# ANALYSIS & PDE

Volume 18 No. 8 2025

msp

# Analysis & PDE

msp.org/apde


EDITORS-IN-CHIEF

| | |
|---|---|
| Anna L. Mazzucato | Penn State University, USA<br>alm24@psu.edu |
| Clément Mouhot | Cambridge University, UK<br>c.mouhot@dpmms.cam.ac.uk |

BOARD OF EDITORS

| | | | |
|---|---|---|---|
| Massimiliano Berti | Scuola Intern. Sup. di Studi Avanzati, Italy<br>berti@sissa.it | William Minicozzi II | Johns Hopkins University, USA<br>minicozz@math.jhu.edu |
| Zbigniew Błocki | Uniwersytet Jagielloński, Poland<br>zbigniew.blocki@uj.edu.pl | Werner Müller | Universität Bonn, Germany<br>mueller@math.uni-bonn.de |
| Charles Fefferman | Princeton University, USA<br>cf@math.princeton.edu | Igor Rodnianski | Princeton University, USA<br>irod@math.princeton.edu |
| David Gérard-Varet | Université de Paris, France<br>david.gerard-varet@imj-prg.fr | Yum-Tong Siu | Harvard University, USA<br>siu@math.harvard.edu |
| Colin Guillarmou | Université Paris-Saclay, France<br>colin.guillarmou@universite-paris-saclay.fr | Terence Tao | University of California, Los Angeles, USA<br>tao@math.ucla.edu |
| Ursula Hamenstaedt | Universität Bonn, Germany<br>ursula@math.uni-bonn.de | Michael E. Taylor | Univ. of North Carolina, Chapel Hill, USA<br>met@math.unc.edu |
| Peter Hintz | ETH Zurich, Switzerland<br>peter.hintz@math.ethz.ch | Gunther Uhlmann | University of Washington, USA<br>gunther@math.washington.edu |
| Vadim Kaloshin | Institute of Science and Technology, Austria<br>vadim.kaloshin@gmail.com | András Vasy | Stanford University, USA<br>andras@math.stanford.edu |
| Izabella Laba | University of British Columbia, Canada<br>ilaba@math.ubc.ca | Dan Virgil Voiculescu | University of California, Berkeley, USA<br>dvv@math.berkeley.edu |
| Richard B. Melrose | Massachussets Inst. of Tech., USA<br>rbm@math.mit.edu | Jim Wright | University of Edinburgh, UK<br>j.r.wright@ed.ac.uk |
| Frank Merle | Université de Cergy-Pontoise, France<br>merle@ihes.fr | Maciej Zworski | University of California, Berkeley, USA<br>zworski@math.berkeley.edu |

**PRODUCTION**

production@msp.org

Silvio Levy, Scientific Editor



Cover image: Eric J. Heller: “Linear Ramp”

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

The subscription price for 2025 is US $475/year for the electronic version, and $735/year (+$70, 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.

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/



© 2025 Mathematical Sciences Publishers

ANALYSIS AND PDE
Vol. 18 (2025), No. 8, pp. 1835–1874
DOI: 10.2140/apde.2025.18.1835

# UNIFORM CONTRACTIVITY OF THE FISHER INFINITESIMAL MODEL WITH STRONGLY CONVEX SELECTION

VINCENT CALVEZ, DAVID POYATO AND FILIPPO SANTAMBROGIO

The Fisher infinitesimal model is a classical model of phenotypic trait inheritance in quantitative genetics. Here, we prove that it encompasses a remarkable convexity structure which is compatible with a selection function having a convex shape. It yields uniform contractivity along the flow, as measured by an $L^\infty$ version of the Fisher information. It induces in turn asynchronous exponential growth of solutions, associated with a well-defined, log-concave, equilibrium distribution. Although the equation is nonlinear and nonconservative, our result shares some similarities with the Bakry–Emery approach to the exponential convergence of solutions to the Fokker–Planck equation with a convex potential. Indeed, the contraction takes place at the level of the Fisher information. Moreover, the key lemma for proving contraction involves the Wasserstein distance $W_\infty$ between two probability distributions of a (dual) backward-in-time process, and it is inspired by a maximum principle by Caffarelli for the Monge–Ampère equation.

## 1. Introduction

Let us consider the nonlinear model

$$F_n = \mathcal{T}[F_{n-1}], \quad n \in \mathbb{N},\ x \in \mathbb{R}, \tag{1-1}$$

describing the evolution of the distribution $F_n = F_n(x)$ of a one-dimensional trait $x \in \mathbb{R}$, subject to sexual reproduction and the effect of selection at each generation. The operator $\mathcal{T}$ above is defined by

$$\mathcal{T}[F](x) := e^{-m(x)}\mathcal{B}[F](x), \qquad x \in \mathbb{R}, \tag{1-2}$$

$$\mathcal{B}[F](x) := \iint_{\mathbb{R}^2} G\left(x - \frac{x_1 + x_2}{2}\right) F(x_1)\frac{F(x_2)}{\|F\|_{L^1}}\, dx_1\, dx_2, \quad x \in \mathbb{R}, \tag{1-3}$$

for any $F \in L^1_+(\mathbb{R}) \setminus \{0\}$. On the one hand, the operator $\mathcal{B}$ describes the distribution of traits of descendants of the previous generation $F_{n-1}$, arising as recombination of parental traits in agreement with *Fisher's infinitesimal model* [1919], which is a classical model in quantitative genetics; see also [Barton et al. 2017]. Accordingly, the mixing kernel $G$ is set to a centered Gaussian distribution with unit *segregation variance* without loss of generality, namely

$$G(x) := \frac{1}{(2\pi)^{1/2}} e^{-x^2/2}, \quad x \in \mathbb{R}. \tag{1-4}$$

*MSC2020:* primary 35B40, 35P30; secondary 35Q92, 47G20, 92D15.
*Keywords:* integrodifferential equations, asymptotic behavior, nonlinear spectral theory, quantitative genetics, Monge–Ampère equation, maximum principle.

© 2025 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

On the other hand, the trait-dependent mortality function $m = m(x) \geq 0$ represents the effect of selection on the population, which acts multiplicatively over the descendants. In other words, the multiplicative factor $e^{-m(x)}$ in (1-2) represents the survival probability to the next generation of individuals having the trait $x$. We note that the time-discrete generations $n \in \mathbb{N}$ are assumed nonoverlapping since, altogether, $F_n$ describes the distribution of those offspring of $F_{n-1}$ having survived after the selection step, and then different generations do not get mixed; see [Calvez et al. 2024] for further insight.

As the model is tracking only one trait distribution, it applies either when individuals are hermaphroditic, or when the traits are equally distributed between male and female individuals within the population. We refer to [Barton et al. 2017] for a comprehensive presentation of the model, its derivation and its limitations.

The goal of this paper is to extend the studies initiated in [Calvez et al. 2024] to a broader class of selection functions. Specifically, when $m$ is a strongly convex function we prove *asynchronous exponential growth* [Webb 1987] of solutions to (1-1). In other words, we derive quantitative rates for the relaxation of the solutions $\{F_n\}_{n\in\mathbb{N}}$ of (1-1) to a strongly log-concave quasiequilibrium of the form $\boldsymbol{\lambda}^n \boldsymbol{F}$, where $\boldsymbol{\lambda} > 0$ and $\boldsymbol{F} \in L^1(\mathbb{R}) \cap \mathcal{P}(\mathbb{R})$ is an appropriate probability density. The fact that the quasiequilibrium is strongly log-concave is crucial in our approach and will be present throughout the paper.

**Definition 1.1** (log-concavity). Consider any nonnegative function $F = e^{-V} : \mathbb{R}^d \to \mathbb{R}_+$:

(i) $F$ is said to be log-concave when $V$ is a convex function.

(ii) $F$ is said to be strongly log-concave with log-concavity parameter $\gamma > 0$ (or $\gamma$-log-concave) when $V$ is a strongly convex function with convexity parameter $\gamma$ (or $\gamma$-convex).

When the potential function $V$ is in $C^2(\mathbb{R}^d)$, we can equivalently formulate log-concavity in terms of second-order derivatives. Namely, $F$ is log-concave when $D^2 V \geq 0$, and $F$ is $\gamma$-log-concave when $D^2 V \geq \gamma I_d$.

We remark that in order for an ansatz of the form $F_n(x) = \lambda^n F(x)$ to define a solution to (1-1), we need that the pair $(\lambda, F)$ solves the nonlinear eigenproblem

$$\begin{aligned} \lambda F = \mathcal{T}[F], \quad & x \in \mathbb{R}, \\ F \geq 0, \quad & \int_{\mathbb{R}} F(x)\,dx = 1. \end{aligned} \tag{1-5}$$

Hence, the possible quasiequilibria are to be found as solutions to (1-5). Note that contrarily to the special quadratic regime treated in [Calvez et al. 2024], the Gaussian structure can no longer be exploited and, in particular, the existence of solutions to (1-5) is unclear. Indeed, the above nonlinear integral operator is 1-homogeneous but nonmonotone, and therefore the Krein–Rutman theorem [Mahadevan 2007] cannot be applied as it has been done in other (usually linear) problems in population dynamics [Berestycki et al. 2016; Li et al. 2017]. Hence, the study of the nonlinear evolution problem (1-1) and the nonlinear eigenproblem (1-5) requires innovative ideas.

