15}$$

In order to bound the last term in the right-hand side, we apply Corollary 2.11: we get, for some constant $C>0$ also depending on d and on p, the estimate

$$\|\partial_{X_{\lambda_\ell}}(\nabla(\eta\mathrm{Id}-\Delta)^{-1}g)\|_{L^\infty}\le C(\|\partial_{X_{\lambda_\ell}}g\|_{L^p}+\|g\|_{L^\infty}\|\nabla X_{\lambda_\ell}\|_{L^p}).$$

Inserting this bound into (2-15) immediately gives us the result. □

We conclude this part by presenting a new estimate concerning tangential regularity. This statement, which will be proved in the Appendix, turns out to be of tremendous importance to exhibiting the Lipschitz regularity of the velocity field u; see Section 4A below.

Proposition 2.13. *Let the hypotheses of Proposition 2.12 be in force. Then there exists a constant $C>0$ such that, for all $\eta>0$, the following estimate holds true*:

$$|\!|\!|\partial_{\mathcal{X}}\nabla^2(\eta\mathrm{Id}-\Delta)^{-1}g|\!|\!|_{L^p}\le C\Bigg(|\!|\!|\nabla\mathcal{X}|\!|\!|_{L^p}\bigg(1+\frac{|\!|\!|\mathcal{X}|\!|\!|_{L^\infty}^{4d-5}|\!|\!|\nabla\mathcal{X}|\!|\!|_{L^p}}{(I(\mathcal{X}))^{4d-4}}\bigg)\|g\|_{L^\infty}$$
$$+\bigg(1+\frac{|\!|\!|\mathcal{X}|\!|\!|_{L^\infty}^{4d-5}|\!|\!|\nabla\mathcal{X}|\!|\!|_{L^p}}{(I(\mathcal{X}))^{4d-4}}\bigg)|\!|\!|\partial_{\mathcal{X}}g|\!|\!|_{L^p}+\frac{|\!|\!|\mathcal{X}|\!|\!|_{L^\infty}^{4d-4}}{(I(\mathcal{X}))^{4d-4}}|\!|\!|\nabla\mathcal{X}|\!|\!|_{L^p}\|g\|_{L^\infty}\Bigg).$$

3. An existence statement for almost critical data and only bounded density

The goal of the present section is to prove Theorem 1.1. After reformulating the original system (0-1), we establish a priori estimates on smooth solutions, and then provide the reader with the construction of a family of approximate smooth solutions to our (new) system. As a last step, we show the convergence of the sequence to a true solution.

3A. ***Reformulation of the system.*** In all that follows, we assume for notational simplicity that $\nu = 1$. This is of course not restrictive, owing to the change of unknowns

$$(\tilde{\rho}, \tilde{u})(t, x) := (\rho, u)(\nu t, \nu x). \tag{3-1}$$

First of all, we want to reformulate our system in terms of new unknowns, to which maximal regularity results of Section 2B apply. As already explained at the beginning of the paper, in order to handle the pressure term, it is convenient to introduce the auxiliary vector field

$$v := -\nabla(\mathrm{Id} - \Delta)^{-1} P \tag{3-2}$$

and the *modified velocity field*

$$w := u - v = u + \nabla(\mathrm{Id} - \Delta)^{-1} P. \tag{3-3}$$

For future use, we observe that

$$\operatorname{div} u = \operatorname{div} w - \Delta(\mathrm{Id} - \Delta)^{-1} P. \tag{3-4}$$

We want to reformulate system (0-1) in terms of the new unknowns (ϱ, w), keeping in mind that estimates for v may be deduced from those for ϱ, and that combining with information on w enables us to bound the original velocity field u. From the first equation of (1-1) and relation (3-4), we immediately deduce that

$$\partial_t \varrho + u \cdot \nabla \varrho = -\rho \operatorname{div} w + \rho \Delta(\mathrm{Id} - \Delta)^{-1} P, \tag{3-5}$$

with $\rho := 1 + \varrho$ and $u := w - \nabla(\mathrm{Id} - \Delta)^{-1} P$.

Regarding w, we see from the second equation of (1-1) and (3-3) that

$$\rho \partial_t w - \mu \Delta w - \lambda \nabla \operatorname{div} w = -(\mathrm{Id} - \Delta)^{-1} \nabla P - \rho u \cdot \nabla u - \rho (\mathrm{Id} - \Delta)^{-1} \nabla \psi_t P.$$

Using once again the mass equation in (0-1), we find

$$\partial_t P + \operatorname{div}(P u) = g(\rho) \operatorname{div} u, \quad \text{with } g(\rho) := P(\rho) - \rho P'(\rho),$$

so that the equation for w can be recast as

$$\rho \partial_t w + \mathcal{L} w = -\rho F, \tag{3-6}$$

with

$$F = \frac{1}{\rho} (\mathrm{Id} - \Delta)^{-1} \nabla P + w \cdot \nabla w + w \cdot \nabla v + v \cdot \nabla w + v \cdot \nabla v + (\mathrm{Id} - \Delta)^{-1} \nabla (g(\rho) \operatorname{div} u - \operatorname{div}(P u)). \tag{3-7}$$

The existence part of Theorem 1.1 will be a consequence of the following statement.

Proposition 3.1. *Let $\varrho_0 \in L^p \cap L^\infty(\mathbb{R}^d)$ and $w_0 \in \dot{B}^{2-2/r}_{p,r}$, with*

$$d < p < \infty \quad \textit{and} \quad 1 < r < \frac{2p}{2p - d}. \tag{3-8}$$

Define r_0 and r_1 by

$$\frac{1}{r_0} := \frac{1}{r} + \frac{d}{2p} - 1 \quad \text{and} \quad \frac{1}{r_1} := \frac{1}{r} - \frac{1}{2}. \tag{3-9}$$

There exists some small enough constant $\varepsilon > 0$ depending only on d, p and r, such that, if in addition ϱ_0 fulfills (1-5), then there exist a time $T > 0$ and a weak solution (ϱ, w) to system (3-5)–(3-6) on $[0, T[\times \mathbb{R}^d$ such that $\varrho \in \mathcal{C}([0, T]; L^p)$ and

$$\|\varrho\|_{L^\infty_T(L^\infty)} \leq 4\varepsilon, \tag{3-10}$$

and $w \in \mathcal{C}([0, T]; \dot{B}^{2-2/r}_{p,r}) \cap L^{r_0}([0, T]; L^\infty(\mathbb{R}^d))$, with $\nabla w \in L^{r_1}([0, T]; L^p(\mathbb{R}^d))$ and $\partial_t w, \nabla^2 w \in L^r([0, T]; L^p(\mathbb{R}^d))$.

If in addition to the above hypotheses, we have $\inf P' > 0$ *on $[1-4\varepsilon, 1+4\varepsilon]$ and $\varrho_0, u_0 \in L^2(\mathbb{R}^d)$, where $u_0 := w_0 + v_0$ and v_0 is defined as in the statement of Theorem 1.1, then $u := w - \nabla(\mathrm{Id} - \Delta)^{-1} P$ fulfills $u \in L^\infty_T(L^2) \cap L^2_T(H^1)$ and the energy equality (1-6) holds true.*

3B. ***A priori bounds for smooth solutions.*** We start by establishing a priori estimates for smooth solutions to the new system (3-5)–(3-6). Our goal is to "close the estimates" in the space

$$E_T := \left\{(\varrho, w) \in L^\infty_T(L^p \cap L^\infty) \times (\mathcal{C}([0, T]; \dot{B}^{2-2/r}_{p,r}) \cap L^{r_0}_T(L^\infty)) \;\middle|\; \nabla w \in L^{r_1}_T(L^p),\ \nabla^2 w \in L^{r_2}_T(L^p)\right\}$$

for some small enough $T > 0$.

We define

$$\mathcal{N}(T) := \|\varrho\|_{L^\infty_T(L^p\cap L^\infty)} + \|w\|_{L^\infty_T(\dot{B}^{2-2/r}_{p,r})\cap L^{r_0}_T(L^\infty)} + \|\nabla w\|_{L^{r_1}_T(L^p)} + \|(\partial_t w, \nabla^2 w)\|_{L^r_T(L^p)}. \tag{3-11}$$

As it fulfills a transport equation, it is easy to propagate any Lebesgue norm L^q for ϱ once we know that $\operatorname{div} u$ is in $L^1_T(L^\infty)$ and that the right-hand side of (3-5) is in $L^1_T(L^q)$. Given the expected properties on w, this will give us the constraint $q \geq p$. As for (3-6), we want to apply the maximal regularity estimates given by Proposition 2.9.

3B1. *Bounds for the density.* Throughout, we fix some $\varepsilon > 0$ and constant $C > 0$ so that (recall that $P(1) = 0$)

$$|P(z)| \leq C|z - 1| \quad \text{for all } z \in [1 - 4\varepsilon, 1 + 4\varepsilon]. \tag{3-12}$$

As a first step, let us establish estimates for the density term. Let us take some $q \in [p, \infty[$. By multiplying (3-5) by $|\varrho|^{q-2}\varrho$ and integrating in space, we easily get

$$\frac{1}{q}\frac{d}{dt}\|\varrho\|^q_{L^q} - \frac{1}{q}\int |\varrho|^q \operatorname{div} u + \int |\varrho|^q \operatorname{div} w + \int \varrho|\varrho|^{q-2} \operatorname{div} w = \int (\varrho+1)\varrho|\varrho|^{q-2}\Delta(\mathrm{Id} - \Delta)^{-1} P.$$

By (3-4), one can rewrite the previous relation as

$$\frac{1}{q}\frac{d}{dt}\|\varrho\|^q_{L^q} + \left(1 - \frac{1}{q}\right)\int |\varrho|^q(\operatorname{div} w - \Delta(\mathrm{Id} - \Delta)^{-1} P) = \int \varrho|\varrho|^{q-2}(\Delta(\mathrm{Id} - \Delta)^{-1} P - \operatorname{div} w),$$

which immediately implies

$$\|\varrho(t)\|_{L^q} \le \|\varrho_0\|_{L^q} + \Big(1-\frac{1}{q}\Big)\int_0^t \|\varrho\|_{L^q}\|\Delta(\mathrm{Id}-\Delta)^{-1}P - \operatorname{div} w\|_{L^\infty}\,d\tau + \int_0^t \|\Delta(\mathrm{Id}-\Delta)^{-1}P - \operatorname{div} w\|_{L^q}\,d\tau.$$

Of course, passing to the limit $q \to \infty$, we see that the same inequality holds true for $q = \infty$. Then, using Lemma 2.3 and the fact that, under (3-12), we have

$$\|P\|_{L^r} \le C\|\varrho\|_{L^r} \quad \text{for all } r \in [1,\infty], \tag{3-13}$$

an application of Gronwall's lemma implies that, for all $q \in [p,\infty]$, there holds

$$\|\varrho(t)\|_{L^q} \le e^{Ct+\int_0^t \|\operatorname{div} w\|_{L^\infty}d\tau}\Big(\|\varrho_0\|_{L^q} + \int_0^t \|\operatorname{div} w\|_{L^q}\,d\tau\Big). \tag{3-14}$$

Define now the time $T > 0$ as

$$T := \sup\Big\{t > 0 \;\Big|\; Ct + \int_0^t \|\operatorname{div} w\|_{L^\infty}\,d\tau \le \log 2 \text{ and } \int_0^t \|\operatorname{div} w\|_{L^\infty} \le \varepsilon\Big\}; \tag{3-15}$$

then, on $[0, T]$ one has, owing to (3-14),

$$\|\varrho(t)\|_{L^\infty} \le 2\|\varrho_0\|_{L^\infty} + 2\varepsilon.$$

Hence, if we take the initial density satisfying (1-5), for ε fixed in (3-12) above, then we get (3-10).

3B2. *Bounds for w.* Throughout we fix the time $T > 0$ as defined in (3-15) and assume that (3-10) is fulfilled. Then, applying Proposition 2.9 to (3-6) with $(p_0, p_1, p_2) = (\infty, p, p)$, $r_2 = r$ and (r_0, r_1) according to (3-9), we get, treating the term $\varrho\partial_t w$ as a perturbation,

$$\|w\|_{L^\infty_T(\dot B^{2-2/r}_{p,r})\cap L^{r_0}_T(L^\infty)} + \|\nabla w\|_{L^{r_1}_T(L^p)} + \|(\partial_t w, \nabla^2 w)\|_{L^r_T(L^p)} \le C_0(\|w_0\|_{\dot B^{2-2/r}_{p,r}} + \|\varrho\|_{L^\infty_T(L^\infty)}\|\partial_t w\|_{L^r_T(L^p)} + \|F\|_{L^r_T(L^p)})$$

for a suitable constant $C_0 > 0$. Now, assuming that ε in (1-5) has been fixed so small that

$$8C_0\varepsilon \le 1, \tag{3-16}$$

thanks to (3-10) we gather the estimate

$$\|w\|_{L^\infty_T(\dot B^{2-2/r}_{p,r})\cap L^{r_0}_T(L^\infty)} + \|\nabla w\|_{L^{r_1}_T(L^p)} + \|(\partial_t w, \nabla^2 w)\|_{L^{r_2}_T(L^p)} \le 2C_0(\|w_0\|_{\dot B^{2-2/r}_{p,r}} + \|F\|_{L^r_T(L^p)}). \tag{3-17}$$

Our next goal is to bound F, defined by (3-7). First of all, as the operator $\nabla(\mathrm{Id}-\Delta)^{-1}$ maps continuously L^q into $W^{1,q}$ for any $1 < q < \infty$, one deduces that

$$\|v\|_{W^{1,q}} \le C\|\varrho\|_{L^q}, \tag{3-18}$$

where v is the vector field defined in (3-2). Hence, the first term in (3-7) can be bounded, thanks to (3-13), in the following way:

$$\|(\mathrm{Id}-\Delta)^{-1}\nabla P\|_{L^r_T(L^p)} \le CT^{1/r}\|\varrho\|_{L^\infty_T(L^p)}. \tag{3-19}$$

Next, we estimate the transport terms of F by means of Hölder inequality, using that

$$\frac{1}{r}-\frac{1}{r_0}-\frac{1}{r_1}=\frac{3}{2}-\frac{1}{r}-\frac{d}{2p}=\frac{1}{2}-\frac{1}{r_0}>0.$$