Throughout this paper, we address jointly the following two problems: (i) existence of a strongly log-concave solution $(\boldsymbol{\lambda}, \boldsymbol{F})$ to (1-5), and (ii) quantitative relaxation of the solutions to (1-1) towards the quasiequilibrium $\boldsymbol{\lambda}^n \boldsymbol{F}$. We make the crucial hypothesis that $m$ is a strongly convex function,

$$m'' \geq \alpha \quad \text{for some } \alpha > 0. \tag{H1}$$

The function $m$ necessarily reaches its minimum value over $\mathbb{R}$. For convenience, we assume the following additional hypothesis without loss of generality:

$$m \geq 0 \quad \text{and} \quad m(0) = 0. \tag{H2}$$

The $L^\infty$ relative Fisher information $\mathcal{I}_\infty$ plays a pivotal role in our analysis, as it measures the contractivity along the flow (see methodological notes below). It is defined as follows, for a pair of functions $P, Q \in L^1_+(\mathbb{R}) \cap C^1(\mathbb{R})$:

$$\mathcal{I}_\infty(P \,\|\, Q) := \left\| \frac{d}{dx}\left(\log \frac{P}{Q}\right) \right\|_{L^\infty}. \tag{1-6}$$

**Theorem 1.2.** *Let $m \in C^2(\mathbb{R}^d)$ satisfy* (H1)–(H2). *Then, the following statements hold true*:

(i) (existence of quasiequilibrium) *There is at least one solution* $(\boldsymbol{\lambda}, \boldsymbol{F})$ *to* (1-5). *In addition, $\boldsymbol{F} = e^{-\boldsymbol{V}} \in L^1_+(\mathbb{R}) \cap C^\infty(\mathbb{R})$ is $\beta$-log-concave, where $\beta > \frac{1}{2}$ is uniquely defined by the relationship*

$$\beta = \alpha + \frac{2\beta}{1 + 2\beta}. \tag{1-7}$$

*Moreover*, $(\boldsymbol{\lambda}, \boldsymbol{F})$ *is the unique solution to* (1-5) *among all pairs* $(\lambda, F)$ *such that*

$$\frac{d}{dx}\left(\log \frac{F}{\boldsymbol{F}}\right) \in L^\infty(\mathbb{R}). \tag{1-8}$$

(ii) (one-step contraction) *Consider any $F_0 \in L^1_+(\mathbb{R}) \cap C^1(\mathbb{R})$ such that*

$$\frac{d}{dx}\left(\log \frac{F_0}{\boldsymbol{F}}\right) \in L^\infty(\mathbb{R}), \tag{H3}$$

*and let $\{F_n\}_{n\in\mathbb{N}}$ be the solution to* (1-1) *issued at $F_0$. Then, we have*

$$\mathcal{I}_\infty(F_n \,\|\, \boldsymbol{F}) \leq \frac{2}{1 + 2\beta}\, \mathcal{I}_\infty(F_{n-1} \,\|\, \boldsymbol{F}) \tag{1-9}$$

*for any $n \in \mathbb{N}$.*

(iii) (asynchronous exponential growth) *Consider any $F_0 \in L^1_+(\mathbb{R}) \cap C^1(\mathbb{R})$ satisfying the assumption* (H3) *above*, *and let $\{F_n\}_{n\in\mathbb{N}}$ be the solution to* (1-1) *issued at $F_0$. Then, we have*

$$\left| \frac{\|F_n\|_{L^1}}{\|F_{n-1}\|_{L^1}} - \boldsymbol{\lambda} \right| \leq C \left( \frac{2}{1 + 2\beta} \right)^n, \tag{1-10}$$

$$\mathcal{D}_{\mathrm{KL}}\left( \frac{F_n}{\|F_n\|_{L^1}} \,\middle\|\, \boldsymbol{F} \right) \leq C \left( \frac{2}{1 + 2\beta} \right)^{2n} \tag{1-11}$$

*for every $n \in \mathbb{N}$, where $C > 0$ is a explicit constant depending on $F_0$, and $\mathcal{D}_{\mathrm{KL}}$ is the Kullback–Leibler divergence (or relative entropy), that is,*

$$\mathcal{D}_{\mathrm{KL}}(P\,\|\,Q) := \int_{\mathbb{R}} \log\left(\frac{P(x)}{Q(x)}\right) P(x)\,dx, \quad P, Q \in L^1_+(\mathbb{R}) \cap \mathcal{P}(\mathbb{R}). \tag{1-12}$$

**Remark 1.3** (case of quadratic selection). For quadratic selection $m(x) = \frac{1}{2}\alpha\,|x|^2$, we have that $m$ satisfies the hypotheses (H1)–(H2) in Theorem 1.2, and then our new result applies. Such a special case was studied in detail in [Calvez et al. 2024], where in particular it was proven that there is a unique eigenpair $(\boldsymbol{\lambda}, \boldsymbol{F})$ of (1-5), which involves a Gaussian eigenfunction $\boldsymbol{F}(x) = G_{0,\sigma^2}(x)$ with variance $\sigma^2 > 0$ satisfying

$$\frac{1}{\sigma^2} = \alpha + \frac{1}{1+\sigma^2/2}. \tag{1-13}$$

In particular, $\boldsymbol{F}$ is $(1/\sigma^2)$-log-concave (see Definition 1.1), which is compatible with our new result in view of the identity $\sigma^2 = \beta^{-1}$ stemming from (1-7) and (1-13). Furthermore, the contraction factor in (1-9) predicted by Theorem 1.2 also recovers the one obtained in [Calvez et al. 2024] for quadratic selection. Specifically,

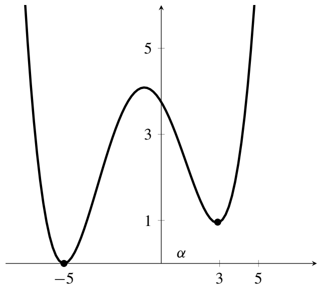
$$\frac{2}{1+2\beta} = \frac{(3+2\alpha) - \sqrt{(3+2\alpha)^2 - 8}}{2},$$

which agrees precisely with the contraction factor found in [Calvez et al. 2024, Lemma 6.3].

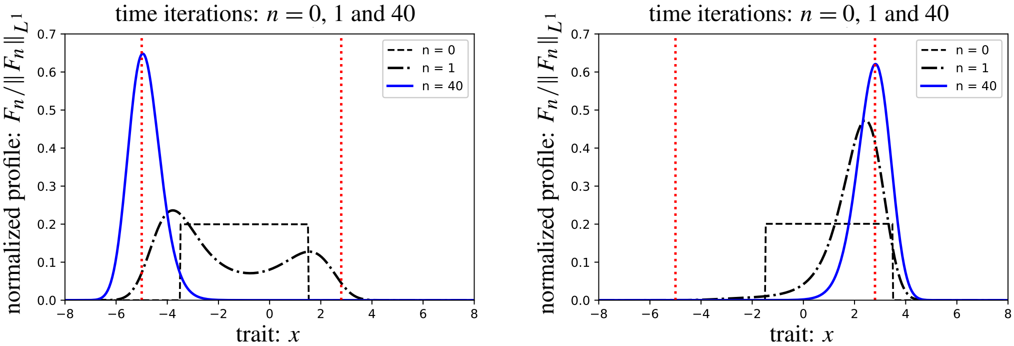
**Remark 1.4** (close-to-equilibrium initial data). In contrast with [Calvez et al. 2024], where the above framework was restricted to $m(x) = \frac{1}{2}\alpha\,|x|^2$ but generic $F_0 \in \mathcal{M}_+(\mathbb{R})$, Theorem 1.2 applies to a broader class of selection functions satisfying (H1)–(H2) at the cost of restricting to initial data fulfilling the hypothesis (H3). Specifically, such a condition imposes a precise behavior of the tails of $F_0$, which must be very close to those of the eigenfunction $\boldsymbol{F}$ (in particular, two Gaussian initial distributions should have the same variance).

**Remark 1.5** (conditional uniqueness). Another difference with [Calvez et al. 2024] is that the current approach does not guarantee global uniqueness of solutions to the eigenproblem (1-5), but only within the class of eigenpairs satisfying (1-8). Nevertheless, we conjecture that global uniqueness holds true, as in the quadratic case $m(x) = \frac{1}{2}\alpha\,|x|^2$. Proving global uniqueness would require a careful control of the behavior at infinity, in the spirit of [Calvez et al. 2024], which is beyond the scope of this paper.

**Remark 1.6** (on the convexity assumption). The convexity assumption (H1) ensures that $m$ must have a unique minimum. It implies that the quasiequilibrium $\boldsymbol{F}$ obtained in Theorem 1.2 is log-concave, as a consequence of the Prékopa–Leindler inequality. In the presence of multiple local minima of $m$, it was proven in [Calvez et al. 2019, Corollary 1.5] that several quasiequilibria could coexist in the time-continuous version of (1-1) provided that the variance of kernel (1-4) is small enough (in original units). That is, in the case of nonconvex $m$ there is evidence that the generalized eigenproblem (1-5) may admit nonunique solutions, in contrast with general conclusions of the Krein–Rutman theory in the linear case. This is illustrated by numerical simulations shown in Figure 1, where two different quasiequilibria (one of them bimodal) are found numerically if $m$ has two minima. A similar behavior can be observed




(a) Double-well selection function.




(b) Nonuniqueness of quasiequilibria for the double-well selection function.

**Figure 1.** (a) Double-well selection function $m(x) = 0.015((x-3)^2+1)(x+5)^2$ used in the simulations. (b) Time-evolution of the normalized profiles $F_n/\|F_n\|_{L^1}$ up to generation $n = 40$ (solid line) for two different choices of initial datum $F_0$. On the left, $F_0 = \mathbb{1}_{[-3.5,1.5]}$ leads to concentration near the left (globally) optimal trait. On the right, $F_0 = \mathbb{1}_{[-1.5,3.5]}$ leads to concentration near the right (locally) optimal trait.

in a population adapting to a heterogeneous, patchy environment, when each patch is associated with a different optimal trait [Dekens 2022]. The same conclusions also hold for the (continuous) time-marching problem in [Raoul 2021; Patout 2023; Guerand et al. 2025].

**Remark 1.7** (log-concavity and contraction factor). For any $\alpha > 0$, we have that the log-concavity parameter $\beta$ in (1-7) and the corresponding contraction factor $\frac{2}{1+2\beta}$ in (1-9) satisfy the properties

$$\alpha \searrow 0 \quad \Longrightarrow \quad \beta \searrow \frac{1}{2} \text{ and } \frac{2}{1+2\beta} \nearrow 1,$$
$$\alpha \nearrow \infty \quad \Longrightarrow \quad \beta \nearrow \infty \text{ and } \frac{2}{1+2\beta} \searrow 0.$$

See Figure 2. In particular, we have genuine contraction in (1-9) since $0 < \frac{2}{1+2\beta} < 1$ for every $\alpha > 0$.

**Remark 1.8** (one-dimensional traits). In this paper we restrict to one-dimensional traits, but note that an analogous version of (1-1) and (1-5) makes sense in higher dimensions yet. In fact, these were studied in [Calvez et al. 2024] for quadratic selection functions. However, a higher-dimensional version of our result for generic strongly convex selection function would require some nontrivial improvements of the

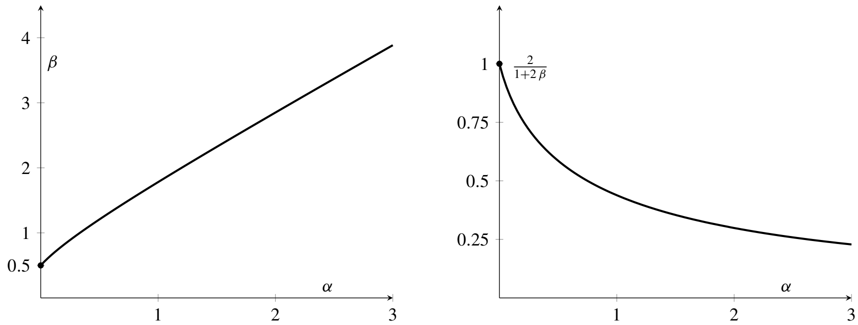



**Figure 2.** Plot of the log-concavity parameter $\beta$ of the eigenfunction $\boldsymbol{F}$ (left) and the contraction parameter $\frac{2}{1+2\beta}$ in Theorem 1.2 as a function of $\alpha$ (right).

present methods. Just to emphasize some nontrivial obstructions, we remark that our approach exploits a maximum principle for the Monge–Ampère equation in convex but not uniformly convex domains, as described below. In this setting, it is not even clear why the standard elliptic regularity should hold up to the boundary, as in the seminal work [Caffarelli 1996]. In two-dimensional domains with special symmetries, this theory has been developed recently in [Jhaveri 2019], but a higher-dimensional extension would require further work which goes beyond the scope of this paper. The extension to any dimension was achieved in [Khudiakova et al. 2024], which was released during the time of revision of the present work.

***Bibliographical notes.*** This work can be viewed as another step in using optimal transportation tools for nonconservative problems arising in biology. The connection between the Fisher infinitesimal model and the $L^2$ Wasserstein distance was spotted by G. Raoul [2017] (see also [Mirrahimi and Raoul 2013] for similar results in a different context of protein exchanges between cells). In fact, when there is no selection (that is, $m \equiv \text{const.}$), the operator $\mathcal{T}$ is nonexpansive for the latter distance. Contraction cannot be expected because of translational invariance. Nevertheless, it is contractive with rate $1/\sqrt{2}$ in the class of distributions having the same center of mass (the latter being preserved by the flow) [Raoul 2017, Theorem 4.1 and Corollary 4.2]. This remarkable structure was further exploited by G. Raoul [2021] in a perturbative setting, when selection is small (in amplitude), and restricted to a compact interval ($m$ is constant beyond a certain range). More precisely, G. Raoul proved that the dynamics is well captured by some averaged quantities ("moments") of the Gaussian distribution coupled with the selection function, provided that the initial data is well-prepared, in the basin of attraction of the stationary state, and the amplitude of selection is small enough. For that purpose, he carefully established that the contraction issued from the infinitesimal operator was robust enough to dominate detrimental effects due to selection. Note that the later references consider overlapping generations, that is, a continuous-in-time rather than discrete dynamics. However, some fruitful analogy can be drawn between the results and methodology.

In parallel, the regime of small segregation variance (when $G$ (1-4) has variance $\varepsilon^2$ and $\varepsilon$ is small enough) was investigated by [Calvez et al. 2019; Patout 2023] in another perturbative setting, without exploiting the Wasserstein metric structure. This methodology built upon the seminal works on vanishing viscosity limits associated with linear (asexual) modes of reproduction in quantitative genetics models [Diekmann et al. 2005; Perthame and Barles 2008; Barles et al. 2009]. Interestingly, it was proven in

[Calvez et al. 2019] that the problem (1-5) lacks uniqueness in full generality. More precisely, it was possible to build a solution to (1-5) centered in the vicinity of any local minimum of $m$, provided that the selection value at the local minimum is close enough to the global minimum. This result gives a clear separation with linear, order-preserving operators (and nonlinear extensions [Mahadevan 2007; Nussbaum 1988]) for which (1-5) genuinely admits a unique solution (under standard irreducibility assumptions); see Remark 1.6. The Cauchy problem initialized with some concentrated initial data was further investigated in [Patout 2023] (in a multiplicative perturbative approach) and more recently in [Guerand et al. 2025] (in a moment-based approach), still in the regime of small segregation variance. The case of zero segregation variance was the subject of the recent [Frouvelle and Taing 2025].

Heuristically, uniqueness of the (nonlinear) eigenpair $(\boldsymbol{\lambda}, \boldsymbol{F})$ is rather clear when the selection function $m$ is convex, and [Calvez et al. 2024] was a first contribution in this direction, restricted to $m(x) = \frac{1}{2}\alpha\,|x|^2$. By exploiting the quadratic structure of the operator $\mathcal{T}$ in (1-2) (which involves products and convolutions by Gaussian density functions), it was possible to prove asynchronous exponential growth towards the explicit Gaussian distribution of equilibrium $\boldsymbol{F}$, starting from any initial configuration $F_0$. This was achieved by a careful study of the binary tree of ancestors, together with explicit change of variables in a high-dimensional integral, to prove a sort of concentration of measure estimates. More precisely, it was shown that the traits of the ancestors decorrelate sufficiently fast, backward in the tree, from the trait of the individual at generation $n$. This implies that the dependence of the trait distribution $F_n$ at generation $n$ upon the initial distribution $F_0$ diminishes exponentially fast. Asynchronous exponential growth is a consequence of this observation, which is a backward feature.

Last, but not least, let us mention that both the infinitesimal model (1-2), and the relative information (1-6) (or rather (1-18) below) date back to a couple of seminal works [Fisher 1919; 1922] respectively on seemingly different purposes; see [Stigler 2005] for a discussion.

***Methodological notes.*** In the present study, we push further the observations of [Calvez et al. 2024]. We identify a key mechanism ensuring a one-step contraction for the flow (1-1). This can be summarized roughly as follows:

> *For any two given individuals with traits $X$ and $X'$ respectively, the associated parental traits $(X_1, X_2)$ and $(X_1', X_2')$ are closer to each other than $X$ and $X'$ are, in some sense.*

See also [Garnier et al. 2023, Appendix F.2] for a visual explanation. To make sense of this contraction, we shall work with the $L^\infty$ Wasserstein distance, denoted by $W_\infty$ (in contrast with the $L^2$ Wasserstein distance). This naturally leads to estimates on the so-called $L^\infty$ relative Fisher information $\mathcal{I}_\infty$ (1-6) (in contrast with the ($L^2$) relative Fisher information $\mathcal{I}_2$, see (1-18) below). The core estimate (1-9) is forward-in-time, and it naturally arises as a dual estimate of a backward-in-time estimate analogous to the work in [Calvez et al. 2024].

***A forward-backward argument.*** We propose a short warm-up to this argument, which may help the reader follow our method (without details of the proofs). Indeed, one complication of our setting is that each individual has two parents, so that the dimension of the distribution doubles at each generation.

Nonetheless, the same methodology can be applied to the case of a single parent, which boils down to a *linear operator*. We thus consider, temporarily, the linear operator

$$\mathcal{A}[F](x) := e^{-m(x)} \int_{\mathbb{R}} G(x-y)F(y)\,dy, \quad x \in \mathbb{R}, \tag{1-14}$$

in place of the above nonlinear operator $\mathcal{T}$ in (1-2). In this simpler case, the Krein–Rutman theorem can be applied (at least formally), and there exists an eigenpair $(\boldsymbol{\lambda}, \boldsymbol{F})$ of the linear eigenproblem (1-5) with $\mathcal{T}$ replaced by $\mathcal{A}$. Now, consider any solution $\{F_n\}_{n\in\mathbb{N}}$ to the time-discrete problem (1-1) with $\mathcal{T}$ replaced again by the linear operator $\mathcal{A}$. We may introduce the associated relative distribution $u_n = F_n/(\boldsymbol{\lambda}^n \boldsymbol{F})$ to follow the trend of $F_n$ across generations. It satisfies the equation

$$u_n(x) = \frac{\int_{\mathbb{R}} G(x-y)u_{n-1}(y)\,\boldsymbol{F}(y)\,dy}{\int_{\mathbb{R}} G(x-z)\,\boldsymbol{F}(z)\,dz} = \int_{\mathbb{R}} \boldsymbol{P}(x;y)u_{n-1}(y)\,dy, \quad n \in \mathbb{N},\ x \in \mathbb{R},$$

where the $x$-dependent probability distribution function $\boldsymbol{P}(x;\cdot)$ is defined as

$$\boldsymbol{P}(x;y) = \frac{G(x-y)\,\boldsymbol{F}(y)}{\int_{\mathbb{R}} G(x-z)\,\boldsymbol{F}(z)\,dz}, \quad x, y \in \mathbb{R}, \tag{1-15}$$

and it can be interpreted as the transition probability from trait $y$ to trait $x$. The fact that it is a probability distribution function, $\int \boldsymbol{P}(x;y)\,dy = 1$, is immediate by the choice of the normalization, which is such that constant functions $u_n \equiv \text{const.}$ are invariant by the flow.

Next, it can be proven that, if $\boldsymbol{F}$ is strongly log-concave, then we have

$$W_\infty(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(x';\cdot)) \le \kappa |x-x'|, \tag{1-16}$$

where $\kappa \in (0,1)$ is related to the modulus of convexity of $\boldsymbol{V} = -\log \boldsymbol{F}$. By duality, this backward contraction estimate results in the forward estimate (see Lemma 2.4)

$$\left\| \frac{d}{dx}(\log u_n) \right\|_{L^\infty} \le \kappa \left\| \frac{d}{dx}(\log u_{n-1}) \right\|_{L^\infty},$$

which, by iteration and using the $L^\infty$ relative Fisher information, can be expressed as

$$\mathcal{I}_\infty(F_n \,\|\, \boldsymbol{F}) \le \kappa^n \,\mathcal{I}_\infty(F_0 \,\|\, \boldsymbol{F}). \tag{1-17}$$

As mentioned in Remark 1.8, the key estimate (1-16) is a consequence of the maximum principle on the Monge–Ampère equation for the optimal transportation plan between $\boldsymbol{P}(x;\cdot)$ and $\boldsymbol{P}(x';\cdot)$. Interestingly, this is an argument borrowed from the theory of conservative equations, whereas our problem is not. The trick is to match an individual to its ancestor, which is obviously a conservative process, backward-in-time.