We get the inequality

$$\|w\cdot\nabla w+w\cdot\nabla v+v\cdot\nabla w+v\cdot\nabla v\|_{L^r_T(L^p)} \le CT^{1/2-1/r_0}(\|v\|_{L^{r_0}_T(L^\infty)}+\|w\|_{L^{r_0}_T(L^\infty)})(\|\nabla v\|_{L^{r_1}_T(L^p)}+\|\nabla w\|_{L^{r_1}_T(L^p)}). \tag{3-20}$$

Hence, using the definition of $\mathcal{N}(T)$ and (3-18), inequality (3-20) becomes

$$\|w\cdot\nabla w+w\cdot\nabla v+v\cdot\nabla w+v\cdot\nabla v\|_{L^r_T(L^p)}\le CT^{1/2-1/r_0}(1+T^{1/r_0})(1+T^{1/r_1})\mathcal{N}^2(T). \tag{3-21}$$

Let us now consider the term $\nabla(\mathrm{Id}-\Delta)^{-1}\operatorname{div}(Pu)$ occurring in the definition of F. From Corollary 2.4 and (3-10), combined with the continuity of the function P, we deduce that

$$\begin{aligned}\|\nabla(\mathrm{Id}-\Delta)^{-1}\operatorname{div}(Pu)\|_{L^r_T(L^p)}&\le C\|Pu\|_{L^r_T(L^p)}\\&\le CT^{1/r-1/r_0-1/r_1}\|P\|_{L^{r_1}_T(L^p)}(\|v\|_{L^{r_0}_T(L^\infty)}+\|w\|_{L^{r_0}_T(L^\infty)})\\&\le CT^{1/r-1/r_0}\|\varrho\|_{L^\infty_T(L^p)}(T^{1/r_0}\|\varrho\|_{L^\infty_T(L^\infty)}+\|w\|_{L^{r_0}_T(L^\infty)}).\end{aligned}$$

Hence, using the definition of $\mathcal{N}(T)$, we get

$$\|\nabla(\mathrm{Id}-\Delta)^{-1}\operatorname{div}(Pu)\|_{L^r_T(L^p)}\le CT^{1/r-1/r_0}(1+T^{1/r_0})\mathcal{N}^2(T). \tag{3-22}$$

To handle the last term of F, we write

$$\nabla(\mathrm{Id}-\Delta)^{-1}(g(\rho)\operatorname{div}u)=g(1)\nabla(\mathrm{Id}-\Delta)^{-1}\operatorname{div}u+\nabla(\mathrm{Id}-\Delta)^{-1}((g(\rho)-g(1))\operatorname{div}u).$$

To bound the first term, we use that, thanks to (3-4),

$$\nabla(\mathrm{Id}-\Delta)^{-1}\operatorname{div}u=\nabla(\mathrm{Id}-\Delta)^{-1}\operatorname{div}w+\nabla(\mathrm{Id}-\Delta)^{-2}\Delta P.$$

Because both $\nabla(\mathrm{Id}-\Delta)^{-1}$ and $\nabla(\mathrm{Id}-\Delta)^{-2}\Delta$ map L^p to itself, we get

$$\|\nabla(\mathrm{Id}-\Delta^{-1})\operatorname{div}u\|_{L^r_T(L^p)}\le C(T^{1/2}\|\nabla w\|_{L^{r_1}_T(L^p)}+T^{1/r}\|P\|_{L^\infty_T(L^p)})\le C(T^{1/2}+T^{1/r})\mathcal{N}(T).$$

Similarly, we have

$$\begin{aligned}\|\nabla(\mathrm{Id}-\Delta)^{-1}((g(\rho)-g(1))\operatorname{div}u)\|_{L^r_T(L^p)}&\le C\|(g(\rho)-g(1))\operatorname{div}u\|_{L^r_T(L^p)}\\&\le CT^{1/r-1/r_0-1/r_1}\|\varrho\|_{L^{r_0}_T(L^\infty)}\|\operatorname{div}u\|_{L^{r_1}_T(L^p)};\end{aligned}$$

so, keeping in mind (3-4), one can conclude that

$$\big\|\nabla(\mathrm{Id}-\Delta)^{-1}\big((g(\rho)-g(1))\operatorname{div}u\big)\big\|_{L^r_T(L^p)}\le CT^{1/2}(1+T^{1/r_1})\mathcal{N}^2(T). \tag{3-23}$$

In the end, plugging the inequalities (3-14) and (3-19)–(3-23) into (3-17), whenever relation (3-10) is fulfilled, we get

$$\mathcal{N}(T)\le C\big(\|\varrho_0\|_{L^p\cap L^\infty}+\|w_0\|_{\dot B^{2-2/r}_{p,r}}+(T^{1/2}+T^{1/r})\mathcal{N}(T)+(T^{1/2-1/r_0}+T^{1/r})\mathcal{N}^2(T)\big)$$

for some constant C depending only on the pressure function and on the regularity parameters. From that inequality and a standard bootstrap argument, one may conclude that there exists a time $T > 0$, depending only on the norm of the initial data, such that

$$\mathcal{N}(T) \le 2C(\|\varrho_0\|_{L^p\cap L^\infty} + \|w_0\|_{\dot{B}^{2-2/r}_{p,r}}). \tag{3-24}$$

3B3. *Classical energy estimates.* We establish here energy estimates for solutions to system (0-1), under the additional assumption that $u_0 \in L^2$. The computations being quite standard, we only sketch the arguments, and refer the reader to, e.g., Chapter 5 of [Lions 1998] for the details: we take the L^2 scalar product of the momentum equation in (0-1) with u, integrate by parts and make use of the mass equation. Defining Π as in the statement of Theorem 1.1, we end up with the relation

$$\frac{1}{2}\frac{d}{dt}\int \rho|u|^2\,dx + \frac{d}{dt}\int \Pi(\rho)\,dx + \mu\int |\nabla u|^2\,dx + \lambda\int |\operatorname{div} u|^2\,dx = 0.$$

The previous relation, after integration in time, leads to the classical energy balance (1-6).

Now, keeping in mind the smallness assumption (1-5), we gather that if $\inf P' > 0$ on $[1-4\varepsilon, 1+4\varepsilon]$ then it holds on $[0, T]$ that

$$C^{-1}\|\varrho(t)\|^2_{L^2} \le \|\Pi(\rho(t))\|_{L^1} \le C\|\varrho(t)\|^2_{L^2}.$$

Hence, by the hypotheses on the initial data, we get that the right-hand side of (1-6) is finite, and then, for all $t \in [0, T]$, one has that $\sqrt{\rho}u$ belongs to $L^\infty_t(L^2)$, $\varrho \in L^\infty_t(L^2)$ and $\nabla u \in L^2_t(L^2)$.

To make a long story short, one can eventually assert that for all $t \in [0, T]$, we have

$$\|u(t)\|^2_{L^2} + \int_0^t \|\nabla u(\tau)\|^2_{L^2}\,d\tau + \|\varrho(t)\|^2_{L^2} \le C \tag{3-25}$$

for some $C > 0$ just depending on the initial energy, on P and on ε.

3C. ***The proof of existence.*** In this subsection, we derive, from the estimates of the previous part, the existence of a weak solution to system (3-5)–(3-6).

We start by smoothing out the initial data (ϱ_0, u_0) by convolution with a family of nonnegative mollifiers:

$$\varrho_0^n := \chi^n * \varrho_0 \quad \text{and} \quad u_0^n := \chi^n * u_0.$$

Then ϱ_0^n still satisfies (1-5), and both ϱ_0^n and u_0^n belong to all Sobolev spaces $W^{k,p}$, with $k \in \mathbb{N}$. Note that one can in addition multiply the regularized data by a family of cut-off functions, to have ϱ_0^n and u_0^n in L^2, which enables us to apply Theorem A of [Mucha 2001].[4] We get a sequence of solutions (ϱ^n, u^n) on $[0, T^n]$ (with $T^n > 0$) to (0-1), supplemented with initial data $(1+\varrho_0^n, u_0^n)$, fulfilling (1-5), the energy balance (1-6), and the properties

$$\varrho^n \in \mathcal{C}([0, T^n]; W^{1,p}), \quad u^n \in \mathcal{C}([0, T^n]; L^2), \quad \partial_t w^n, \nabla^2 w^n \in L^r([0, T^n]; L^p).$$

[4] Actually, [Mucha 2001] concentrates on the $\mathbb{R}^3$ case but very small changes allow one to get a similar result in $\mathbb{R}^d$ provided $p > d$. One can alternatively use [Danchin 2001], which proves local-in-time existence in Besov spaces with subcritical regularity.

Furthermore, by taking advantage of (3-24), one can exhibit some $T>0$, depending only on norms of (ϱ_0,u_0), such that $T^n\geq T$ for all $n\in\mathbb{N}$, and eventually get, for some constant C depending only on p, r, d and ε, the bound

$$\|\varrho^n\|_{L^\infty_T(L^p\cap L^\infty)}+\|w^n\|_{L^\infty_T(\dot B^{2-2/r}_{p,r})\cap L^{r_0}_T(L^\infty)}+\|\nabla w^n\|_{L^{r_1}_T(L^p)}+\|(\partial_t w^n,\nabla^2 w^n)\|_{L^r_T(L^p)}\\ \leq C(\|\varrho_0\|_{L^p\cap L^\infty}+\|w_0\|_{\dot B^{2-2/r}_{p,r}}).$$

The previous inequality ensures the weak-$\star$ convergence (up to an extraction of a subsequence) of (ϱ^n,w^n) to some (ϱ,w) in the space E_T.

Strong convergence properties are still needed, in order to pass to the limit in the weak formulation of the equations, and show that (ϱ,w) is indeed a solution of system (3-5)–(3-6). In order to glean strong compactness, it suffices to use the fact that the above uniform bound also provides a control on the first-order time derivatives in sufficiently negative Sobolev spaces, through the equation fulfilled by (ϱ^n,w^n). Then one can combine with the Ascoli theorem and interpolation, to get strong convergence, which turns out to be enough to pass to the limit in the equation satisfied by w. In order to justify that ϱ satisfies the mass equation, one can repeat the arguments of [Huang et al. 2013] (see also [Lions 1998]) which in particular imply that $\varrho^n\to\varrho$ (up to subsequence) in $\mathcal{C}([0,T];L^q)$ for all $p\leq q<\infty$. The details are left to the reader.

Finally, once it is known that (ϱ,w) satisfy the desired equation, one can recover time continuity for w by taking advantage of Proposition 2.9.

To prove that the energy balance is fulfilled in the case where $\inf P'>0$ on $[1-4\varepsilon,1+4\varepsilon]$ and $\varrho_0,u_0\in L^2$, we just have to observe that it is satisfied by (ρ^n,u^n) (with the regularized data) for all $n\in\mathbb{N}$ and that having $\varrho_0\in L^2$ guarantees that $\varrho^n\to\varrho$ in $\mathcal{C}([0,T];L^2)$. This implies that $v^n\to v$ in $\mathcal{C}([0,T];H^1)$. Furthermore, the compactness properties of (w^n) that have been pointed out just above ensure that $w^n\to w$ in $L^2([0,T];H^1_{\mathrm{loc}})\cap L^\infty([0,T];L^2_{\mathrm{loc}})$. Finally, since

$$\tfrac12\|\sqrt{\rho(t)}u(t)\|^2_{L^2}+\|\Pi(\rho(t))\|_{L^1}+\mu\|\nabla u\|^2_{L^2_t(L^2)}+\lambda\|\operatorname{div}u\|^2_{L^2_t(L^2)}\\ =\lim_{R\to\infty}\biggl(\frac12\int_{B(0,R)}(\rho(t,x)|u(t,x)|^2+\Pi(\rho(t,x)))\,dx+\int_0^t\int_{B(0,R)}(\mu|\nabla u|^2+\lambda(\operatorname{div}u)^2)\,dx\,d\tau\biggr),$$

and because the aforementioned properties of convergence enable us to pass to the limit in the right-hand side for any $R>0$, we get the desired energy balance.

Proposition 3.1 is thus completely proven, and so is Theorem 1.1 (apart from uniqueness if $d=1$, which will be discussed at the beginning of the next section).

4. Tangential regularity and uniqueness

The main goal of this section is to prove uniqueness of solutions to (0-1) in the previous functional framework. According to the pioneering work [Hoff 2006] or to the recent paper [Danchin 2014] by the first author, the condition $\nabla u\in L^1_T(L^\infty)$ seems to be the minimal requirement in order to get uniqueness. Recall that (still assuming $\nu=1$ for notational simplicity)

$$\nabla u=\nabla w+\nabla^2(\mathrm{Id}-\Delta)^{-1}P(\rho). \tag{4-1}$$

By Proposition 3.1 and Sobolev embeddings, we immediately get that ∇w is in $L^1_T(L^\infty)$. So is the last term of (4-1) if the space dimension is 1 since $P(\rho)$ is bounded and $\partial^2_{xx}(\mathrm{Id}-\partial^2_{xx})^{-1}$ maps L^∞ to L^∞. If $d\geq 2$, however, then the property that P is bounded ensures only (by Calderón–Zygmund theory) that $\nabla^2(\mathrm{Id}-\Delta)^{-1}P$ is in BMO. Having Proposition 2.12 in mind, this prompts us to make an additional tangential regularity type assumption so as to guarantee that, indeed, $\nabla^2(\mathrm{Id}-\Delta)^{-1}P$ belongs to L^∞.

In the rest of this section, we thus assume the following tangential regularity hypothesis: for $p\in\,]d,\infty[$, there exists a nondegenerate family $\mathcal{X}_0=(X_{0,\lambda})_{1\leq\lambda\leq m}$ of vector fields in $\mathbb{L}^{\infty,p}$ such that the initial density ρ_0 belongs to the space $\mathbb{L}^p_{\mathcal{X}_0}$ (see Definition 1.2 above).