***Analogy with the Bakry–Emery argument.*** There is some analogy between our results and the standard Bakry–Emery method for exponential relaxation towards equilibrium for the gradient flow of some displacement convex “entropy”, for instance, the Fokker–Planck equation with a convex potential [Bakry 1994; Arnold et al. 2001; Villani 2003; Bakry et al. 2014]. Indeed, from (1-9) (alternatively (1-17) in the linear case) we obtain exponential convergence on a quantity which is the $L^\infty$ analog of the usual ($L^2$) relative Fisher information,

$$I_2(P\,\|\,Q) := \int_{\mathbb{R}} \left| \frac{d}{dx}\left(\log \frac{P}{Q}\right)(x) \right|^2 P(x)\,dx. \tag{1-18}$$

Recall that, in the usual Bakry–Emery argument, the exponential convergence is established at the level of the dissipation of entropy, that is, the usual relative Fisher information [Villani 2003]. In turn, the exponential relaxation of the dissipation is intimately linked with the displacement convexity of the entropy functional (essentially because the gradient flow is differentiated, which leads to the second derivative of the entropy functional). In our argument, it is the convexity of $\boldsymbol{V} = -\log \boldsymbol{F}$ which induces the geometrical relaxation of the uniform relative Fisher information.

***Connection with another projective metric.*** The uniform relative Fisher information (1-6) may also be viewed as a kind of first-order version of the *Hilbert's projective distance* associated with the cone of nonnegative functions, that is,

$$\mathfrak{H}(P, Q) := \operatorname{osc}\left(\log \frac{P}{Q}\right) \equiv \sup_{x\in\mathbb{R}} \log \frac{P(x)}{Q(x)} - \inf_{x\in\mathbb{R}} \log \frac{P(x)}{Q(x)}.$$

The latter distance is well-suited for the analysis of 1-positively homogeneous, order-preserving operators [Nussbaum 1988]. An obvious reason is the projective character of that metric [Nussbaum 1994], which makes it insensitive to the exponential growth (or decay) $\mathcal{O}(\boldsymbol{\lambda}^n)$. This character is also shared by $\mathcal{I}_\infty$ (in contrast with $\mathcal{I}_2$).

***A linear argument, even in the nonlinear case.*** The previous discussion focused on the linear operator (1-14) for the sake of clarity. Interestingly, the nonlinear case under study (1-2) also involves a linear argument when formulated backward in time. Similarly, define the relative distribution $u_n = F_n/(\boldsymbol{\lambda}^n \boldsymbol{F})$, where the pair $(\boldsymbol{\lambda}, \boldsymbol{F})$ is the strongly log-concave solution to (1-5) from part (i) in Theorem 1.2. Then, $u_n$ satisfies the forward-in-time nonlinear problem

$$u_n(x) = \frac{1}{\|u_{n-1}\boldsymbol{F}\|_{L^1}} \iint_{\mathbb{R}^d} \boldsymbol{P}(x; x_1, x_2) u_{n-1}(x_1) u_{n-1}(x_2)\,dx_1\,dx_2, \quad n \in \mathbb{N},\ x \in \mathbb{R}, \tag{1-19}$$

where the function $\boldsymbol{P}(x; x_1, x_2)$ is explicitly defined as

$$\boldsymbol{P}(x; x_1, x_2) = \frac{G\big(x - \frac{1}{2}(x_1 + x_2)\big)\,\boldsymbol{F}(x_1)\,\boldsymbol{F}(x_2)}{\iint_{\mathbb{R}^2} G\big(x - \frac{1}{2}(x_1' + x_2')\big)\,\boldsymbol{F}(x_1')\,\boldsymbol{F}(x_2')\,dx_1'\,dx_2'}, \quad x \in \mathbb{R},\ (x_1, x_2) \in \mathbb{R}^2. \tag{1-20}$$

Since $\boldsymbol{P}$ is normalized with respect to the variables $(x_1, x_2)$, it can be regarded as a Markov kernel with source $x \in \mathbb{R}$ and target $(x_1, x_2) \in \mathbb{R}^2$ representing the probability of transitioning from the trait of the offspring $x$ to the traits of the parents $(x_1, x_2)$. In Lemma 2.6, we prove the very same contraction estimate as in (1-16) for the family of Markov kernels $\boldsymbol{P}$ indexed by its first variable $x$. The key difference is that this Markov kernel makes the transition between $u_n$ and $u_{n-1} \otimes u_{n-1}$ due to the joint distribution of parental traits (the nonlinearity, in fact). This is rescued by an appropriate tensorization property of the relative Fisher information, which is expressed in Lemma 2.4.

***A close-to-optimal result despite a nonoptimal argument.*** The rate of contraction $\frac{2}{1+2\beta}$ coincides with the optimal one in the quadratic case (see Remark 1.3). However, there is a nonoptimal step in the proof. Indeed, our key contraction estimate (1-16) is a consequence of the maximum principle on the Monge–Ampère equation satisfied by the Brenier transportation map between the joint distributions of the parental traits $(X_1, X_2)$ and $(X_1', X_2')$. There is some subtlety here to be noticed, as the contraction is set for the $L^\infty$ Wasserstein distance (maximum of the optimal transportation displacement), whereas the Brenier transportation map used in our argument is optimal for the $L^2$ Wasserstein distance. Nevertheless, in the quadratic case, the transportation map is simply a translation, so that it comes with the same cost, measured either in (weighted) $L^2$ or in $L^\infty$.

In the recent contribution [Khudiakova et al. 2024], the authors used a different approach based on Langevin dynamics to make the connection between the two joint distributions. Hence, they bypassed the use of the Brenier map. Their approach is much simpler, and it enabled them to extend the result readily to higher dimensions. These results were originally motivated by a computation in a previous version of our paper, where we obtained an upper bound on the displacement $\|T(x)-x\|_2$ for the Brenier map between a strongly log-concave density and a perturbation of it. In the current version, such an estimate cited by [Khudiakova et al. 2024] is not crucial, as the important one concerns the displacement $\|T(x)-x\|_1$ (see Sections 2.3 and 2.4) and interpolating $\ell_1$ estimates from $\ell_2$ ones worsens the coefficients (see Remark 3.1). We have moved the $\ell_2$ estimates to Appendix C for an easier readability. In [Khudiakova et al. 2024], the authors bypass this delicate issue of choosing $\ell_1$- rather than $\ell_2$-based distances by establishing some fruitful anisotropic version of our Corollary C.2.

***Organization of the paper.*** In Section 2 we provide a sketch of the proof of the one-step contraction property in Theorem 1.2(ii) under an additional technical condition. In Section 3 we derive the fundamental contraction property of the one-step transition probability of the problem under the $W_{\infty,1}$ Wasserstein distance (see definition below), thus removing the technical condition used in the sketch of proof of Section 2. In Section 4 we analyze a truncated version of the time-marching problem (1-5) to bounded intervals, which will be necessary in the next part. Section 5 focuses on proving the existence of strongly log-concave solutions of the nonlinear eigenproblem (1-5) as claimed in Theorem 1.2(i). In Section 6 we prove asymptotic exponential growth of (1-5) for restricted initial data (H3) as in Theorem 1.2(iii). Finally, Appendices A and B contain some technical results to alleviate the reading of the paper.

***Notation.*** • (vector norms) Throughout paper, $\mathbb{R}^d$ will be endowed with the various $\ell_q$ norms, namely, for any $z=(z_1,\dots,z_d)\in\mathbb{R}^d$ and any $1\le q\le\infty$ we define

$$\|z\|_q := \begin{cases} \left(\sum_{i=1}^d |z_i|^q\right)^{1/q} & \text{if } 1\le q<\infty, \\ \max_{1\le i\le d}|z_i| & \text{if } q=\infty. \end{cases} \tag{1-21}$$

The associated $\ell_2$ and $\ell_\infty$ open balls centered at 0 with radius $R>0$ are respectively denoted by

$$\begin{aligned} B_R &:= \{z\in\mathbb{R}^d : \|z\|_2 < R\}, \\ Q_R &:= \{z\in\mathbb{R}^d : \|z\|_\infty < R\}. \end{aligned} \tag{1-22}$$

• (characteristic function) Given any set $A \subset \mathbb{R}^d$, we will denote the associated characteristic function of convex analysis by $\chi_A : \mathbb{R}^d \to (-\infty, +\infty]$, which is the mapping defined by

$$\chi_A(z) := \begin{cases} 0 & \text{if } z \in A, \\ +\infty & \text{if } z \in \mathbb{R}^d \setminus A. \end{cases} \tag{1-23}$$

• (measure spaces) We denote by $\mathcal{M}(\mathbb{R}^d)$ the space of finite Radon measures, endowed with the total variation norm, and $\mathcal{M}^+(\mathbb{R}^d)$ represents the cone of nonnegative finite Radon measures. Similarly, $\mathcal{P}(\mathbb{R}^d)$ is the subspace of probability measures, endowed with the narrow topology except otherwise specified.

• (Wasserstein metrics) For any $1 \le p \le \infty$, we define the $L^p$ Wasserstein space

$$\mathcal{P}_p(\mathbb{R}^d) := \left\{ P \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |z|^p \, P(dz) < \infty \right\} \quad \text{if } 1 \le p < \infty,$$
$$\mathcal{P}_\infty(\mathbb{R}^d) := \{ P \in \mathcal{P}(\mathbb{R}^d) : \operatorname{supp} P \text{ is compact} \}.$$

Similarly, we consider the $L^p$ Wasserstein metric associated with the $\ell_q$ vector norm of $\mathbb{R}^d$. Specifically, for any $P, Q \in \mathcal{P}(\mathbb{R}^d)$ and any $1 \le p, q \le \infty$ we define

$$\begin{aligned} W_{p,q}(P, Q) &:= \Big( \inf_{\gamma \in \Gamma(P,Q)} \int_{\mathbb{R}^{2d}} \|z - \tilde{z}\|_q^p \, \gamma(dz, d\tilde{z}) \Big)^{1/p} \quad \text{if } 1 \le p < \infty, \\ W_{\infty,q}(P, Q) &:= \inf_{\gamma \in \Gamma(P,Q)} \gamma\text{-ess}\sup_{z,\tilde{z} \in \mathbb{R}^d} \|z - \tilde{z}\|_q, \end{aligned} \tag{1-24}$$

where $\Gamma(P, Q)$ is the family of transference plan $\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R})^d$ with marginals $P$ and $Q$. Whilst the $L^p$ Wasserstein distances could be infinitely valued over $\mathcal{P}(\mathbb{R}^d)$, note that they take finite values over $\mathcal{P}_p(\mathbb{R}^d)$ at least, although not exclusively. In particular, note that the $L^\infty$ Wasserstein distances take finite values over distributions $P$ and $Q$ that only differ on a space translation independently of their supports being compact or not. For this reason, throughout paper we shall not restrict to compactly supported distributions, but in all our computations the involved $L^\infty$ Wasserstein distances will take finite values, as it will become clear later in the proofs.