4A. ***Propagation of tangential regularity.*** In this subsection we establish a priori estimates for striated regularity of the density of a smooth enough solution (ϱ,u) to system (3-5)–(3-6). From those bounds, we will infer a control on the Lipschitz norm of u. Throughout this section, we shall use the notation

$$U(t):=\int_0^t\|\nabla u(\tau)\|_{L^\infty}\,d\tau.$$

4A1. *Bounds for the tangential vector fields.* Let us generically denote by X_0 one of the vector fields of the family $\mathcal{X}_0$. It is well known that the evolution of X_0 along the velocity flow is the solution to the transport equation

$$\begin{cases}(\partial_t+u\cdot\nabla)X=\partial_X u,\\ X_{|t=0}=X_0,\end{cases}\tag{4-2}$$

where the notation $\partial_X u$ was introduced in (1-7).

For all $t\geq 0$, we then define the family $\mathcal{X}(t):=(X_\lambda(t))_{1\leq\lambda\leq m}$, where X_λ stands for the solution to system (4-2), supplemented with initial datum $X_{0,\lambda}$.

Our goal, now, is to establish some bounds on each vector field $X(t)$ of the family $\mathcal{X}(t)$. They are based on classical estimates for transport equations in the spirit of those of Section 3B1.

First of all, the standard L^∞ estimate for (4-2) leads us to the inequality

$$\|X(t)\|_{L^\infty}\leq\|X_0\|_{L^\infty}e^{U(t)}.\tag{4-3}$$

Next, arguing exactly as in Proposition 4.1 of [Danchin 1999] (one needs to pass through the flow associated to u) yields

$$I(\mathcal{X}(t))\geq I(\mathcal{X}_0)e^{-U(t)},\tag{4-4}$$

which ensures that the family $\mathcal{X}(t)$ remains nondegenerate whenever $U(t)$ stays bounded.

Finally, differentiating (4-2) with respect to the space variable x_j, we get

$$\partial_t\partial_j X+u\cdot\nabla\partial_j X=-\partial_j u\cdot\nabla X+\partial_j\partial_X u,$$

which leads to

$$\|\nabla X(t)\|_{L^p}\leq\|\nabla X(0)\|_{L^p}+\int_0^t\big(C\|\nabla u\|_{L^\infty}\|\nabla X\|_{L^p}+\|\nabla\partial_X u\|_{L^p}\big)\,d\tau.\tag{4-5}$$

Observe that, for all $1 \le j \le d$, we have the relation

$$\partial_j \partial_X u = \partial_j X \cdot \nabla u + \partial_X \partial_j u,$$

where the former term is easily bounded in L^p by the quantity $\|\nabla X\|_{L^p} \|\nabla u\|_{L^\infty}$. Then, taking advantage of Gronwall's inequality, we get

$$\|\nabla X(t)\|_{L^p} \le e^{CU(t)} \left(\|\nabla X_0\|_{L^p} + \int_0^t e^{-CU(\tau)} \|\partial_X \nabla u\|_{L^p} \, d\tau \right). \tag{4-6}$$

4A2. *Propagation of striated regularity for the density.* We now show propagation of tangential regularity for the density function. To begin, we recast the first equation of (0-1) in the form

$$\partial_t \rho + u \cdot \nabla \rho = -\rho \operatorname{div} u. \tag{4-7}$$

Next, we multiply the previous relation by X: by virtue of (4-2), we find

$$\partial_t (\rho X) + u \cdot \nabla (\rho X) + \rho X \operatorname{div} u - \rho \partial_X u = 0.$$

Taking the divergence of the obtained relation, straightforward computations lead to the equation

$$\partial_t \operatorname{div}(\rho X) + u \cdot \nabla \operatorname{div}(\rho X) = -\operatorname{div} u \operatorname{div}(\rho X).$$

From it, repeating the computations of Section 3B1, we deduce that

$$\|\operatorname{div}(\rho(t) X(t))\|_{L^p} \le e^{CU(t)} \|\operatorname{div}(\rho_0 X_0)\|_{L^p}. \tag{4-8}$$

Thanks to the previous estimate and to the relation

$$\operatorname{div}(P(\rho) X) = (P(\rho) - \rho P'(\rho)) \operatorname{div} X + P'(\rho) \operatorname{div}(\rho X),$$

one easily gathers the propagation of tangential regularity also for the pressure term:

$$\begin{aligned} \|\operatorname{div}(P(\rho) X)\|_{L^p} &\le \|P(\rho) - \rho P'(\rho)\|_{L^\infty} \|\operatorname{div} X\|_{L^p} + \|P'(\rho)\|_{L^\infty} \|\operatorname{div}(\rho X)\|_{L^p} \\ &\le C(\|\nabla X\|_{L^p} + \|\operatorname{div}(\rho X)\|_{L^p}), \end{aligned}$$

where, in writing the last inequality, we have used (3-12) and (3-10). For future use, we also notice that, by the previous estimate and (1-7), we have

$$\|\partial_X P(\rho)\|_{L^p} \le C(\|\nabla X\|_{L^p} + \|\operatorname{div}(\rho X)\|_{L^p}). \tag{4-9}$$

4A3. *Final estimates for the gradient of the velocity.* In this subsection, we complete the proof of propagation of striated regularity, and exhibit a bound for ∇u in $L^1_{T_0}(L^\infty)$, for some time $T_0 > 0$ depending only on suitable norms of the data.

First, we want to control the L^p norm of ∇X which, in light of inequality (4-6), requires bounding the quantity $\|\partial_X \nabla u\|_{L^p}$. Here lies the main difficulty, compared to the standard result of propagation of striated regularity for incompressible flows. On the one hand, the part of this term corresponding to $\mathbb{P}u$ may be bounded quite easily since $\mathbb{P}u = \mathbb{P}w$ (note that $u - w$ is a gradient) and estimates on the second derivatives are thus available through the maximal regularity results that have been proved before. On the other hand, $\mathbb{Q}u$ has a part involving w (which is fine, exactly as before) and another one depending on $P(\rho)$. Here Proposition 2.13 will come into play. More precisely, we use relation (4-1) to write

$$\partial_X \nabla u = \partial_X \nabla w + \partial_X \nabla^2 (\mathrm{Id} - \Delta)^{-1} P(\rho). \tag{4-10}$$

Bounding the first term is easy: thanks to Section 3B2 and (4-3), we have

$$\|\partial_X \nabla w\|_{L^p} \le \|X\|_{L^\infty} \|\nabla^2 w\|_{L^p} \le e^U \|X_0\|_{L^\infty} \|\nabla^2 w\|_{L^p}. \tag{4-11}$$

Recall that, thanks to Theorem 1.1 and especially estimate (3-24), the quantity $\|\nabla^2 w\|_{L^p}$ is in L^r_T, and thus in L^1_T.

For the last term in (4-10), Proposition 2.13, guarantees that

$$\begin{aligned}\|\partial_X \nabla^2 (\mathrm{Id}-\Delta)^{-1} P(\rho)\|_{L^p} \le C\Bigg(\|\!|\nabla \mathcal{X}|\!\|_{L^p} \bigg(1 + \frac{\|\!|\mathcal{X}|\!\|_{L^\infty}^{4d-5} \|\!|\nabla \mathcal{X}|\!\|_{L^p}}{(I(\mathcal{X}))^{4d-4}}\bigg) \|P(\rho)\|_{L^\infty} \\ + \bigg(1 + \frac{\|\!|\mathcal{X}|\!\|_{L^\infty}^{4d-5} \|\!|\nabla \mathcal{X}|\!\|_{L^p}}{(I(\mathcal{X}))^{4d-4}}\bigg) \|\!|\partial_{\mathcal{X}} P(\rho)|\!\|_{L^p} \\ + \frac{\|\!|\mathcal{X}|\!\|_{L^\infty}^{4d-4}}{(I(\mathcal{X}))^{4d-4}} \|\!|\nabla \mathcal{X}|\!\|_{L^p} \|P(\rho)\|_{L^\infty}\Bigg).\end{aligned}$$

In view of estimates (3-24) and (4-9), this implies that (taking C larger if need be)

$$\begin{aligned}&\|\partial_X \nabla^2 (\mathrm{Id}-\Delta)^{-1} P(\rho)\|_{L^p} \\ &\qquad \le C\Bigg((\|\!|\nabla \mathcal{X}|\!\|_{L^p} + \|\!|\mathrm{div}(\rho\mathcal{X})|\!\|_{L^p}) \bigg(1 + \frac{\|\!|\mathcal{X}|\!\|_{L^\infty}^{4d-5} \|\!|\nabla \mathcal{X}|\!\|_{L^p}}{(I(\mathcal{X}))^{4d-4}}\bigg) + \frac{\|\!|\mathcal{X}|\!\|_{L^\infty}^{4d-4}}{(I(\mathcal{X}))^{4d-4}} \|\!|\nabla \mathcal{X}|\!\|_{L^p}\Bigg).\end{aligned}$$

At this point, we use the bounds (4-3), (4-4) and (4-8). Including a dependence on $\|\!|\mathcal{X}_0|\!\|_{L^\infty}$, $I(\mathcal{X}_0)$ and $\|\!|\mathrm{div}(\rho_0\mathcal{X}_0)|\!\|_{L^p}$ in C, we deduce for some constant c_d depending only on d

$$\begin{aligned}&\|\partial_X \nabla^2 (\mathrm{Id}-\Delta)^{-1} P(\rho)\|_{L^p} \\ &\qquad \le C\big(\|\!|\nabla \mathcal{X}|\!\|_{L^p} (1 + e^{c_d U} \|\!|\nabla \mathcal{X}|\!\|_{L^p}) + (1 + e^{c_d U} \|\!|\nabla \mathcal{X}|\!\|_{L^p}) e^U + e^{(c_d+1)U} \|\!|\nabla \mathcal{X}|\!\|_{L^p}\big).\end{aligned}$$

Changing c_d to $c_d + 1$, the previous relation finally leads us to the bound

$$\|\partial_X \nabla^2 (\mathrm{Id}-\Delta)^{-1} P(\rho)\|_{L^p} \le C e^{c_d U} (1 + \|\!|\nabla \mathcal{X}|\!\|_{L^p}^2). \tag{4-12}$$

Putting estimates (4-11) and (4-12) together, we finally gather

$$\|\partial_X \nabla u\|_{L^p} \le C e^{c_d U} (1 + \|\!|\nabla \mathcal{X}|\!\|_{L^p}^2 + \|\nabla^2 w\|_{L^p}). \tag{4-13}$$

At this point, define the time T_1 to satisfy

$$T_1 := \sup\{t \in \,]0, T] \mid U(t) \le \log 2\}, \tag{4-14}$$

where $T > 0$ is the time given by Proposition 3.1.

Then, changing C if need be, estimate (4-13) implies that, on $[0, T_1]$, one has

$$\|\partial_X \nabla u\|_{L^p} \le C(1 + \|\!|\nabla \mathcal{X}|\!\|_{L^p}^2 + \|\nabla^2 w\|_{L^p}).$$

Inserting now this bound in (4-6) and using also (3-24), we find that, for all $\lambda \in \{1, \dots, m\}$,

$$\begin{aligned}\|\nabla X_\lambda(t)\|_{L^p} &\le 2\biggl(\|\nabla X_{0,\lambda}\|_{L^p} + \int_0^t \|\partial_{X_\lambda}\nabla u\|_{L^p}\,d\tau\biggr)\\ &\le 2\biggl(\|\nabla X_{0,\lambda}\|_{L^p} + C\int_0^t (1+|||\nabla\mathcal{X}|||_{L^p}^2 + \|\nabla^2 w\|_{L^p})\,d\tau\biggr).\end{aligned}$$

Taking the supremum on λ, we find that for some C_0 depending only on the norms of the data, and for some "absolute" constant C, we have for all $t \in [0, T_1]$

$$|||\nabla\mathcal{X}(t)|||_{L^p} \le C\biggl(C_0 + C\int_0^t |||\nabla\mathcal{X}(\tau)|||_{L^p}^2\,d\tau\biggr).$$

Then, using a Gronwall-type argument, we conclude that

$$|||\nabla\mathcal{X}(t)|||_{L^p} \le \frac{C_0}{1-CC_0 t} \quad \text{for all } t \in [0, T_1] \text{ satisfying } CC_0 t < 1. \tag{4-15}$$

This having been established, let us turn our attention to finding a bound for the quantity $U(t)$, as it is needed to close the estimates. Resorting to relation (4-1) again, we see that we have to control the $L_t^1(L^\infty)$ norm of each term appearing on its right-hand side. For the term in w, this is an easy task: thanks to Proposition 3.1 and Sobolev embeddings, a decomposition into low and high frequencies implies

$$\|\nabla w\|_{L_t^1(L^\infty)} \le C(t^{1-1/r_1}\|\nabla w\|_{L_T^{r_1}(L^p)} + t^{1-1/r}\|\nabla^2 w\|_{L_T^r(L^p)}) \tag{4-16}$$

for all $t \in [0, T]$, where we also used relation (1-4).

Bounding the latter term in (4-1) is based on Proposition 2.12, which gives

$$\|\nabla^2(\mathrm{Id}-\Delta)^{-1}P(\rho)\|_{L^\infty} \le C\biggl(\biggl(1+\frac{|||\mathcal{X}|||_{L^\infty}^{4d-5}|||\nabla\mathcal{X}|||_{L^p}}{(I(\mathcal{X}))^{4d-4}}\biggr)\|P(\rho)\|_{L^\infty} + \frac{|||\mathcal{X}|||_{L^\infty}^{4d-5}}{(I(\mathcal{X}))^{4d-4}}|||\partial_{\mathcal{X}}P(\rho)|||_{L^p}\biggr).$$

In view of (3-24), (4-3), (4-4), (4-8) and (4-9), omitting once again the explicit dependence of the multiplicative constants on the norms of the initial data, the previous inequality allows us to get

$$\begin{aligned}\|\nabla^2(\mathrm{Id}-\Delta)^{-1}P(\rho)\|_{L^\infty} &\le C(1+e^{c_d U}|||\nabla\mathcal{X}|||_{L^p} + e^{c_d U}(|||\nabla\mathcal{X}|||_{L^p}+e^U))\\ &\le Ce^{c_d U}(1+|||\nabla\mathcal{X}|||_{L^p}).\end{aligned} \tag{4-17}$$

Recalling definition (4-14) of T_1 and (4-15) and taking $0 < T_0 \le T_1$ so that $2CC_0T_0 \le 1$, we gather

$$\|\nabla^2(\mathrm{Id}-\Delta)^{-1}P(\rho)\|_{L_{T_0}^\infty(L^\infty)} \le C(1+C_0), \tag{4-18}$$

which implies, together with (4-16), the following control, for all fixed $t \in [0, T_0]$:

$$\|\nabla u\|_{L_t^1(L^\infty)} \le C(t^{1-1/r_1}\|\nabla w\|_{L_T^{r_1}(L^{p_1})} + t^{1-1/r}\|\nabla^2 w\|_{L_T^r(L^p)} + t(1+C_0)). \tag{4-19}$$

Up to taking a smaller T_0, we then see that the requirement $\|\nabla u\|_{L_{T_0}^1(L^\infty)} \le \log 2$ can be fulfilled. Then, a classical bootstrap argument, which we do not detail here, finally allows us to deduce the boundedness of ∇u in $L_{T_0}^1(L^\infty)$.

In order to prove rigorously the existence part of Theorem 1.3, one may proceed as in Section 3C. There, we constructed a sequence (ϱ^n, u^n) of smooth solutions that is uniformly bounded in the space E_T. Therefore, it is only a matter of checking that one can get uniform bounds, too, for the striated regularity. To do this, we smooth out the reference family of vector fields $\mathcal{X}_0$ into $\mathcal{X}_0^n$ (paying attention to keep the nondegeneracy condition), and then define the family $\mathcal{X}^n := (X_\lambda^n)_{1\le\lambda\le m}$ transported by the flow of u^n according to (4-2), taking u^n instead of u and starting from the initial vector field $X_{0,\lambda}^n$. Then, repeating the computations that have been carried out just above, we get uniform bounds for all the quantities involving the striated regularity, and thus also for ∇u^n in $L^1_{T_0}(L^\infty)$. That (ϱ^n, u^n) tends to some solution (ϱ, u) of (1-1) belonging to E_{T_0} has already been justified before. Furthermore, combining our new bounds with compactness arguments allows us to pass to the limit in (4-2) as well, and to get the crucial information that ∇u is in $L^1_{T_0}(L^\infty)$.

4B. ***The proof of uniqueness.*** With (4-19) established, one can now tackle the proof of uniqueness of solutions. The basic idea is to perform a Lagrangian change of coordinates in system (0-1), in order to bypass the hyperbolic nature of the mass equation, which otherwise would cause the loss of one derivative in the stability estimates. In fact, we will perform stability estimates for the Lagrangian formulation of (0-1).

4B1. *Lagrangian formulation.* The goal of this subsection is to recast system (0-1) in Lagrangian variables. Recall that in light of the estimates of Section 4A, we know that for all $k_0 > 0$ there exists a time $T_0 > 0$ such that

$$\int_0^{T_0} \|\nabla u(t)\|_{L^\infty}\, dt \le k_0. \tag{4-20}$$

The value of k_0 will be determined in the course of the computations below.

First of all, we define the flow ψ_u associated to the velocity field u to be the solution of

$$\psi_u(t, y) := y + \int_0^t u(\tau, \psi_u(\tau, y))\, d\tau. \tag{4-21}$$

Thanks to that, any function $f = f(t, x)$ may be rewritten in Lagrangian coordinates (t, y) according to the relation

$$\bar{f}(t, y) := f(t, \psi_u(t, y)). \tag{4-22}$$

A key observation is that, once passing to Lagrangian coordinates, one can forget about the reference *Eulerian* velocity u by rewriting definition (4-21) in terms of the *Lagrangian* velocity $\bar{u}$, defining directly ψ_u by

$$\psi_u(t, y) = y + \int_0^t \bar{u}(\tau, y)\, d\tau.$$

In what follows, we set $J_u := \det(D\psi_u)$ and $A_u := (D\psi_u)^{-1}$. Observe that, by the standard chain rule,

$$\overline{D_x f} = D_y \bar{f} \cdot A_u. \tag{4-23}$$

Lemma A.2 of [Danchin 2014] provides us with the following alternative expressions.[5]

[5]From now on we agree that $\mathrm{adj}(M)$ designates the adjugate matrix of M. Of course, if M is invertible, then $\mathrm{adj}(M) = (\det M)M^{-1}$.

Lemma 4.1. *For any $\mathcal{C}^1$ function K and any $\mathcal{C}^1$ vector field H defined over $\mathbb{R}^d$, one has*

$$\overline{\nabla_x K} = J_u^{-1}\operatorname{div}_y(\operatorname{adj} D\psi_u \bar{K}),$$
$$\overline{\operatorname{div}_x H} = J_u^{-1}\operatorname{div}_y(\operatorname{adj} D\psi_u \bar{H}).$$

Moreover, since our diffeomorphism ψ_u is the flow of the time-dependent vector field u, we also get, for any function f,

$$\overline{\partial_t f + \operatorname{div}(fu)} = J_u^{-1}\partial_t(J_u\bar{f}). \tag{4-24}$$

The next statement is in the spirit of Lemma A.3 of [Danchin 2014]; also its proof follows the same steps, up to a straightforward adaptation to our functional framework.

Lemma 4.2. *Let u be a velocity field with $\nabla u \in L^1([0,T_0];L^\infty(\mathbb{R}^d))$, and let ψ_u be its flow, defined by (4-21). Suppose that condition (4-20) is fulfilled with $k_0 < 1$.*

Then there exists a constant $C > 0$, just depending on k_0, such that the following estimates hold true, for all times $t \in [0,T_0]$:

$$\|\mathrm{Id} - \operatorname{adj} D\psi_u(t)\|_{L^\infty} \le C\|Du\|_{L^1_{T_0}(L^\infty)},$$
$$\|\mathrm{Id} - A_u(t)\|_{L^\infty} \le C\|Du\|_{L^1_{T_0}(L^\infty)},$$
$$\|J_u^{\pm 1}(t) - 1\|_{L^\infty} \le C\|Du\|_{L^1_{T_0}(L^\infty)}.$$

We also state the following lemma, the proof of which is straightforward.

Lemma 4.3. *For any function $f = f(x)$, define $\bar{f}$ according to (4-22). Then for any $p \in [1,\infty[$ one has*

$$\|\bar{f}\|_{L^p} \le \|J_u\|_{L^\infty}^{1/p}\|f\|_{L^p} \quad \text{and} \quad \|f\|_{L^p} \le \|J_u^{-1}\|_{L^\infty}^{1/p}\|\bar{f}\|_{L^p}.$$

After these preliminaries, we can recast our system in Lagrangian coordinates. First of all, from the mass equation in (0-1) and (4-24), we discover that

$$\partial_t(J_u\bar{\rho}) = 0, \quad \text{whence} \quad J_u\bar{\rho} = \rho_0. \tag{4-25}$$

Second, we notice that, in Lagrangian coordinates, the operator $\mathcal{L}$ can be written as

$$\widetilde{\mathcal{L}} f := -J_u^{-1}\big(\mu \operatorname{div}(\operatorname{adj}(D\psi_u)^t A_u \nabla f) - \lambda \operatorname{div}(\operatorname{adj}(D\psi_u)({}^t A_u : \nabla f))\big), \tag{4-26}$$

where we have used the notation $M : N := \operatorname{tr}(MN) = \sum_{ij} M_{ij}N_{ji}$.

Hence, thanks to (4-24) and (4-25), the momentum equation in (0-1) can be rewritten as

$$\rho_0 \psi_t \bar{u} + \widetilde{\mathcal{L}}\bar{u} = -\operatorname{div}(\operatorname{adj} D\psi_u P(J^{-1}\rho_0)). \tag{4-27}$$

4B2. *Stability estimates in Lagrangian coordinates.* In this section, we tackle the proof of uniqueness, by showing stability estimates for the Lagrangian formulation of our system.

More precisely, we consider initial data (ρ_0^j, u_0^j), for $j = 1, 2$, satisfying the hypotheses of Theorem 1.3. For the sake of simplicity and clarity, we focus on the case $\rho_0^1 = \rho_0^2 = \rho_0$, and suppose that ρ_0 satisfies the striated regularity assumption with respect to some fixed nondegenerate family of vector fields $\mathcal{X}_0$. The initial velocities do not need to be equal.

Let (ρ^1, u^1) and (ρ^2, u^2) be two solutions to system (0-1) on the time interval $[0, T]$, fulfilling the properties given by Theorem 1.3 and corresponding to the data (ρ_0, u_0^1) and (ρ_0, u_0^2), respectively. Setting $\varrho^j = \rho^j - 1$ for $j = 1, 2$, and defining w^j according to (3-3), the pairs (ϱ^j, w^j) solve (3-5)–(3-6) and also enjoy the regularity properties stated in Theorem 1.1. Moreover, as shown in Section 4A, for all j tangential regularity is propagated with respect to the nondegenerate family $\mathcal{X}^j$, which corresponds to the family $\mathcal{X}_0$ transported by the flow of u^j. Hence, for all $k_0 > 0$, there exists $T_0 > 0$ such that both ∇u^1 and ∇u^2 fulfill (4-20), which allows us to pass to Lagrangian coordinates, as shown in Section 4B1. Denoting, for $j = 1, 2$, the flow of u^j by ψ_j, setting $J_j := J_{u^j}$ and $A_j := A_{u^j}$, and taking advantage of the previous computations, we discover that $(\bar\rho^j, \bar u^j)_{j=1,2}$ satisfy the relations $J_j \bar\rho^j = \rho_0$ and

$$\rho_0 \psi_t \bar u^j + \widetilde{\mathcal{L}}_j \bar u^j = -\operatorname{div}(\operatorname{adj} D\psi_j P(J_j^{-1}\rho_0)),$$

where $\widetilde{\mathcal{L}}_j$ is the operator corresponding to u^j that has been defined by formula (4-26).

Let $\delta\bar u := \bar u^1 - \bar u^2$ and use similar notation for the other quantities. Taking the difference of the equations respectively for $\bar u^1$ and $\bar u^2$, we find that $\delta\bar u$ satisfies

$$\rho_0 \psi_t \delta\bar u + \mathcal{L}\delta\bar u = (\mathcal{L} - \widetilde{\mathcal{L}}_1)\delta\bar u + \delta\mathcal{L}\bar u^2 - \operatorname{div}(\delta \operatorname{adj} P(J_1^{-1}\rho_0)) \\ - \operatorname{div}\big(\operatorname{adj} D\psi_2 (P(J_1^{-1}\rho_0) - P(J_2^{-1}\rho_0))\big), \quad (4\text{-}28)$$

where we have set $\delta \operatorname{adj} := \operatorname{adj} D\psi_1 - \operatorname{adj} D\psi_2$. A slight adaptation of Lemma A.4 of [Danchin 2014] allows us to get the following bounds.

Lemma 4.4. *If (4-20) is fulfilled by u^1 and u^2 for some $k_0 \in \,]0, 1]$, then there exists a constant $C > 0$ just depending on k_0 such that the following estimates hold true for all times $t \in [0, T_0]$ and all $p \in [1, \infty]$:*

$$\|\operatorname{adj} D\psi_1(t) - \operatorname{adj} D\psi_2(t)\|_{L^p} \le C \int_0^t \|\nabla \delta u(\tau)\|_{L^p}\, d\tau,$$
$$\|A_1(t) - A_2(t)\|_{L^p} \le C \int_0^t \|\nabla \delta u(\tau)\|_{L^p}\, d\tau,$$
$$\|J_1^{\pm 1}(t) - J_2^{\pm 1}(t)\|_{L^p} \le C \int_0^t \|\nabla \delta u(\tau)\|_{L^p}\, d\tau.$$

We now perform energy estimates for (4-28): take the L^2 scalar product of both sides with $\delta\bar u$ and integrate by parts. In view of Lemmas 4.2 and 4.4, we deduce the following controls for the terms coming from the right-hand side:

$$\left|\int (\mathcal{L} - \widetilde{\mathcal{L}}_1)\delta\bar u \cdot \delta\bar u\, dx\right| \le C k_0 \|\nabla \delta\bar u\|_{L^2}^2,$$
$$\left|\int \delta\mathcal{L}\bar u^2 \cdot \delta\bar u\, dx\right| \le C\left(\int_0^t \|\nabla\delta\bar u\|_{L^2}\, d\tau\right)\|\nabla \bar u^2\|_{L^\infty}\|\nabla\delta\bar u\|_{L^2},$$
$$\left|\int \operatorname{div}(\delta \operatorname{adj} P(J_1^{-1}\rho_0)) \cdot \delta\bar u\, dx\right| \le C\left(\int_0^t \|\nabla\delta\bar u\|_{L^2}\, d\tau\right)\|\rho_0\|_{L^\infty}\|\nabla\delta\bar u\|_{L^2},$$
$$\left|\int \operatorname{div}\big(\operatorname{adj} D\psi_2 (P(J_1^{-1}\rho_0) - P(J_2^{-1}\rho_0))\big) \cdot \delta\bar u\, dx\right| \le C\left(\int_0^t \|\nabla\delta\bar u\|_{L^2}\, d\tau\right)\|\rho_0\|_{L^\infty}\|\nabla\delta\bar u\|_{L^2}.$$

Now, if k_0 in (4-20) has been taken small enough, then a repeated use of the Young and Cauchy–Schwarz inequalities leads us to the estimate

$$\frac{d}{dt}\int \rho_0|\delta\bar{u}|^2\,dx + \int |\nabla\delta\bar{u}|^2\,dx \le Ct(1+\|\nabla\bar{u}^2\|_{L^\infty}^2)\int_0^t \|\nabla\delta\bar{u}\|_{L^2}^2\,d\tau \tag{4-29}$$

for a new constant $C > 0$ that depends only on k_0, $\|\rho_0\|_{L^\infty}$, P, λ and μ.

In order to conclude uniqueness on $[0, T_0]$ by applying Gronwall's lemma, we need that

$$\int_0^{T_0} t(1+\|\nabla\bar{u}^2\|_{L^\infty}^2)\,dt < \infty. \tag{4-30}$$

In view of (4-23) and Lemma 4.3, it suffices to show that $t^{1/2}\nabla u^2$ is in $L^2_{T_0}(L^\infty)$. Now, recall

$$\nabla u^2 = \nabla w^2 - \nabla^2(\mathrm{Id}-\Delta)^{-1}P(\rho^2),$$

and, thanks to the estimates of Section 4A, one has that $\nabla^2(\mathrm{Id}-\Delta)^{-1}P(\rho^2)$ belongs to $L^\infty([0, T_0]\times\mathbb{R}^d)$. Therefore, our next (and final) goal is to show that there exists $T_0 > 0$ such that

$$\int_0^{T_0} t\|\nabla w^2\|_{L^\infty}^2\,dt < \infty. \tag{4-31}$$

This will be achieved thanks to the following proposition, the proof of which is postponed to the next subsection.