## 2. Proof of the one-step contraction property

For the reader's convenience, we provide first the main ingredients behind the proof of the fundamental one-step contraction property in Theorem 1.2(ii). Here, we shall assume that Theorem 1.2(i) holds true, i.e., there exists a $\beta$-log-concave solution $(\boldsymbol{\lambda}, \boldsymbol{F})$ to (1-5) with $\beta$ given by (1-7) (recall the precise notion of strong log-concavity in Definition 1.1). We remark that its use will be crucial in our following argument, but its proof is not apparent with regards to classical approaches based on the application of the Krein–Rutman theorem. For this reason, a major part of this paper is devoted to rigorously addressing this question, which will be introduced in full detail in Section 5 of this paper.

**2.1.** ***Sharp log-concavity parameter.*** First, we elaborate on the precise value of $\beta$ given in (1-7). Specifically, we prove that it amounts to the sharpest possible log-concavity parameter of a generic solution $(\lambda, F)$ to (1-5). To this end, it is worthwhile to note that the nonlinear operator $\mathcal{T}$ in (1-2) can be restated as the composition of a multiplicative operator and a double convolution operator, namely,

$$\mathcal{T}[F]=\frac{e^{-m}}{\|F\|_{L^1}}(G*\bar{F}*\bar{F}) \tag{2-1}$$

for every $F\in L^1_+(\mathbb{R})\setminus\{0\}$, where we define $\bar{F}(x):=2F(2x)$ for $x\in\mathbb{R}$. The starting point is to realize that strong log-concavity is stable under convolutions. This is a classical corollary of the celebrated Prékopa–Leindler inequality, which reads as follows (see [Saumard and Wellner 2014, Proposition 7.1] for further details).

**Lemma 2.1** (stability of log-concavity under convolutions). *Assume that $F_1, F_2\in L^1_+(\mathbb{R})$ satisfy that $F_i$ are $\gamma_i$-log-concave for some $\gamma_1,\gamma_2>0$. Then $F_1*F_2$ is also $\gamma$-log-concave for $\gamma>0$ given by*

$$\frac{1}{\gamma}=\frac{1}{\gamma_1}+\frac{1}{\gamma_2}.$$

Let us remark that the above result could be applied to any pair of Gaussian distributions $F_1$ and $F_2$ with respective variances $\sigma_1^2$ and $\sigma_2^2$ since they are in particular $\gamma_i$-log-concave with parameters $\gamma_i=1/\sigma_i^2$ for $i=1,2$. In doing so one finds that the above result is consistent with the classical fact that the convolution $F_1*F_2$ of two Gaussian distributions is again Gaussian with variance $\sigma^2=\sigma_1^2+\sigma_2^2$.

In addition, note that the mortality function $m$ has been chosen $\alpha$-convex by the hypothesis (H1) in Theorem 1.2, and then $e^{-m}$ is $\alpha$-log-concave. Since strong log-concavity is also preserved under multiplication, and $\bar{F}$ is $4\gamma$-log-concave whenever $F$ is $\gamma$-log-concave, we obtain that log-concavity must also be preserved under the full operator $\mathcal{T}$.

**Lemma 2.2** (stability of log-concavity under $\mathcal{T}$). *Assume that $F\in L^1_+(\mathbb{R})\setminus\{0\}$ is $\gamma$-log-concave for some $\gamma>0$. Then, $\mathcal{T}[F]$ is also $\delta$-log-concave for $\delta>0$ given by*

$$\delta=\alpha+\frac{2\gamma}{1+2\gamma}.$$

Thereby, log-concavity is preserved by the dynamics in (1-1), and we also obtain that the sharpest log-concavity coefficient of the eigenfunction $\boldsymbol{F}$ must be the one given in (1-7).

**Lemma 2.3** (propagation of log-concavity). (i) *Assume that $F_0\in L^1_+(\mathbb{R})\setminus\{0\}$ is $\beta_0$-log-concave for some $\beta_0>0$. Then, the solution $\{F_n\}_{n\in\mathbb{N}}$ to the evolution problem* (1-1) *satisfies that $F_n$ is $\beta_n$-log-concave for $\beta_n>0$ satisfying the recurrence*

$$\beta_n=\alpha+\frac{2\beta_{n-1}}{1+2\beta_{n-1}},\quad n\in\mathbb{N}. \tag{2-2}$$

(ii) *Assume that $(\lambda, F)$ is any solution to the nonlinear eigenproblem* (1-5) *and that $F$ is strongly log-concave. Then, $F$ is $\beta$-log-concave with $\beta$ given by* (1-7), *that is,*

$$\beta=\alpha+\frac{2\beta}{1+2\beta}.$$

*Proof.* Since (i) is clear by Lemma 2.2, we just prove (ii). Recall that for any solution $(\lambda, F)$ of (1-5) with $\gamma$-log-concave $F$, we can build $F_n(x)=\lambda^n F(x)$, which solves the evolution problem (1-1). Therefore, the above applied to $\{F_n\}_{n\in\mathbb{N}}$ shows that $F$ is $\beta_n$ log-concave for any $n \in \mathbb{N}$ with $\{\beta\}_{n\in\mathbb{N}}$ satisfying the recurrence (2-2) above and $\beta_0 = \gamma$. Since $\beta_n \to \beta$, then $F$ is also $\beta$-log-concave. □

**2.2.** ***The renormalized problem.*** We introduce a renormalized version of the evolution problem (1-1). Specifically, for any solution $\{F_n\}_{n\in\mathbb{N}}$ to (1-1) we renormalize by the strongly log-concave quasiequilibrium $\boldsymbol{\lambda}^n \boldsymbol{F}$ granted in Theorem 1.2(i). Namely, we set

$$u_n(x) := \frac{F_n(x)}{\boldsymbol{\lambda}^n \boldsymbol{F}(x)}, \quad n \in \mathbb{N},\ x \in \mathbb{R}. \tag{2-3}$$

By inspection, we obtain that $\{u_n\}_{n\in\mathbb{N}}$ must solve the evolution problem

$$u_n(x) = \frac{1}{\|u_{n-1}\boldsymbol{F}\|_{L^1}} \iint_{\mathbb{R}^2} \boldsymbol{P}(x; x_1, x_2) u_{n-1}(x_1) u_{n-1}(x_2)\, dx_1\, dx_2 \tag{2-4}$$

for any $x \in \mathbb{R}$, where $\boldsymbol{P}(x; x_1, x_2)$ is the *one-step transition probability* of transitioning from the parental traits $(x_1, x_2)$ to the descendant trait $x$. More, specifically, $\boldsymbol{P}(x;\cdot) \in L^1_+(\mathbb{R}^2) \cap \mathcal{P}(\mathbb{R}^2)$ is a probability density on two variables $(x_1, x_2)$ depending on the parameter $x \in \mathbb{R}$ which takes the form (recall the notation $\boldsymbol{F} = e^{-\boldsymbol{V}}$)

$$\begin{aligned} \boldsymbol{P}(x; x_1, x_2) &:= \frac{1}{\boldsymbol{Z}(x)} e^{-\boldsymbol{W}(x;x_1,x_2)}, \quad x \in \mathbb{R},\ (x_1, x_2) \in \mathbb{R}^2, \\ \boldsymbol{W}(x; x_1, x_2) &:= \tfrac{1}{2}\left|x - \tfrac{1}{2}(x_1 + x_2)\right|^2 + \boldsymbol{V}(x_1) + \boldsymbol{V}(x_2), \\ \boldsymbol{Z}(x) &:= \iint_{\mathbb{R}^2} e^{-\boldsymbol{W}(x;x_1,x_2)}\, dx_1\, dx_2. \end{aligned} \tag{2-5}$$

Inspired by our method in [Calvez et al. 2024], we plan to study the relaxation to zero of $\left\| \frac{d}{dx}(\log u_n) \right\|_{L^\infty}$ as $n$ grows. Nevertheless, contrarily to the aforementioned paper, we do not need to accumulate a large enough amount of generations in order to observe some ergodic behavior, but we rather find a precise contraction of such a quantity after a single step.

**2.3.** ***A nonlinear Kantorovich-type duality.*** Our new approach exploits a nice nonlinear version of a Kantorovich-type duality which relates the $L^\infty$ transport distance to the Lipschitz norm of the log of test functions. This nonlinear extension is reminiscent of the usual Kantorovich duality theorem, which relates the $L^1$ transport distance to the Lipschitz norm of test functions; see [Ambrosio et al. 2008, Theorem 6.1.1]. More specifically, we remark that the usual Kantorovich duality is fundamental in the linear setting to establish a general equivalence between the contraction of a forward semigroup under the Lipschitz norm, and the contraction of its backward (or dual) semigroup under the $L^1$ transport distance. We refer to [Kuwada 2010] for further extensions, yet in a linear setting. In our case, our nonlinear relation provides a method to derive contraction of a forward semigroup under the Lipschitz norms of the log of tests functions, once we know that there is contraction of the backward semigroup under a suitable $L^\infty$ transport distance. Interestingly, our nonlinear relation does not only apply to the linear setting, but also to our nonlinear setting. To the best of our knowledge, this relation appears to be new. Moreover, it does not represent an isolated example but there is a full family of related inequalities interpolating between the (classical) $L^1$ result and the (seemingly new) $L^\infty$ result, and which further adapt to $L^p$ transport distances; see Appendix A.