Proposition 4.5. *Under the hypotheses of Theorem 1.3, let us fix some*

$$R_2 > \max\left\{\frac{r_0}{2}, \frac{r_1}{2}, 2\right\} \quad \text{such that} \quad \frac{1}{2R_2} < \frac{3}{2} - \frac{1}{r_2} + \frac{d}{2p}, \quad \text{with } r_2 := r,$$

and set $R_0 = R_1 = 2R_2$. For $j \in \{0, 1, 2\}$, define moreover the indices

$$\alpha_j := \frac{1}{r_j} - \frac{1}{R_j} \quad \text{and} \quad \gamma_j := \frac{1}{r_j} - \frac{1}{2R_j}. \tag{4-32}$$

Then, there exists a positive time T_* such that one has the properties

$$t^{\alpha_2}\nabla^2 w \in L^{R_2}_{T_*}(L^p), \quad t^{\gamma_1}\nabla w \in L^{R_1}_{T_*}(L^p), \quad t^{\gamma_0} w \in L^{R_0}_{T_*}(L^\infty).$$

From the previous proposition, we immediately deduce the following corollary.

Corollary 4.6. *Under the assumptions of Theorem 1.3, one has*

$$\int_0^{T_*} t\|\nabla w(t)\|_{L^\infty}^2\,dt < \infty.$$

Proof. By Sobolev embedding, the stated inequality is a consequence of the following computation:

$$\begin{aligned}\int_0^{T_*} t\|\nabla^2 w(t)\|_{L^p}^2\,dt &= \int_0^{T_*} t^{1-\eta}t^{\eta}\|\nabla^2 w(t)\|_{L^p}^2\,dt \\ &\le \left(\int_0^{T_*} t^{\eta q}\|\nabla^2 w(t)\|_{L^p}^{2q}\,dt\right)^{1/q}\left(\int_0^{T_*} t^{(1-\eta)q'}\,dt\right)^{1/q'},\end{aligned}$$

where q' is the conjugate exponent of q. Take $q = R_2/2$, so that $1/q' = 1 - 2/R_2$ and impose the relation $q\eta = R_2\alpha_2$, getting in this way $\eta = 2\alpha_2$. With these choices and because $r_2 = r \in]1, 2[$, the condition $(1-\eta)q' > -1$ is satisfied, which completes the proof of the corollary. □

At this point, one can finish the proof of Theorem 1.3 by establishing the uniqueness of solutions. Let us define the function

$$E(t) := \|\sqrt{\rho_0}\delta\bar{u}(t)\|_{L^2}^2 + \int_0^t \|\nabla\delta\bar{u}(\tau)\|_{L^2}^2\, d\tau.$$

Up to choosing a smaller T_0, we can suppose that $T_0 = T_*$. Then, applying Gronwall's inequality to (4-29), we get, for all $t \in [0, T_*]$, the bound

$$E(t) \le E(0)\exp\biggl(C\int_0^t f(\tau)\, d\tau\biggr), \quad \text{where } f(t) := t(1 + \|\nabla\bar{u}^2(t)\|_{L^\infty}^2).$$

Since $E(0) \equiv 0$ and, by Corollary 4.6, $f \in L^1([0, T_*])$, we get uniqueness on $[0, T_*]$. Combining with a standard continuation argument, we then conclude uniqueness on the whole interval $[0, T]$.

4B3. *Maximal regularity with time weights.* For completeness, we still have to prove Proposition 4.5. First, we need the following lemma, which concerns the maximal regularity issue with time weights for the heat semigroup, and is strongly inspired by Lemma 3.2 of [Huang et al. 2013].

Lemma 4.7. *Let the exponents $(R_j, \alpha_j, \gamma_j)_{j\in\{0,1,2\}}$ be chosen as in Proposition 4.5. Let the operators $\mathcal{A}_1$ and $\mathcal{A}_0$ be defined as in Lemmas 2.7 and 2.8. Fix some $T > 0$, and assume that $t^{\alpha_2} f$ belongs to $L_T^{R_2}(L^p)$.*

Then one has $t^{1/r_1}\mathcal{A}_1 f \in L_T^\infty(L^p)$ and $t^{1/r_0}\mathcal{A}_0 f \in L_T^\infty(L^\infty)$, together with the estimates

$$\|t^{1/r_1}\mathcal{A}_1 f\|_{L_T^\infty(L^p)} + \|t^{1/r_0}\mathcal{A}_0 f\|_{L_T^\infty(L^\infty)} \le C\|t^{\alpha_2} f\|_{L_T^{R_2}(L^p)}.$$

Moreover, we have $t^{\alpha_1}\mathcal{A}_1 f(t) \in L_T^{R_1/(1+\delta)}(L^p)$ and $t^{\alpha_0}\mathcal{A}_0 f(t) \in L_T^{R_0/(1+\delta)}(L^\infty)$ for all $\delta > 0$, with the bounds

$$\|t^{\alpha_1}\mathcal{A}_1 f\|_{L_T^{R_1/(1+\delta)}(L^p)} + \|t^{\alpha_0}\mathcal{A}_0 f\|_{L_T^{R_0/(1+\delta)}(L^\infty)} \le C(T^{\delta/R_1} + T^{\delta/R_0})\|t^{\alpha_2} f\|_{L_T^{R_2}(L^p)}.$$

In particular, defining γ_0 and γ_1 according to (4-32), we have $t^{\gamma_1}\mathcal{A}_1 f \in L_T^{R_1}(L^p)$ and $t^{\gamma_0}\mathcal{A}_0 f \in L_T^{R_0}(L^\infty)$, and the following estimate is satisfied:

$$\|t^{\gamma_1}\mathcal{A}_1 f\|_{L_T^{R_1}(L^p)} + \|t^{\gamma_0}\mathcal{A}_0 f\|_{L_T^{R_0}(L^\infty)} \le C(T^{1/(2R_1)} + T^{1/(2R_0)})\|t^{\alpha_2} f\|_{L_T^{R_2}(L^p)}.$$

Proof. Regarding the operator $\mathcal{A}_1$, going along the lines of the proof of Lemma 2.7, one gets

$$\|\nabla e^{(t-s)\Delta} f(s,\cdot)\|_{L^p} \le C(t-s)^{-1/2}\|f(s)\|_{L^p} \quad \text{for all } 0 \le s \le t \le T,$$

which implies, after setting $1/R_2' = 1 - 1/R_2$, the inequality

$$\begin{aligned}\|\mathcal{A}_1 f(t)\|_{L^p} &\le C\int_0^t (t-s)^{-1/2} s^{-\alpha_2}\|s^{\alpha_2} f(s)\|_{L^p}\, ds\\ &\le C\biggl(\int_0^t (t-s)^{-R_2'/2} s^{-\alpha_2 R_2'}\, ds\biggr)^{1/R_2'}\|s^{\alpha_2} f\|_{L_T^{R_2}(L^p)}.\end{aligned}$$

Since $R_2 > 2$, we have $R_2'/2 < 1$, while, by our definition of α_2 in (4-32), we have $\alpha_2 R_2' < 1$. Therefore, performing the change of variable $s = t\tau$ inside the integral yields

$$\|\mathcal{A}_1 f(t)\|_{L^p} \le C t^{1/2-\alpha_2-1/R_2} \|s^{\alpha_2} f\|_{L_T^{R_2}(L^p)}.$$

On the one hand, since $1/2 - \alpha_2 - 1/R_2 = 1/2 - 1/r_2 = -1/r_1$, we have

$$\left\| t^{1/r_1} \|\mathcal{A}_1 f(t)\|_{L^p} \right\|_{L_T^\infty} \le C \|s^{\alpha_2} f\|_{L_T^{R_2}(L^p)}. \tag{4-33}$$

On the other hand, since $1/2 - \alpha_2 - 1/R_2 = -\alpha_1 - 1/R_1$, we also get that $t^{\alpha_1} \|\mathcal{A}_1 f(t)\|_{L^p}$ belongs to $L_T^{R_1/(1+\delta)}$ for all $\delta > 0$, and satisfies

$$\left\| t^{\alpha_1} \|\mathcal{A}_1 f(t)\|_{L^p} \right\|_{L_T^{R_1/(1+\delta)}} \le C_\delta \|s^{\alpha_2} f\|_{L_T^{R_2}(L^p)} T^{\delta/R_1}. \tag{4-34}$$

Taking $\delta = 1$ and interpolating between estimates (4-33) and (4-34), we get that $t^{\gamma_1} \|\mathcal{A}_1 f(t)\|_{L^p} \in L_T^{R_1}$, with the estimate

$$\left\| t^{\gamma_1} \|\mathcal{A}_1 f(t)\|_{L^p} \right\|_{L_T^{R_1}} \le C \|s^{\alpha_2} f\|_{L_T^{R_2}(L^p)} T^{1/(2R_1)}. \tag{4-35}$$

Proving the claimed bound for the term $\mathcal{A}_0$ follows along the same lines. First of all, setting p' to be the conjugate exponent of p, we can write

$$\begin{aligned}\|e^{(t-s)\Delta} f(s,\cdot)\|_{L^\infty} &\le C(t-s)^{-d/2} \left\| K_0\left(\frac{\cdot}{\sqrt{4\pi(t-s)}}\right)\right\|_{L^{p'}} \|f(s)\|_{L^p} \\ &\le C(t-s)^{-d/(2p)} s^{-\alpha_2} \|s^{\alpha_2} f(s)\|_{L^p}.\end{aligned}$$

Integrating this expression in time and applying the Hölder inequality once give us, much as above,

$$\|\mathcal{A}_0 f(t)\|_{L^\infty} \le C \left(\int_0^t (t-s)^{-d/(2p)R_2'} s^{-\alpha_2 R_2'}\, ds\right)^{1/R_2'} \|s^{\alpha_2} f\|_{L_T^{R_2}(L^p)}.$$

Once again, thanks to our choice of R_2 we have that $d/(2p)R_2' < 1$ (recall that $p > d$); hence, repeating the change of variable $s = t\tau$ we find

$$\|\mathcal{A}_0 f(t)\|_{L^\infty} \le C t^{1-d/(2p)-\alpha_2-1/R_2} \|t^{\alpha_2} f\|_{L_T^{R_2}(L^p)}.$$

Now, first we remark that $1 - d/(2p) - \alpha_2 - 1/R_2 = -1/r_0$, and hence $t^{1/r_0} \|\mathcal{A}_0 f(t)\|_{L^\infty} \in L_T^\infty$, with

$$\left\| t^{1/r_0} \|\mathcal{A}_0 f(t)\|_{L^\infty} \right\|_{L_T^\infty} \le C \|s^{\alpha_2} f\|_{L_T^{R_2}(L^p)}. \tag{4-36}$$

Then, we also notice that $1 - d/(2p) - \alpha_2 - 1/R_2 = -\alpha_0 - 1/R_0$, so that $t^{\alpha_0} \|\mathcal{A}_0 f(t)\|_{L^\infty}$ belongs to $L_T^{R_0/(1+\delta)}$ for all $\delta > 0$, and satisfies the estimate

$$\left\| t^{\alpha_0} \|\mathcal{A}_0 f(t)\|_{L^\infty} \right\|_{L_T^{R_0/(1+\delta)}} \le C_\delta \|s^{\alpha_2} f\|_{L_T^{R_2}(L^p)} T^{\delta/R_0}. \tag{4-37}$$

As above, taking $\delta = 1$ and interpolating between estimates (4-36) and (4-37), we finally deduce the property $t^{\gamma_0} \|\mathcal{A}_0 f(t)\|_{L^\infty} \in L_T^{R_0}$, together with the estimate

$$\left\| t^{\gamma_0} \|\mathcal{A}_0 f(t)\|_{L^\infty} \right\|_{L_T^{R_0}} \le C \|s^{\alpha_2} f\|_{L_T^{R_2}(L^p)} T^{1/(2R_0)}. \tag{4-38}$$

The lemma is now proved. □

Finally, we need the following lemma, which was established in [Huang et al. 2013]:

Lemma 4.8. *Let $1 < R, p < \infty$, and let $\alpha \geq 0$ be such that $\alpha + 1/R < 1$. Suppose that $t^\alpha f$ belongs to $L^R_T(L^p)$ for some $T \in]0, \infty]$.*

Then also $t^\alpha \mathcal{A}_2 f$ belongs to $L^R_T(L^p)$, and one has the estimate

$$\|t^\alpha \mathcal{A}_2 f\|_{L^R_T(L^p)} \leq C\|t^\alpha f\|_{L^R_T(L^p)}.$$

Now, we are in the position of proving Proposition 4.5.

Proof of Proposition 4.5. Recall that w satisfies (3-6), and thus

$$w(t) = e^{-t\mathcal{L}} w_0 - \int_0^t e^{(s-t)\mathcal{L}}(\rho F)(s)\,ds \quad \text{with } F \text{ given by (3-7)}. \tag{4-39}$$

Let us first study the term containing the initial data. By hypothesis, $\nabla^2 w_0 \in \dot{B}^{-2/r}_{p,r} \hookrightarrow \dot{B}^{-2/r}_{p,R_2}$, since $R_2 > 2 > r$ by our definitions. Thanks to Proposition 2.9, this implies that $t^{\alpha_2}\nabla^2 e^{-t\mathcal{L}} w_0$ belongs to $L^{R_2}(\mathbb{R}_+; L^p(\mathbb{R}^d))$.

In the same way, we have $\nabla w_0 \in \dot{B}^{-2/r_1}_{p,r_1} \hookrightarrow \dot{B}^{-2/r_1}_{p,R_1}$ and $w_0 \in \dot{B}^{-2/r_0}_{\infty,r_0} \hookrightarrow \dot{B}^{-2/r_0}_{\infty,R_0}$, because we have taken $R_1 = R_0 = 2R_2 > \max\{r_0, r_1\}$. From these properties we deduce that $t^{\alpha_1}\nabla e^{-t\mathcal{L}} w_0 \in L^{R_1}(\mathbb{R}_+; L^p(\mathbb{R}^d))$ and that $t^{\alpha_0} e^{-t\mathcal{L}} w_0 \in L^{R_0}(\mathbb{R}_+; L^\infty(\mathbb{R}^d))$. Since now both γ_1 and γ_0 are greater than α_1 and α_0 respectively, we get that, for all $T > 0$ fixed, $t^{\gamma_1}\nabla e^{-\mathcal{L}t} w_0 \in L^{R_1}_T(L^p)$ and $t^{\gamma_0} e^{-t\mathcal{L}} w_0 \in L^{R_0}_T(L^\infty)$.

As for the forcing term of (4-39), we apply Lemma 4.8 with $R = R_2$ and $\alpha = \alpha_2$ (note that $\alpha_2 + 1/R_2 = 1/r < 1$). We also apply Lemma 4.7. Therefore, if we set

$$\widetilde{\mathcal{N}}(T) := \|\varrho\|_{L^\infty_T(L^p \cap L^\infty)} + \|t^{\gamma_0} w\|_{L^{R_0}_T(L^\infty)} + \|t^{\gamma_1}\nabla w\|_{L^{R_1}_T(L^p)} + \|t^{\alpha_2}\nabla^2 w\|_{L^{R_2}_T(L^p)},$$

arguing exactly as in Section 3B2, we get for some constant C_T bounded by a positive power of T,

$$\widetilde{\mathcal{N}}(T) \leq C_T(\|\varrho_0\|_{L^p \cap L^\infty} + \|w_0\|_{\dot{B}^{2-2/r}_{p,r}} + \|t^{\alpha_2}\rho F\|_{L^{R_2}_T(L^p)}), \tag{4-40}$$

where F is defined in (3-7). At this point, we bound the term $\|t^{\alpha_2}\rho(t)F(t)\|_{L^{R_2}_T(L^p)}$ by following the computations of Section 3B2: first of all, (3-19) is now replaced by the control

$$\|t^{\alpha_2}(\mathrm{Id} - \Delta)^{-1}\nabla P\|_{L^{R_2}_T(L^p)} \leq CT^{\alpha_2 + 1/R_2}\|\varrho\|_{L^\infty_T(L^p)} \leq C_T\widetilde{\mathcal{N}}(T).$$

Next, we have, noting that our conditions on the exponents imply $\alpha_2 > \gamma_0 + \gamma_1$,

$$\begin{aligned}
&\|t^{\alpha_2}(w \cdot \nabla w + w \cdot \nabla v + v \cdot \nabla w + v \cdot \nabla v)\|_{L^{R_2}_T(L^p)} \\
&\quad\leq CT^{\alpha_2 - (\gamma_0+\gamma_1)}(\|t^{\gamma_0} v\|_{L^{R_0}_T(L^\infty)} + \|t^{\gamma_0} w\|_{L^{R_0}_T(L^\infty)})(\|t^{\gamma_1}\nabla v\|_{L^{T_1}_T(L^p)} + \|t^{\gamma_1}\nabla w\|_{L^{R_1}_T(L^p)}) \\
&\quad\leq CT^{\alpha_2 - (\gamma_0+\gamma_1)}(1 + T^{\gamma_0})(1 + T^{\gamma_1})\widetilde{\mathcal{N}}^2(T),
\end{aligned}$$

and estimate (3-22) becomes

$$\begin{aligned}
\|t^{\alpha_2}\nabla(\mathrm{Id} - \Delta)^{-1}\operatorname{div}(Pu)\|_{L^{R_2}_T(L^p)} &\leq CT^{\alpha_2 - \gamma_0}\|P\|_{L^{R_1}_T(L^p)}(\|t^{\gamma_0} v\|_{L^{R_0}_T(L^\infty)} + \|t^{\gamma_0} w\|_{L^{R_0}_T(L^\infty)}) \\
&\leq CT^{\alpha_2 - \gamma_0 + 1/R_1}(1 + T^{\gamma_0})\widetilde{\mathcal{N}}^2(T).
\end{aligned}$$

Finally, we have

$$\|t^{\alpha_2}\nabla(\mathrm{Id}-\Delta^{-1})\,\mathrm{div}\,u\|_{L_T^{R_2}(L^p)} \le CT^{\alpha_2-\gamma_1}(\|t^{\gamma_1}\nabla w\|_{L_T^{R_2}(L^p)}+\|t^{\gamma_1}P\|_{L_T^{R_2}(L^p)})$$
$$\le CT^{\alpha_2-\gamma_1+1/(2R_2)}(1+T^{1/(2R_2)+\gamma_1})\widetilde{\mathcal{N}}(T),$$

and, arguing in a pretty similar way, we also get

$$\|t^{\alpha_2}\nabla(\mathrm{Id}-\Delta)^{-1}\big((g(\rho)-g(1))\,\mathrm{div}\,u\big)\|_{L_T^{R_2}(L^p)}$$
$$\le CT^{\alpha_2-\gamma_1}\|\varrho\|_{L_T^\infty(L^\infty)}(\|t^{\gamma_1}\Delta(\mathrm{Id}-\Delta)^{-1}P(\rho)\|_{L_T^{R_2}(L^p)}+\|t^{\gamma_1}\,\mathrm{div}\,w\|_{L_T^{R_2}(L^p)})$$
$$\le T^{\alpha_2-\gamma_1+1/(2R_2)}(1+T^{1/(2R_2)+\gamma_1})\widetilde{\mathcal{N}}^2(T).$$

Putting all these bounds together, we end up with

$$\|t^{\alpha_2}\rho F\|_{L_T^{R_2}(L^p)} \le C_T(\widetilde{\mathcal{N}}(T)+\widetilde{\mathcal{N}}^2(T)).$$

Therefore, we can insert the previous inequality into (4-40): the application of a standard bootstrap argument allows us to find a time $T_* > 0$ such that, for all $t \in [0, T_*]$, one has

$$\widetilde{\mathcal{N}}(t) \le C(\|\varrho_0\|_{L^p\cap L^\infty}+\|w_0\|_{\dot B^{2-2/r}_{p,r}})$$

for a suitable positive constant C, which completes the proof of Proposition 4.5. □

Appendix: Harmonic analysis estimates

This appendix is devoted to the proofs of Lemma 2.10 and Proposition 2.13.

Proof of Lemma 2.10. It is based on the following Bony's paraproduct decomposition (first introduced in [Bony 1981]) for the (formal) product of two tempered distributions u and v:

$$uv = T_u v + T_v u + R(u, v), \tag{A-1}$$

where we have defined

$$T_u v := \sum_j S_{j-1}u\Delta_j v \qquad \text{and} \qquad R(u, v) := \sum_j \tilde\Delta_j u\Delta_j v, \quad \text{with } \tilde\Delta_j := \sum_{|j'-j|\le 1}\Delta_{j'}.$$

The above operator T is called *paraproduct*, whereas R is called *remainder*. We refer to Chapter 2 of [Bahouri et al. 2011] for a presentation of continuity properties of the previous operators in the class of Besov spaces. For the time being, we limit ourselves to pointing out that the generic term $S_{j-1}u\Delta_j v$ of $T_u v$ is spectrally supported on dyadic annuli with radius of size about 2^j, while the generic term $\tilde\Delta_j u\Delta_j v$ of $R(u, v)$ is supported on dyadic balls of size about 2^j.

One can now start the proof of Lemma 2.10. By using Bony's decomposition (A-1) and a commutator's process, we get, setting $\widetilde X := (\mathrm{Id}-S_0)X$,

$$\psi_X\sigma(D)g = \sigma(D)\,\mathrm{div}(Xg)+[T_{X^k};\sigma(D)\psi_k]g-\sigma(D)\psi_k T_g X^k-\sigma(D)\psi_k R(\widetilde X^k, g)$$
$$+T_{\sigma(D)\psi_k g}X^k+R(\widetilde X^k,\sigma(D)\psi_k g)+\big(R(S_0X^k,\sigma(D)\psi_k g)-\sigma(D)\psi_k R(S_0X^k, g)\big). \tag{A-2}$$

Bounding the first term relies on the fact that multiplier operators in S^{-1} map $B^0_{p,\infty}$ to $B^1_{p,\infty}$ (see [Bahouri et al. 2011, Proposition 2.78]) and that L^p is embedded in $B^0_{p,\infty}$. We thus have

$$\begin{aligned}\|\sigma(D)\operatorname{div}(Xg)\|_{B^1_{p,\infty}} &\le C\|\operatorname{div}(Xg)\|_{B^0_{p,\infty}}\\ &\le C\|\operatorname{div}(Xg)\|_{L^p} \le C(\|\psi_X g\|_{L^p} + \|\nabla X\|_{L^p}\|g\|_{L^\infty}).\end{aligned}$$

Next, to handle the third term of (A-2), we use the fact that, being in S^0, the operator $\sigma(D)\partial_k$ maps $B^1_{p,\infty}$ to itself (again, see [Bahouri et al. 2011, Proposition 2.78]), that the paraproduct operator T maps $L^\infty \times B^1_{p,\infty}$ to $B^1_{p,\infty}$ and [Bahouri et al. 2011, Remark 2.83], and that L^p is embedded in $B^0_{p,\infty}$. We eventually get

$$\begin{aligned}\|\sigma(D)\psi_k T_g X^k\|_{B^1_{p,\infty}} &\le C\|T_g X\|_{B^1_{p,\infty}}\\ &\le C\|g\|_{L^\infty}\|\nabla X\|_{B^0_{p,\infty}} \le C\|g\|_{L^\infty}\|\nabla X\|_{L^p}.\end{aligned}$$

Similarly, since the remainder operator R maps $L^\infty \times B^1_{p,\infty}$ to $B^1_{p,\infty}$ and because, owing to the low-frequency cut-off, we have

$$\|\widetilde{X}\|_{B^1_{p,\infty}} \le C\|\nabla X\|_{B^0_{p,\infty}} \le C\|\nabla X\|_{L^p}, \tag{A-3}$$

we readily get

$$\|\sigma(D)\psi_k R(\widetilde{X}^k, g)\|_{B^1_{p,\infty}} \le C\|g\|_{L^\infty}\|\nabla X\|_{L^p}.$$

Regarding the term $R(\widetilde{X}^k, \sigma(D)\psi_k g)$, we just have to use (A-3) and that R maps also $B^0_{\infty,\infty} \times B^1_{p,\infty}$ to $B^1_{p,\infty}$, to get

$$\|R(\widetilde{X}^k, \sigma(D)\psi_k g)\|_{B^1_{p,\infty}} \le C\|\sigma(D)\psi_k g\|_{B^0_{\infty,\infty}}\|\nabla X\|_{L^p}.$$

Since $\sigma(D)\psi_k$ maps $B^0_{\infty,\infty}$ to itself, and because $L^\infty \hookrightarrow B^0_{\infty,\infty}$, that term also satisfies the required inequality.

The term $T_{\sigma(D)\partial_k g} X^k$ turns out to be the only one that cannot be bounded in $B^1_{p,\infty}$ under our assumptions. In fact, for that term, we use that the paraproduct maps $B^{s-1}_{\infty,\infty} \times B^1_{p,\infty}$ to $B^s_{p,\infty}$ (as $s-1<0$) to write (still using [Bahouri et al. 2011, Remark 2.83]),

$$\|T_{\sigma(D)\partial_k g} X^k\|_{B^s_{p,\infty}} \le C\|\sigma(D)\partial_k g\|_{B^{s-1}_{\infty,\infty}}\|\nabla X^k\|_{L^p}.$$

Because $\sigma(D)\partial_k$ maps $B^{s-1}_{\infty,\infty}$ to itself, and $L^\infty \hookrightarrow B^{s-1}_{\infty,\infty}$, we get

$$\|T_{\sigma(D)\partial_k g} X^k\|_{B^s_{p,\infty}} \le C\|g\|_{L^\infty}\|\nabla X\|_{L^p}.$$

To conclude the proof, it is only a matter of bounding suitably the two commutator terms in (A-2). First of all, notice that since the general term of the paraproduct is spectrally supported in dyadic annuli, one may find a smooth function ψ supported in some annulus centered at the origin, and such that

$$[T_{X^k}; \sigma(D)\psi_k]g = \sum_{j\in\mathbb{Z}} [S_{j-1}X^k, \psi(2^{-j}D)\sigma(D)\psi_k]\Delta_j g. \tag{A-4}$$

For each fixed $j\in\mathbb{Z}$ and $k\in\{1,\dots,d\}$, let us define $h^k_j := i\mathcal{F}^{-1}(\xi_k\psi(2^{-j}\cdot)\sigma)$. Then we have, thanks to the definition of $h^k_j(D)$ and the mean value formula,

$$\begin{aligned}[S_{j-1}X^k,\psi(2^{-j}D)\sigma(D)\psi_k]\Delta_jg(x) &= \int_{\mathbb{R}^d} h^k_j(y)(S_{j-1}X^k(x)-S_{j-1}X^k(x-y))\Delta_jg(x-y)\,dy\\ &= -\int_0^1\int_{\mathbb{R}^d} h^k_j(y)y\cdot\nabla S_{j-1}X^k(x-\tau y)\Delta_jg(x-y)\,dy\,d\tau\\ &= -\int_0^1\int_{\mathbb{R}^d} h^k_j\Big(\frac{z}{\tau}\Big)\Big(\frac{z}{\tau}\Big)\cdot\nabla S_{j-1}X^k(x-z)\Delta_jg\Big(x-\frac{z}{\tau}\Big)\frac{dz}{\tau^d}\,d\tau.\end{aligned}$$

From the last line and convolution inequalities, we get

$$\|[S_{j-1}X^k,\psi(2^{-j}D)\sigma(D)\psi_k]\Delta_jg\|_{L^p}\le\||\cdot|h^k_j\|_{L^1}\|\Delta_jg\|_{L^\infty}\|\nabla S_{j-1}X^k\|_{L^p},$$

which, admitting for a while that

$$\||\cdot|h^k_j\|_{L^1}\le C2^{-j} \tag{A-5}$$

and using the definition of the norm in $B^1_{p,\infty}$ implies

$$\|[T_{X^k};\sigma(D)\psi_k]g\|_{B^1_{p,\infty}}\le C\|g\|_{L^\infty}\|\nabla X\|_{L^p}.$$

In order to prove (A-5), we use the fact that, performing integration by parts,

$$(1+|z|^2)^d(zh^k_j(z)) = (2\pi)^{-d}\int e^{iz\cdot\xi}(\mathrm{Id}-\Delta)^d\nabla(\xi_k\psi(2^{-j}\cdot)\sigma)(\xi)\,d\xi.$$

As integration may be restricted to those $\xi\in\mathbb{R}^d$ such that $|\xi|\sim 2^j$ and since σ is in S^{-1}, routine computations lead to

$$(1+|z|^2)^d|zh^k_j(z)|\le C2^{-j}\quad\text{for all } z\in\mathbb{R}^d,$$

whence (A-5) follows.