**Lemma 2.4** ($L^\infty$-type Kantorovich duality). *Consider the one-step transition from $u_0$ to $u_1$ in* (2-4), *where it is assumed that $u_0 \in C^1(\mathbb{R})$ with $u_0 > 0$ and $\frac{d}{dx}(\log u_0) \in L^\infty(\mathbb{R})$. Then, we have*

$$|\log u_1(x) - \log u_1(\tilde{x})| \le \left\| \frac{d}{dx}(\log u_0) \right\|_{L^\infty} W_{\infty,1}(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(\tilde{x};\cdot)) \tag{2-6}$$

*for any $x, \tilde{x} \in \mathbb{R}$. Here, the metric $W_{\infty,1}$ represents the $L^\infty$ Wasserstein distance associated with the $\ell_1$ norm*; *see* (1-24).

*Proof.* Set $x, \tilde{x} \in \mathbb{R}$ and assume that

$$W_{\infty,1}(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(\tilde{x};\cdot)) < \infty$$

(otherwise the inequality is obvious). Indeed, this will always be the case as we prove later in Section 3. Then, consider any $\gamma \in \Gamma(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(\tilde{x};\cdot))$ minimizing the $W_{\infty,1}$ transport distance (1-24) and note that

$$\begin{aligned}
u_1(x) &= \frac{1}{\|u_0 \boldsymbol{F}\|_{L^1}} \iint_{\mathbb{R}^2} u_0(x_1) u_0(x_2) \gamma(dx_1, dx_2, d\tilde{x}_1, d\tilde{x}_2) \\
&= \frac{1}{\|u_0 \boldsymbol{F}\|_{L^1}} \iint_{\mathbb{R}^2} \exp\big(\log u_0(x_1) - \log u_0(\tilde{x}_1) + \log u_0(x_2) - \log u_0(\tilde{x}_2)\big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \times u_0(\tilde{x}_1) u_0(\tilde{x}_2) \gamma(dx_1, dx_2, d\tilde{x}_1, d\tilde{x}_2) \\
&\le \frac{1}{\|u_0 \boldsymbol{F}\|_{L^1}} \iint_{\mathbb{R}^2} \exp\Big( \left\| \frac{d}{dx}(\log u_0) \right\|_{L^\infty} \|(x_1, x_2) - (\tilde{x}_1, \tilde{x}_2)\|_1 \Big) u_0(\tilde{x}_1) u_0(\tilde{x}_2) \gamma(dx_1, dx_2, d\tilde{x}_1, d\tilde{x}_2) \\
&\le \exp\Big( \left\| \frac{d}{dx}(\log u_0) \right\|_{L^\infty} W_{\infty,1}(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(\tilde{x};\cdot)) \Big) u_1(\tilde{x}),
\end{aligned}$$

where in the next-to-last line we have used the mean value theorem and in the last one we have exploited the fact that $\gamma$ is minimizer. Then, taking the logarithm at each side of the above inequality ends the proof. □

**Remark 2.5** (the choice of $\ell_1$ norm). We note that Lemma 2.4 is a particular instance of Proposition A.1 in Appendix A which can be recovered by setting $d_1 = 1$, $d_2 = 2$, $q = 1$ and

$$u(x_1, x_2) := u_0(x_1) u_0(x_2), \quad (x_1, x_2) \in \mathbb{R}^2.$$

However, the special choice $q = 1$ (that is $\ell_1$ norms) is apparently less clear at this stage since in fact choosing any other $1 \le q \le \infty$ would be possible in Proposition A.1 and it would yield more generally

$$|\log u_1(x) - \log u_1(\tilde{x})| \le 2^{1/q'} \left\| \frac{d}{dx}(\log u_0) \right\|_{L^\infty} W_{\infty,q}(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(\tilde{x};\cdot)) \tag{2-7}$$

for every $x, \tilde{x} \in \mathbb{R}$. Here, the metric $W_{\infty,q}$ represents the $L^\infty$ Wasserstein distance associated with the $\ell_q$ norm; see (1-24). By the natural relation between $\ell_1$ and $\ell_q$ vector norms, we infer that the above estimate (2-6) is sharper than (2-7), namely

$$W_{\infty,1}(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(\tilde{x};\cdot)) \le 2^{1/q'} W_{\infty,q}(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(\tilde{x};\cdot)).$$

Therefore, it is clear that whenever $q > 1$, the additional factor $2^{1/q'}$ makes the one-step contraction factor in next section nonoptimal as compared to the explicit one-step contraction for quadratic selection $m(x) = \frac{1}{2}\,\alpha\,|x|^2$, as illustrated in Remark 2.7 and detailed later in Remark 3.1.

**2.4.** ***Contraction of the one-step transition probability.*** The last step of our argument requires showing that the mapping $x \in \mathbb{R} \mapsto \boldsymbol{P}(x;\cdot) \in L^1_+(\mathbb{R}^2) \cap \mathcal{P}(\mathbb{R}^2)$ is a contraction when the space $\mathcal{P}(\mathbb{R}^2)$ is endowed with the $W_{\infty,1}$ Wasserstein distance in (1-24). Specifically, in the following result we quantify the exact Lipschitz constant, which will account for the precise contraction factor in Theorem 1.2(ii).

**Lemma 2.6** ($W_{\infty,1}$-contraction). *Consider the one-step transition probability $\boldsymbol{P} = \boldsymbol{P}(x; x_1, x_2)$ defined in* (2-5) *in terms of the potential $\boldsymbol{V}$ of the $\beta$-log-concave quasiequilibrium $\boldsymbol{F} = e^{-\boldsymbol{V}}$ in Theorem 1.2*(i). *Then, the following inequality holds true for every $x, \tilde{x} \in \mathbb{R}$*:

$$W_{\infty,1}(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(\tilde{x};\cdot)) \le \frac{2}{1+2\beta}|x-\tilde{x}|.$$

A similar contraction property, with respect to $W_1$ distances instead of $W_\infty$, appeared previously in [Ollivier 2007; 2009] leading to the definition of coarse Ricci curvature of a Markov kernel $\boldsymbol{P}(x;\cdot)$:

$$\kappa(x,\tilde{x}) = 1 - \frac{W_1(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(\tilde{x};\cdot))}{|x-\tilde{x}|}, \quad x,\ \tilde{x} \in \mathbb{R}.$$

Specifically, the above references proved that a positive lower bound on the coarse Ricci curvature amounts to the aforementioned contraction of the forward semigroup under the Lipschitz norm (or equivalently, the contraction of the backward semigroup under the $L^1$ transport distance [Kuwada 2010]). For heat kernels in a linear setting, this hypothesis on the coarse Ricci curvature is compatible with the Bakry–Emery convexity condition and was proved equivalent to the contraction of the backward semigroup in all $W_p$ transport distances [von Renesse and Sturm 2005], including $W_\infty$. However, the decay of the $L^\infty$ relative Fisher information has not been addressed in those works, and a nonlinear adaptation of them does not seem straightforward.

Before entering into the details of the proof of the Lemma 2.6, let us note that putting Lemmas 2.4 and 2.6 together automatically implies the one-step contraction estimate

$$\left\| \frac{d}{dx}(\log u_1) \right\|_{L^\infty} \le \frac{2}{1+2\beta} \left\| \frac{d}{dx}(\log u_0) \right\|_{L^\infty}, \tag{2-8}$$

which can be iterated and propagated into (1-9) in Theorem 1.2(ii) (at generation $n$), thus concluding this section. Nevertheless, we remark that Lemma 2.6 is far from straightforward as one typically cannot even ensure that the above $W_{\infty,1}$ distance must be finite because the probability densities $\boldsymbol{P}(x;\cdot)$ and $\boldsymbol{P}(\tilde{x};\cdot)$ are supported on the full plane $\mathbb{R}^2$.

**Remark 2.7** (quadratic selection). In the case of quadratic selection $m(x) = \frac{1}{2}\,\alpha\,|x|^2$ studied in [Calvez et al. 2024], we recall from Remark 1.3 that the unique eigenfunction of (1-5) is the Gaussian $\boldsymbol{F} = G_{0,\sigma^2}$ with variance $\sigma^2 = \beta^{-1}$. Therefore, one easily obtains from (2-5) that

$$\boldsymbol{P}(x, x_1, x_2) \propto \exp\left(-\tfrac{1}{2}\left|x - \tfrac{1}{2}(x_1+x_2)\right|^2 - \tfrac{1}{2}\beta|x_1|^2 - \tfrac{1}{2}\beta|x_2|^2\right).$$

Completing squares with respect to the variables $(x_1, x_2)$ we readily find that $\boldsymbol{P}(x;\cdot) = G_{\mu_x,\Sigma}$ is the density of a bivariate normal distribution with mean and covariance matrix determined by

$$\mu_x := \frac{1}{1+2\beta}(x,x), \qquad \Sigma^{-1} := \begin{pmatrix} \frac{1}{4}+\beta & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4}+\beta \end{pmatrix}.$$

Since $\Sigma$ is independent of $x$, any couple of Gaussians $\boldsymbol{P}(x;\cdot)$ and $\boldsymbol{P}(\tilde{x};\cdot)$ must agree up to a translation in the direction joining their means. Hence, the transport cost reduces to moving the center $\mu_x$ of $\boldsymbol{P}(x;\cdot)$ to the center $\mu_{\tilde{x}}$ of $\boldsymbol{P}(\tilde{x};\cdot)$, which yields Lemma 2.6 (with identity indeed):

$$W_{\infty,1}(\boldsymbol{P}(x;\cdot), \boldsymbol{P}(\tilde{x};\cdot)) = \|\mu_x - \mu_{\tilde{x}}\|_1 = \frac{2}{1+2\beta}|x-\tilde{x}|.$$