In order to bound the last term of (A-2), we use the fact that, owing to the properties of the localization of the Littlewood–Paley decomposition, we have for some suitable smooth function ψ with compact support and value 1 on some neighborhood of the origin,

$$R(S_0X^k,\sigma(D)\psi_kg)-\sigma(D)\psi_kR(S_0X^k,g)=\sum_{j=-1}^{0}[\Delta_jS_0X^k,\sigma(D)\psi(D)\partial_k]\tilde{\Delta}_jg.$$

Then, arguing as above and setting $h^k := \mathcal{F}^{-1}(i\xi^k\psi\sigma)$, we find that

$$[\Delta_jS_0X^k,\sigma(D)\psi(D)\partial_k]\tilde{\Delta}_jg(x)=\int_0^1\int_{\mathbb{R}^d} h^k(y)y\cdot\nabla\Delta_{j-1}S_0X^k(x-\tau y)\tilde{\Delta}_jg(x-y)\,dy\,d\tau.$$

Hence convolution inequalities and the fact that the only nonzero terms above correspond to $j=0,1$, lead us to

$$\begin{aligned}\|R(S_0X^k,\sigma(D)\psi_kg)-\sigma(D)\psi_kR(S_0X^k,g)\|_{L^p} &\le C2^{-j}\|\nabla\Delta_{j-1}S_0X^k\|_{L^p}\|\tilde{\Delta}_jg\|_{L^\infty}\\ &\le C2^{-j}\|\nabla X\|_{L^p}\|g\|_{L^\infty}.\end{aligned}$$

This completes the proof of Lemma 2.10. □

Proof of Proposition 2.13. For all $1 \le j \le d$ and $\eta > 0$, let us introduce the *modified Riesz transform*

$$\mathcal{R}_j^{(\eta)} := \partial_j(\eta\mathrm{Id} - \Delta)^{-1/2}, \tag{A-6}$$

so that $\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} = \partial_i\partial_j(\eta\mathrm{Id} - \Delta)^{-1}$.

Proposition 2.13 follows from Proposition 2.12 and the following lemma involving the tangential regularity with respect to *only one* vector field.

Lemma A.1. *Let $p \in]1, \infty[$ and take a vector field $X \in \mathbb{L}^{\infty,p}$. Let $g \in L^\infty$ be such that $g \in \mathbb{L}_X^p$ and $\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g \in L^\infty$ for some $\eta > 0$. There exists a constant $C > 0$ such that*

$$\|\partial_X \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g\|_{L^p} \le C\big((\|\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g\|_{L^\infty} + \|g\|_{L^\infty})\|\nabla X\|_{L^p} + \|\partial_X g\|_{L^p}\big).$$

For proving that lemma, a few reminders concerning the Hardy–Littlewood *maximal function* $M[f]$ of a function f in $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ are in order. Recall that it is defined by

$$M[f](x) := \sup_{r>0} \frac{1}{|B(x,r)|}\int_{B(x,r)} |f(y)|\,dy,$$

where $B(x,r)$ denotes the ball in $\mathbb{R}^d$ of center x and radius r, and $|B(x,r)|$ its Lebesgue measure.

The following statement is classical (for the proof, see, e.g., Chapter 1 of [Stein 1993]).

Lemma A.2. *The following properties hold true*:

(a) *For any $1 < p \le \infty$, there exist constants $0 < c < C$ such that for any function g in $L^p(\mathbb{R}^d)$*

$$c\|g\|_{L^p} \le \|M[g]\|_{L^p} \le C\|g\|_{L^p}.$$

(b) *Let $p, q \in]1, \infty[$ or $p = q = \infty$. Let $\{f_j\}_{j\in\mathbb{Z}}$ be a sequence of functions in $L^p(\mathbb{R}^d)$ such that $(f_j)_{\ell^q(\mathbb{Z})} \in L^p(\mathbb{R}^d)$. Then there holds*

$$\|(M[f_j])_{\ell^q}\|_{L^p} \le C\|(f_j)_{\ell^q}\|_{L^p}.$$

(c) *For any fixed $\Phi \in L^1(\mathbb{R}^d)$ such that $\Psi(x) = \sup_{|y|\ge|x|} |\Phi(y)| \in L^1(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \Psi(x)\,dx = A$, and any function g, we have for all $x \in \mathbb{R}^d$*

$$\sup_{t>0} |g * \Phi_t(x)| \le CM[g](x), \quad \textit{with } \Phi_t(x) := t^{-d}\Phi\left(\frac{x}{t}\right).$$

We shall also need the following definition.

Definition A.3. Let $s \in \mathbb{R}$ and $p \in]1, \infty[$. The *homogeneous Sobolev space* $\dot{W}^{s,p}$ is defined as the set of $u \in \mathcal{S}'_h$ such that

$$\|u\|_{\dot{W}^{s,p}} := \|(-\Delta)^{s/2}u\|_{L^p} < \infty.$$

The spaces L^p and $\dot{W}^{s,p}$ may be characterized in terms of Littlewood–Paley decomposition as they come up as special Triebel–Lizorkin spaces (see, e.g., [Runst and Sickel 1996, Chapter 2]).

Proposition A.4. *Let $s \in \mathbb{R}$ and $p \in]1, \infty[$. Then one has the following equivalences of norms*:

$$\|u\|_{\dot{W}^{s,p}} \sim \big\|\|(2^{js}\dot{\Delta}_j u)_j\|_{\ell^2(\mathbb{Z})}\big\|_{L^p} \quad \textit{and} \quad \|u\|_{L^p} \sim \big\|\|(\Delta_j u)_j\|_{\ell^2(j\ge -1)}\big\|_{L^p}.$$

Proof of Lemma A.1. We start the proof by remarking that

$$\psi_X \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g = \operatorname{div}(X\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g) - \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g \operatorname{div} X.$$

Since

$$\|\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g \operatorname{div} X\|_{L^p} \le C\|\nabla X\|_{L^p}\|\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g\|_{L^\infty},$$

we just need to bound the term $\operatorname{div}(X\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g)$ in L^p.

To this end, we resort again to Bony's decomposition (A-1) and get

$$\operatorname{div}(X\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g) = \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} \operatorname{div}(Xg) + \psi_k[T_{X^k}; \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)}]g + \psi_k T_{\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g} X^k \\ - \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} \psi_k T_g X^k + \psi_k R(X^k, \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g) - \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} \psi_k R(X^k, g). \quad \text{(A-7)}$$

The first term in the right-hand side of the previous relation may be bounded by means of Corollary 2.4 and identity (1-7):

$$\|\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} \operatorname{div}(Xg)\|_{L^p} \le C\|\operatorname{div}(Xg)\|_{L^p} \le C(\|\partial_X g\|_{L^p} + \|\nabla X\|_{L^p}\|g\|_{L^\infty}).$$

Next, since we have

$$\|S_{m-1}\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g\|_{L^\infty} \le C\|\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g\|_{L^\infty},$$

we get for all $\ell \ge -1$, thanks to Lemma A.2(c), and using the fact that $S_{m-1} = 0$ for $m \le 0$,

$$\begin{aligned} |(\Delta_\ell \psi_k T_{\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g} X^k)(x)| &\le C 2^\ell \sum_{|m-\ell|\le 4} M[S_{m-1}\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g \Delta_m X^k](x) \\ &\le C \sum_{|m-\ell|\le 4} \|S_{m-1}\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g\|_{L^\infty} M[2^m \Delta_m X^k](x) \\ &\le C\|\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g\|_{L^\infty} \sum_{\substack{|m-\ell|\le 4 \\ m\ge 1}} M[2^m \Delta_m X^k](x). \end{aligned}$$

As a result, thanks to Proposition A.4 (where we take $s = 0$) and point (b) of Lemma A.2, we get

$$\begin{aligned} \|\psi_k T_{\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g} X^k\|_{L^p} &\le C\|\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g\|_{L^\infty} \left\| \left(\sum_{\ell\ge 1} (M[2^\ell \dot\Delta_\ell X^k])^2 \right)^{1/2} \right\|_{L^p} \\ &\le C\|\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g\|_{L^\infty}\|\nabla X\|_{L^p}. \end{aligned}$$

Using the same strategy for handling $\psi_k T_g X^k$, we also obtain

$$\|\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} \psi_k T_g X^k\|_{L^p} \le C\|\psi_k T_g X^k\|_{L^p} \le C\|g\|_{L^\infty}\|\nabla X\|_{L^p}.$$

For the third term of the second line in (A-7), we remove the low frequencies of X and consider the *modified* remainder defined by

$$\widetilde{R}(X^k, \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g) = \sum_{m\ge 0} \Delta_m X^k \tilde\Delta_m \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g.$$

Then we write that for all $\ell \geq -1$, taking advantage of Lemma A.2,

$$
\begin{aligned}
|\Delta_\ell \psi_k \widetilde{R}(X^k, \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g)(x)| &\leq C 2^\ell \sum_{m \geq \max(0,\ell-5)} M[\Delta_m X^k \tilde{\Delta}_m \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g](x) \\
&\leq C \|\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g\|_{L^\infty} \sum_{m \geq \max(0,\ell-5)} 2^{\ell-m} M[2^m \Delta_m X^k](x).
\end{aligned}
$$

Hence, from Proposition A.4 and Young's inequality for convolutions, we infer that

$$
\begin{aligned}
\|\psi_k \widetilde{R}(X^k, \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} g)\|_{L^p} &\leq C \|g\|_{L^\infty} \bigg(\sum_{\ell \geq -5} 2^{-\ell} \bigg) \Big\| \|(M[2^j \Delta_j X](x))_{j \in \mathbb{N}}\|_{\ell^2} \Big\|_{L^p} \\
&\leq C \|g\|_{L^\infty} \|\nabla X\|_{L^p}.
\end{aligned}
$$

Obviously, the same estimate holds true for the term $\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} \psi_k \widetilde{R}(X^k, g)$.

The low-frequency terms that have been discarded have to be treated together; that is, we have to bound $\psi_k[\Delta_{-1} X^k, \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)}]\tilde{\Delta}_{-1} g$ in L^p, and this may be done in the same way as at the end of the proof of Lemma 2.10. We end up with

$$
\|\psi_k[\Delta_{-1} X^k, \mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)}]\tilde{\Delta}_{-1} g\|_{L^p} \leq C \|\Delta_{-1} \nabla X\|_{L^p} \|\tilde{\Delta}_{-1} g\|_{L^\infty} \leq C \|\nabla X\|_{L^p} \|g\|_{L^\infty}.
$$

Finally, it remains to handle the commutator term on the right-hand side of (A-7). We start by decomposing $\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)}$ into

$$
\mathcal{R}_i^{(\eta)}\mathcal{R}_j^{(\eta)} = \mathcal{R}_i \mathcal{R}_j - \eta(\eta \mathrm{Id} - \Delta)^{-1} \mathcal{R}_i \mathcal{R}_j, \tag{A-8}
$$

where $\mathcal{R}_j$ stands for the classical Riesz transform (which corresponds to taking $\eta = 0$ in (A-6)). Let us first consider the part of the commutator corresponding to $\mathcal{R}_i \mathcal{R}_j$. We have

$$
([T_{X^k}; \mathcal{R}_i \mathcal{R}_j] g)(x) = \sum_{m \geq 1} [S_{m-1} X^k, \mathcal{R}_i \mathcal{R}_j] \Delta_m g.
$$

Let $\theta(\xi) = -\xi_i \xi_j |\xi|^{-2} \varphi(\xi)$, with φ being the function used in the Littlewood–Paley decomposition. Since the generic term $[S_{m-1} X^k, \mathcal{R}_i \mathcal{R}_j] \Delta_m g$ is supported on dyadic annuli of size 2^m, one can write

$$
\begin{aligned}
&[S_{m-1} X^k, \mathcal{R}_i \mathcal{R}_j] \Delta_m g \\
&\quad = \sum_{|m-\ell| \leq 4} (S_{m-1} X^k \Delta_m \mathcal{R}_i \mathcal{R}_j \Delta_\ell g - \mathcal{R}_i \mathcal{R}_j \Delta_\ell (S_{m-1} X^k \Delta_m g)) \\
&\quad = \sum_{|m-\ell| \leq 4} \big(S_{m-1} X^k (S_{m+1} \theta(2^{-\ell} D) g - S_m \theta(2^{-\ell} D) g) - \theta(2^{-\ell} D)(S_{m-1} X^k (S_{m+1} g - S_m g)) \big).
\end{aligned}
$$

Therefore, by applying Abel's rearrangement technique and using that $\Delta_j = \dot{\Delta}_j$ for $j \in \mathbb{N}$, we get

$$
[T_{X^k}; \mathcal{R}_i \mathcal{R}_j] g = -\sum_{m \geq 2} \sum_{|m-\ell| \leq 4} [\dot{\Delta}_{m-2} X, \theta(2^{-\ell} D)] S_m g.
$$

The general term of the above series is spectrally supported in dyadic annuli of size about 2^m. Therefore, there exists some universal integer N_0 so that for all $q \in \mathbb{Z}$,

$$\dot\Delta_q \partial_k [T_{X^k}; \mathcal{R}_i \mathcal{R}_j] g = \sum_{|m-q|\le N_0} \sum_{|\ell-m|\le 4} \dot\Delta_q \partial_k [\theta(2^{-\ell} D), \dot\Delta_{m-2} X] S_m g. \tag{A-9}$$

Now, part (c) of Lemma A.2 ensures that for all $x \in \mathbb{R}^d$

$$\left|\dot\Delta_q \partial_k [\theta(2^{-\ell} D), \dot\Delta_{m-2} X] S_m g(x)\right| \le C 2^q M[[\theta(2^{-\ell} D), \dot\Delta_{m-2} X] S_m g](x) \tag{A-10}$$

and the mean value formula gives us, setting $\check\theta := \mathcal{F}^{-1}\theta$,

$$([\theta(2^{-\ell} D), \dot\Delta_{m-2} X^k] S_m g)(x) = -2^{d\ell} \int_0^1 \int_{\mathbb{R}^d} \check\theta(2^\ell z) z \cdot \nabla \dot\Delta_{m-2} X^k (x - \tau z) S_m g(x-z)\, dz\, d\tau,$$

whence, performing a change of variables and setting $\Psi(z) := z\check\theta(z)$,

$$([\theta(2^{-\ell} D), \dot\Delta_{m-2} X^k] S_m g)(x) = -2^{-\ell} \int_0^1 \int_{\mathbb{R}^d} \left(\frac{2^\ell}{\tau}\right)^d \Psi\left(\frac{2^\ell z}{\tau}\right) \cdot \nabla \dot\Delta_{m-2} X^k (x-z) S_m g\left(x - \frac{z}{\tau}\right) dz\, d\tau.$$

From that latter relation, we deduce that

$$\begin{aligned} M[[\theta(2^{-\ell} D), \dot\Delta_{m-2} X] S_m g](x) &\le C 2^{-\ell} \|g\|_{L^\infty} \int_0^1 \int_{\mathbb{R}^d} \left(\frac{2^\ell}{\tau}\right)^d \left|\Psi\left(\frac{2^\ell z}{\tau}\right)\right| M[\nabla \dot\Delta_{m-2} X^k (x-z)]\, dz\, d\tau \\ &\le C 2^{-\ell} \|g\|_{L^\infty} M[\dot\Delta_{m-2} \nabla X](x). \end{aligned}$$

We now plug that inequality into (A-10) and (A-9), then take the norm in $\ell^2(\mathbb{Z})$ with respect to q and eventually compute the norm in $L^p(\mathbb{R}^d)$. We end up with

$$\|\partial_k \dot\Delta_q [T_{X^k}; \mathcal{R}_i \mathcal{R}_j] g\|_{L^p(\ell^2(\mathbb{Z}))} \le C \|M(\dot\Delta_q \nabla X)\|_{L^p(\ell^2(\mathbb{Z}))}.$$

Therefore, by applying Proposition A.4 with $s = 0$ and part (b) of Lemma A.2, we finally get

$$\|\psi_k [T_{X^k}; \mathcal{R}_i \mathcal{R}_j] g\|_{L^p} \le C \|g\|_{L^\infty} \|\nabla X\|_{L^p}. \tag{A-11}$$

To complete the proof, we have to bound the commutator term corresponding to the last part of (A-8). To do this, we use the fact that

$$\eta [T_{X^k}; (\eta \mathrm{Id} - \Delta)^{-1} \mathcal{R}_i \mathcal{R}_j] g = \eta [T_{X^k}; (\eta \mathrm{Id} - \Delta)^{-1}] \mathcal{R}_i \mathcal{R}_j g + \eta (\eta \mathrm{Id} - \Delta)^{-1} [T_{X^k}; \mathcal{R}_i \mathcal{R}_j] g.$$

To handle the last term, we just have to use (A-11) and the fact that $\eta(\eta \mathrm{Id} - \Delta)^{-1}$ maps L^p to itself (uniformly with respect to η). For the other term, we use that, by embedding,

$$\left\|\eta \partial_k [T_{X^k}; (\eta \mathrm{Id} - \Delta)^{-1}] \mathcal{R}_i \mathcal{R}_j g\right\|_{L^p} \le C\eta \left\|[T_{X^k}; (\eta \mathrm{Id} - \Delta)^{-1}] \mathcal{R}_i \mathcal{R}_j g\right\|_{B^1_{p,1}}.$$

Then, using the fact that the multiplier $\eta(\eta \mathrm{Id} - \Delta)^{-1}$ is in S^{-2} (uniformly with respect to $\eta \le 1$) and arguing as for bounding (A-4), one ends up with

$$\left\|\eta \partial_k [T_{X^k}; (\eta \mathrm{Id} - \Delta)^{-1}] \mathcal{R}_i \mathcal{R}_j g\right\|_{L^p} \le C \|\nabla X\|_{L^p} \|\mathcal{R}_i \mathcal{R}_j g\|_{B^{-1}_{\infty,1}} \le C \|\nabla X\|_{L^p} \|\mathcal{R}_i \mathcal{R}_j g\|_{B^0_{\infty,\infty}}.$$

Clearly, in the above computations, the low frequencies of g are not involved. Hence, we actually have, using that $\mathcal{R}_i$ maps $\dot{B}^0_{\infty,\infty}$ to itself and that $\dot{B}^0_{\infty,\infty} \hookrightarrow L^\infty$,

$$\big\| \eta \partial_k [T_{X^k}; (\eta \mathrm{Id} - \Delta)^{-1}] \mathcal{R}_i \mathcal{R}_j g \big\|_{L^p} \leq C \|\nabla X\|_{L^p} \|g\|_{L^\infty}.$$

Summing up all the above estimates concludes the proof of the lemma. □

Acknowledgements

Danchin was partially supported by the project INFAMIE (ANR-15-CE40-0011) operated by the French National Research Agency (ANR). Fanelli was partially supported by the LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, within the program "Investissement d'Avenir" (ANR-11-IDEX-0007), by the project BORDS (ANR-16-CE40-0027-01) and the programme "Oberwolfach Leibniz Fellows" by the Mathematisches Forschungsinstitut Oberwolfach in 2017. Paicu was supported by the project IFSMACS (ANR-15-CE40-0010) operated by the French National Research Agency (ANR).

References

[Bahouri et al. 2011] H. Bahouri, J.-Y. Chemin, and R. Danchin, *Fourier analysis and nonlinear partial differential equations*, Grundlehren der Math. Wissenschaften **343**, Springer, 2011. MR Zbl

[Bony 1981] J.-M. Bony, "Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires", *Ann. Sci. École Norm. Sup.* (4) **14**:2 (1981), 209–246. MR Zbl

[Chemin 1991] J.-Y. Chemin, "Sur le mouvement des particules d'un fluide parfait incompressible bidimensionnel", *Invent. Math.* **103**:3 (1991), 599–629. MR Zbl

[Chemin 1993] J.-Y. Chemin, "Persistance de structures géométriques dans les fluides incompressibles bidimensionnels", *Ann. Sci. École Norm. Sup.* (4) **26**:4 (1993), 517–542. MR Zbl

[Danchin 1997] R. Danchin, "Poches de tourbillon visqueuses", *J. Math. Pures Appl.* (9) **76**:7 (1997), 609–647. MR Zbl

[Danchin 1999] R. Danchin, "Persistance de structures géométriques et limite non visqueuse pour les fluides incompressibles en dimension quelconque", *Bull. Soc. Math. France* **127**:2 (1999), 179–227. MR Zbl

[Danchin 2001] R. Danchin, "Local theory in critical spaces for compressible viscous and heat-conductive gases", *Comm. Partial Differential Equations* **26**:7-8 (2001), 1183–1233. MR Zbl

[Danchin 2010] R. Danchin, "On the solvability of the compressible Navier–Stokes system in bounded domains", *Nonlinearity* **23**:2 (2010), 383–407. MR Zbl

[Danchin 2014] R. Danchin, "A Lagrangian approach for the compressible Navier–Stokes equations", *Ann. Inst. Fourier (Grenoble)* **64**:2 (2014), 753–791. MR Zbl

[Danchin and He 2016] R. Danchin and L. He, "The incompressible limit in L^p type critical spaces", *Math. Ann.* **366**:3-4 (2016), 1365–1402. MR Zbl

[Danchin and Mucha 2012] R. Danchin and P. B. Mucha, "A Lagrangian approach for the incompressible Navier–Stokes equations with variable density", *Comm. Pure Appl. Math.* **65**:10 (2012), 1458–1480. MR Zbl

[Danchin and Mucha 2017] R. Danchin and P. B. Mucha, "The incompressible Navier–Stokes equations in vacuum", preprint, 2017. To appear in *Comm. Pure Appl. Math.* arXiv

[Danchin and Zhang 2014] R. Danchin and P. Zhang, "Inhomogeneous Navier–Stokes equations in the half-space, with only bounded density", *J. Funct. Anal.* **267**:7 (2014), 2371–2436. MR Zbl

[Danchin and Zhang 2017] R. Danchin and X. Zhang, "On the persistence of Hölder regular patches of density for the inhomogeneous Navier–Stokes equations", *J. Éc. Polytech. Math.* **4** (2017), 781–811. MR Zbl

[Depauw 1999] N. Depauw, "Poche de tourbillon pour Euler 2D dans un ouvert à bord", *J. Math. Pures Appl.* (9) **78**:3 (1999), 313–351. MR Zbl

[Fanelli 2012] F. Fanelli, "Conservation of geometric structures for non-homogeneous inviscid incompressible fluids", *Comm. Partial Differential Equations* **37**:9 (2012), 1553–1595. MR Zbl

[Gamblin and Saint Raymond 1995] P. Gamblin and X. Saint Raymond, "On three-dimensional vortex patches", *Bull. Soc. Math. France* **123**:3 (1995), 375–424. MR Zbl

[Haspot 2011] B. Haspot, "Existence of global strong solutions in critical spaces for barotropic viscous fluids", *Arch. Ration. Mech. Anal.* **202**:2 (2011), 427–460. 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

[Hoff 1991] D. Hoff, "Discontinuous solutions of the Navier–Stokes equations for compressible flow", *Arch. Ration. Mech. Anal.* **114**:1 (1991), 15–46. MR Zbl

[Hoff 1995] D. Hoff, "Global solutions of the Navier–Stokes equations for multidimensional compressible flow with discontinuous initial data", *J. Differential Equations* **120**:1 (1995), 215–254. MR Zbl

[Hoff 2002] D. Hoff, "Dynamics of singularity surfaces for compressible, viscous flows in two space dimensions", *Comm. Pure Appl. Math.* **55**:11 (2002), 1365–1407. MR Zbl

[Hoff 2006] D. Hoff, "Uniqueness of weak solutions of the Navier–Stokes equations of multidimensional, compressible flow", *SIAM J. Math. Anal.* **37**:6 (2006), 1742–1760. MR Zbl

[Huang et al. 2013] J. Huang, M. Paicu, and P. Zhang, "Global well-posedness of incompressible inhomogeneous fluid systems with bounded density or non-Lipschitz velocity", *Arch. Ration. Mech. Anal.* **209**:2 (2013), 631–682. MR Zbl

[Lemarié-Rieusset 2002] P. G. Lemarié-Rieusset, *Recent developments in the Navier–Stokes problem*, Chapman & Hall/CRC Res. Notes in Math. **431**, Chapman & Hall/CRC, Boca Raton, FL, 2002. MR Zbl

[Liao and Zhang 2016] X. Liao and P. Zhang, "On the global regularity of the two-dimensional density patch for inhomogeneous incompressible viscous flow", *Arch. Ration. Mech. Anal.* **220**:3 (2016), 937–981. MR Zbl

[Liao and Zhang 2019] X. Liao and P. Zhang, "Global regularity of 2-D density patches for viscous inhomogeneous incompressible flow with general density: low regularity case", *Comm. Pure Appl. Math.* **72**:4 (2019), 835–884. MR Zbl

[Lions 1998] P.-L. Lions, *Mathematical topics in fluid mechanics, II: Compressible models*, Oxford Lecture Series in Math. Appl. **10**, Oxford Univ. Press, 1998. MR Zbl

[Mucha 2001] P. B. Mucha, "On almost global existence for the Cauchy problem for compressible Navier–Stokes equations in the L_p-framework", *J. Appl. Anal.* **7**:1 (2001), 61–79. MR Zbl

[Nash 1962] J. Nash, "Le problème de Cauchy pour les équations différentielles d'un fluide général", *Bull. Soc. Math. France* **90** (1962), 487–497. MR Zbl

[Paicu and Zhang 2017] M. Paicu and P. Zhang, "Striated regularity of 2-D inhomogeneous incompressible Navier–Stokes system with variable viscosity", preprint, 2017. arXiv

[Runst and Sickel 1996] T. Runst and W. Sickel, *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, de Gruyter Series in Nonlin. Anal. Appl. **3**, de Gruyter, Berlin, 1996. MR Zbl

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Math. Series **43**, Princeton Univ. Press, 1993. MR Zbl

Received 24 Apr 2018. Accepted 30 Nov 2018.

RAPHAËL DANCHIN: raphael.danchin@u-pec.fr
LAMA UMR 8050, Université Paris-Est Créteil Val-de-Marne, Créteil, France

FRANCESCO FANELLI: fanelli@math.univ-lyon1.fr
Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, Villeurbanne, France

MARIUS PAICU: marius.paicu@math.u-bordeaux.fr
Institut de Mathématiques de Bordeaux, Université de Bordeaux, Talence, France

mathematical sciences publishers msp

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.org/apde.

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 in APDE are usually in English, 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 make any reference to the bibliography. 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 subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should 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 details about how your graphics were generated.

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.

ANALYSIS & PDE

Volume 13 No. 1 2020

Absence of Cartan subalgebras for right-angled Hecke von Neumann algebras 1
MARTIJN CASPERS

A vector field method for radiating black hole spacetimes 29
JESÚS OLIVER and JACOB STERBENZ

Stable ODE-type blowup for some quasilinear wave equations with derivative-quadratic nonlinearities 93
JARED SPECK

Asymptotic expansions of fundamental solutions in parabolic homogenization 147
JUN GENG and ZHONGWEI SHEN

Capillary surfaces arising in singular perturbation problems 171
ARAM L. KARAKHANYAN

A spiral interface with positive Alt–Caffarelli–Friedman limit at the origin 201
MARK ALLEN and DENNIS KRIVENTSOV

Infinite-time blow-up for the 3-dimensional energy-critical heat equation 215
MANUEL DEL PINO, MONICA MUSSO and JUNCHENG WEI

A well-posedness result for viscous compressible fluids with only bounded density 275
RAPHAËL DANCHIN, FRANCESCO FANELLI and MARIUS PAICU