The goal of this section is to prove Lemma 2.6. To alleviate the notation, throughout this section we let $z := (x_1, x_2) \in \mathbb{R}^2$, we fix $x, \tilde{x} \in \mathbb{R}$ with $x \neq \tilde{x}$ and then we simplify the notation on the one-step transition probability in (2-5) by setting $\boldsymbol{p}(z) := \boldsymbol{P}(x; x_1, x_2)$ and $\tilde{\boldsymbol{p}}(z) := \boldsymbol{P}(\tilde{x}; x_1, x_2)$, that is,

$$\boldsymbol{p}(z) = \frac{1}{\boldsymbol{Z}} e^{-\boldsymbol{W}(z)}, \qquad \tilde{\boldsymbol{p}}(z) = \frac{1}{\widetilde{\boldsymbol{Z}}} e^{-\widetilde{\boldsymbol{W}}(z)}, \tag{2-9}$$

where the potentials $\boldsymbol{W}$ and $\widetilde{\boldsymbol{W}}$, and the normalizing constants $\boldsymbol{Z}$ and $\widetilde{\boldsymbol{Z}}$ are then given by

$$\begin{aligned} \boldsymbol{W}(z) &:= \boldsymbol{W}(x; x_1, x_2) = \tfrac{1}{2}\left|x - \tfrac{1}{2}(x_1+x_2)\right|^2 + \boldsymbol{V}(x_1) + \boldsymbol{V}(x_2), \\ \widetilde{\boldsymbol{W}}(z) &:= \boldsymbol{W}(\tilde{x}; x_1, x_2) = \tfrac{1}{2}\left|\tilde{x} - \tfrac{1}{2}(x_1+x_2)\right|^2 + \boldsymbol{V}(x_1) + \boldsymbol{V}(x_2), \\ \boldsymbol{Z} &:= \boldsymbol{Z}(x) = \iint_{\mathbb{R}^2} e^{-\boldsymbol{W}(z)}\,dz, \qquad \widetilde{\boldsymbol{Z}} := \boldsymbol{Z}(\tilde{x}) = \iint_{\mathbb{R}^2} e^{-\widetilde{\boldsymbol{W}}(z)}\,dz. \end{aligned} \tag{2-10}$$

For any transport map $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T_\# \boldsymbol{p} = \tilde{\boldsymbol{p}}$, note that a possible strategy in order to estimate the $W_{\infty,1}$ distance is to compute an $L^\infty$ bound for the $\ell_1$ associated displacement, namely,

$$W_{\infty,1}(\boldsymbol{p}, \tilde{\boldsymbol{p}}) \leq \big\| \|T - I\|_1 \big\|_{L^\infty}. \tag{2-11}$$

Whilst the choice of $T$ is somehow arbitrary at this point, a comfortable one is usually the Brenier map $T : \mathbb{R}^2 \to \mathbb{R}^2$ from the density $\boldsymbol{p}$ to the density $\tilde{\boldsymbol{p}}$, which is characterized as the unique transport map satisfying $T_\# \boldsymbol{p} = \tilde{\boldsymbol{p}}$ and solving the Monge problem [Brenier 1991]

$$\iint_{\mathbb{R}^2} \|T(z) - z\|_2^2\, \boldsymbol{p}(z)\,dz = W_{2,2}^2(\boldsymbol{p}, \tilde{\boldsymbol{p}}),$$

where $W_{2,2}$ is the $L^2$ Wasserstein distance associated with the $\ell_2$ norm of $\mathbb{R}^2$; see (1-24). As we anticipated in the Methodological notes in Section 1, in many cases this nonoptimal argument leads to no loss of generality since the $W_{\infty,1}$ and the uniform bound of the $\ell_1$ displacement of the Brenier map have the same order. This was further depicted in the example of the Gaussians from Remark 2.7, where the Brenier map is a translation, and therefore the transport cost is indeed identical to the displacement.

Our proof of Lemma 2.6 is based on the derivation of a novel $L^\infty$ bound of the $\ell_1$ displacement $\|T - I\|_1$ associated with the Brenier map $T$ between the densities $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$. We derive those bounds by reformulating such a Brenier map as a solution to a Monge–Ampère equation and using a version of *Caffarelli's maximum principle* along with the strong log-concavity of our densities. Indeed, by the strong log-concavity of $\boldsymbol{F}$ in Theorem 1.2(i) we have

$$-D^2_{(x_1,x_2)} \log \boldsymbol{p} = -D^2_{(x_1,x_2)} \log \tilde{\boldsymbol{p}} \geq \begin{pmatrix} \frac{1}{4}+\beta & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4}+\beta \end{pmatrix} \geq \beta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and then $\boldsymbol{p}$, $\tilde{\boldsymbol{p}}$ are $\beta$-log-concave. The aforementioned strategy recalls the one applied in *Caffarelli's contraction principle* [2000] (see also [Colombo and Fathi 2021; Colombo et al. 2017]) to find Lipschitz bounds of the Brenier map between strongly log-concave probability densities. Yet, in order to obtain Lipschitz bounds on the map (i.e., bounds on the Hessian of the potential), it is necessary to differentiate twice the Monge–Ampère equation; here we only require bounds on the displacement, and we need to differentiate only once. This recalls more what was done in [Ferrari and Santambrogio 2021], where the goal was to obtain Lipschitz bounds on the logarithm of the solution of a JKO scheme or, equivalently, $L^\infty$ bounds of the displacement associated with the Brenier map between two subsequent measures in the same JKO scheme. Among the important differences, [Ferrari and Santambrogio 2021] was not concerned with log-concave measures, but required one of the two to be obtained from the other via the JKO scheme. As another important difference, [Ferrari and Santambrogio 2021] was concerned with $\ell_2$ displacement bounds, and the choice of the Euclidean ball played a special role. In our setting, in view of the definition (1-24) of $W_{\infty,1}$, the choice of $\ell_2$ is not suitable and we focus on $\ell_1$. For the $\ell_1$ norm, we obtain new bounds on the Monge–Ampère equation, which lead to the sharp contraction factor, and which cannot be recovered by interpolation from known $\ell_2$ estimates; see Remark 3.1.

For the reader's convenience, we provide below a formal proof of Lemma 2.6 under the strong additional assumption that the maximal $\ell_1$ displacement associated with the Brenier map is attained. Whilst true in particular situations (see Remark 2.7), unfortunately this hypothesis is not necessarily always true, and thus the rigorous derivation requires further work which we provide in detail in Section 3.

*Formal proof of Lemma 2.6.* It is well known that the Brenier map $T : \mathbb{R}^2 \to \mathbb{R}^2$ from $\boldsymbol{p}$ to $\tilde{\boldsymbol{p}}$ takes the form $T = \nabla\phi$ for some convex function $\phi : \mathbb{R}^2 \to \mathbb{R}$. Since $\boldsymbol{p}$, $\tilde{\boldsymbol{p}} > 0$ and $\boldsymbol{p}$, $\tilde{\boldsymbol{p}} \in C^\infty(\mathbb{R}^2)$, the regularity results in [Caffarelli 1992b] imply that $\phi \in C^\infty(\mathbb{R}^2)$. Moreover, the change of variable formula implies

$$\det(D^2\phi) = \frac{\boldsymbol{p}}{\tilde{\boldsymbol{p}} \circ \nabla\phi}, \quad z \in \mathbb{R}^2. \tag{2-12}$$

As usual we make the change of variables through the displacement potential

$$\psi(z) := \phi(z) - \tfrac{1}{2}\|z\|_2^2, \quad z \in \mathbb{R}^2. \tag{2-13}$$

In view of the relation (2-11), we note that the core of the proof then reduces to obtaining $L^\infty$ bounds for the $\ell_1$ norm of the displacement of the Brenier map, that is,

$$H(z) := \|T(z) - z\|_1 = \|\nabla\psi(z)\|_1 = |\partial_{x_1}\psi(z)| + |\partial_{x_2}\psi(z)|, \quad z \in \mathbb{R}^2. \tag{2-14}$$

We start by restating the Monge–Ampère equation (2-12) by taking its logarithm,

$$\log\det(D^2\psi(z) + I) = \widetilde{\boldsymbol{W}}(\nabla\psi(z) + z) - \boldsymbol{W}(z) + \log\frac{\widetilde{\boldsymbol{Z}}}{\boldsymbol{Z}}, \quad z \in \mathbb{R}^2. \tag{2-15}$$

Taking partial derivatives $\partial_{x_k}$ in (2-15) we have

$$\operatorname{tr}((D^2\phi)^{-1}\partial_{x_k} D^2\psi) = \nabla \widetilde{\boldsymbol{W}}(\nabla\psi + z)\cdot \partial_{x_k}\nabla\psi + (\nabla \widetilde{\boldsymbol{W}}(\nabla\psi+z) - \nabla \boldsymbol{W})\cdot e_k, \quad z\in\mathbb{R}^2, \tag{2-16}$$

for $k = 1, 2$. Let us assume that $H$ attains its maximum at some $z^* = (x_1^*, x_2^*) \in \mathbb{R}^2$ (for the general case where the maximum is not attained we refer to Section 3) and let us also define the auxiliary function

$$\widetilde{H}(z) := \operatorname{sgn}(\partial_{x_1}\psi(z^*))\,\partial_{x_1}\psi(z) + \operatorname{sgn}(\partial_{x_2}\psi(z^*))\,\partial_{x_2}\psi(z), \quad z\in\mathbb{R}^2. \tag{2-17}$$

Then, $\widetilde{H}$ must also attain its maximum at $z^*$ and it agrees with the maximum of $H$. In particular, we have the necessary optimality conditions

$$\nabla\widetilde{H}(z^*) = 0, \quad D^2\widetilde{H}(z^*) \le 0. \tag{2-18}$$

Now, we perform an appropriate convex combination of (2-16) depending on the signs of $\partial_{x_1}\psi(z^*)$ and $\partial_{x_2}\psi(z^*)$ in order to make the auxiliary function $\widetilde{H}$ in (2-14) appear.

**Case 1:** $\partial_{x_1}\psi(z^*) \ge 0$ and $\partial_{x_2}\psi(z^*) \ge 0$. In this case we have $\widetilde{H} := \partial_{x_1}\psi + \partial_{x_2}\psi$. Evaluating (2-16) at $z^*$ and summing over $k \in \{1, 2\}$ we have

$$\operatorname{tr}((D^2\phi(z^*))^{-1}D^2\widetilde{H}(z^*)) = \nabla\widetilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*)\cdot\nabla\widetilde{H}(z^*) + \big(\nabla\widetilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*) - \nabla \boldsymbol{W}(z^*)\big)\cdot(1,1).$$

By the optimality conditions (2-18) and since $D^2\phi(z^*)^{-1}$ is positive definite, the term in the left-hand side above is nonpositive, and we obtain

$$\big(\nabla\widetilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*) - \nabla\widetilde{\boldsymbol{W}}(z^*)\big)\cdot(1,1) \le \nabla(\boldsymbol{W}-\widetilde{\boldsymbol{W}})(z^*)\cdot(1,1) = \tilde{x} - x.$$

By expanding the left-hand side we obtain

$$\begin{aligned}
&\big(\nabla\widetilde{\boldsymbol{W}}(\nabla\psi(z^*)+z^*) - \nabla\widetilde{\boldsymbol{W}}(z^*)\big)\cdot(1,1)\\
&\qquad= \frac{\partial_{x_1}\psi(z^*)+\partial_{x_2}\psi(z^*)}{2} + \boldsymbol{V}'(\partial_{x_1}\psi(z^*)+x_1^*) - \boldsymbol{V}'(x_1^*) + \boldsymbol{V}'(\partial_{x_2}\psi(z^*)+x_2^*) - \boldsymbol{V}'(x_2^*)\\
&\qquad\ge \frac{\partial_{x_1}\psi(z^*)+\partial_{x_2}\psi(z^*)}{2} + \beta(\partial_{x_1}\psi(z^*)+\partial_{x_2}\psi(z^*)) = \frac{1+2\beta}{2}\widetilde{H}(z^*),
\end{aligned}$$

where we have used that in this case $\partial_{x_1}\psi(z^*) \ge 0$ and $\partial_{x_2}\psi(z^*) \ge 0$, along with the $\beta$-convexity of $\boldsymbol{V}$. Therefore, we conclude that $\tilde{x} > x$ and

$$\|H\|_{L^\infty} = H(z^*) = \widetilde{H}(z^*) \le \frac{2}{1+2\beta}|x-\tilde{x}|.$$

**Case 2:** $\partial_{x_1}\psi(z^*) < 0$ and $\partial_{x_2}\psi(z^*) < 0$. This case follows the same argument as Case 1. Indeed, note now that $\widetilde{H} = -\partial_{x_1}\psi - \partial_{x_2}\psi$. Then, we sum over $k \in 1, 2$, multiply by $-1$ on (2-16) and we obtain

$$\frac{1+2\beta}{2}\widetilde{H}(z^*) \le x - \tilde{x}.$$

Hence, in this case we obtain $x > \tilde{x}$ and we recover

$$\|H\|_{L^\infty} = H(z^*) = \widetilde{H}(z^*) \le \frac{2}{1+2\beta}|x-\tilde{x}|.$$

We show below that the other two cases (namely, $\partial_{x_1}\psi(z^*) \geq 0$ and $\partial_{x_2}\psi(z^*) < 0$, or $\partial_{x_1}\psi(z^*) < 0$ and $\partial_{x_2}\psi(z^*) \geq 0$) cannot happen.

**Case 3:** $\partial_{x_1}\psi(z^*) \geq 0$ and $\partial_{x_2}\psi(z^*) < 0$. Our goal is to show that this case cannot take place. In this case, we have $\widetilde{H} := \partial_{x_1}\psi - \partial_{x_2}\psi$. Taking the difference of (2-16) with $k = 1$ and $k = 2$ we obtain

$$\operatorname{tr}\big((D^2\phi(z^*))^{-1} D^2 \widetilde{H}(z^*)\big) = \nabla \widetilde{\boldsymbol{W}}(\nabla\psi(z^*) + z^*) \cdot \nabla \widetilde{H}(z^*) + \big(\nabla \widetilde{\boldsymbol{W}}(\nabla\psi(z^*) + z^*) - \nabla\, \boldsymbol{W}(z^*)\big) \cdot (1, -1).$$

Since $z^*$ is a maximizer of $\widetilde{H}$, we have

$$\big(\nabla \widetilde{\boldsymbol{W}}(\nabla\psi(z^*) + z^*) - \nabla \widetilde{\boldsymbol{W}}(z^*)\big) \cdot (1, -1) \leq \nabla(\boldsymbol{W} - \widetilde{\boldsymbol{W}})(z^*) \cdot (1, -1) = 0$$

The expansion on the left-hand side is now radically different because the above factor $\frac{1}{2}(\partial_{x_1}\psi(z^*) + \partial_{x_2}\psi(z^*))$ cancels and now we obtain

$$\begin{aligned}\big(\nabla \widetilde{\boldsymbol{W}}(\nabla\psi(z^*) + z^*) - \nabla \widetilde{\boldsymbol{W}}(z^*)\big) \cdot (1, -1) &= \boldsymbol{V}'(\partial_{x_1}\psi(z^*) + x_1^*) - \boldsymbol{V}'(x_1^*) - \boldsymbol{V}'(\partial_{x_2}\psi(z^*) + x_2^*) + \boldsymbol{V}'(x_2^*) \\ &\geq \beta(\partial_{x_1}\psi(z^*) - \partial_{x_2}\psi(z^*)) = \beta\, \widetilde{H}(z^*),\end{aligned}$$

which implies $\|H\|_{L^\infty} = H(z^*) = \widetilde{H}(z^*) = 0$. This is clearly impossible since otherwise $T(z) = z$ for all $z \in \mathbb{R}^2$, that is, $x = \tilde{x}$.

**Case 4:** $\partial_{x_1}\psi(z^*) < 0$ and $\partial_{x_2}\psi(z^*) \geq 0$. This case cannot happen either thanks to the same argument as in Case 3 with $\widetilde{H}$ replaced by $\widetilde{H} = -\partial_{x_1}\psi + \partial_{x_2}\psi$. Thus, we omit the proof. □

**2.5.** ***Proof of the one-step contraction property.*** With all the above machinery in hand, we are finally in position to prove the one-step contraction property (1-9) in Theorem 1.2.

*Proof of Theorem 1.2*(ii). Combining Lemmas 2.4 and 2.6 applied to the solution (2-3) of (2-4) we obtain

$$\left\| \frac{d}{dx}\left(\log \frac{F_n}{\boldsymbol{F}}\right)\right\|_{L^\infty} \leq \frac{2}{1 + 2\beta} \left\| \frac{d}{dx}\left(\log \frac{F_{n-1}}{\boldsymbol{F}}\right)\right\|_{L^\infty}$$

for every $n \in \mathbb{N}$, and this amounts to (1-9). □

## 3. Main contractivity lemma

In this section, we provide a rigorous proof of Lemma 2.6, where the a priori assumption that the maximal displacement associated with the Brenier map must be attained is no longer required. To do so, we shall argue by deriving a local version of the lemma valid for more general strongly log-concave densities $f$ and $g$ compactly supported on an appropriate domain and bounded away from zero on it. More specifically, we propose to adapt the contribution of the maximum principle to the formal argument above (Section 2.4) to compact domains. However, since the maximum may be attained at the boundary, the boundary information is crucial in order to infer information from the nonlinear elliptic PDE (2-12), and therefore the choice of the domain cannot be made arbitrarily.

We refer to Appendix C for a bound on the maximum of $\|T - I\|_2$ (in $\ell_2$ norm) for the Brenier map $T : \overline{B}_R \to \overline{B}_R$ between two generic strongly log-concave probability densities $f = e^{-W}$ and $g = e^{-\widetilde{W}}$,

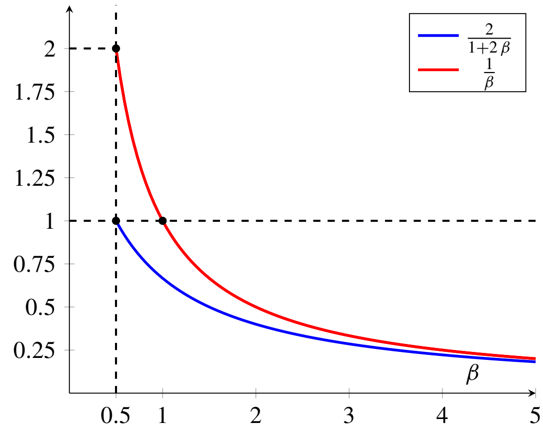



**Figure 3.** Comparison of the theoretical contraction factor $\frac{1}{1+2\beta}$ in Lemma 2.6, and the contraction factor $\frac{1}{\beta}$ obtained by estimating the $\ell_1$ norm with the $\ell_2$ norm in $\mathbb{R}^2$.

supported and strictly positive on an Euclidean ball $\bar{B}_R$. Specifically, we obtain

$$W_{\infty,2}(f,g) \le \big\| \|T-I\|_2 \big\|_{L^\infty(\bar{B}_R)} \le \frac{1}{\gamma} \big\| \|\nabla(W-\widetilde{W})\|_2 \big\|_{L^\infty(\bar{B}_R)}, \tag{3-1}$$

where $\gamma > 0$ is the log-concavity parameter of $f$ and $g$.

**Remark 3.1** (inaccuracy of controlling $\ell_1$ by $\ell_2$ norms). We may be tempted to apply this $\ell_2$ estimate to our setting by setting $f$ and $g$ as truncations of $\boldsymbol{p} \propto e^{-\boldsymbol{W}}$ and $\tilde{\boldsymbol{p}} \propto e^{-\widetilde{\boldsymbol{W}}}$ (see (2-9)–(2-10)) to $\ell_2$ balls and using the Cauchy–Schwarz inequality to get $\ell_1$ estimates. Specifically, consider an increasing sequence of balls $B_R$ and set $f$ and $g$ in (3-1) to be the truncation of $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ on such balls. First, recall that

$$D^2\,\boldsymbol{W}(x_1,x_2) = D^2\widetilde{\boldsymbol{W}}(x_1,x_2) = \begin{pmatrix} \frac{1}{4}+\boldsymbol{V}''(x_1) & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4}+\boldsymbol{V}''(x_2) \end{pmatrix} \ge \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix},$$

because $\boldsymbol{V}'' \ge 0$, and therefore we can set $\gamma = \beta$ in (3-1). Also note that

$$\nabla(\boldsymbol{W} - \widetilde{\boldsymbol{W}})(x_1,x_2) = \tfrac{1}{2}(\tilde{x}-x, \tilde{x}-x).$$

Altogether this implies the $\ell_2$ estimate

$$W_{\infty,2}(f,g) \le \big\| \|T-I\|_2 \big\|_{L^\infty(\bar{B}_R)} \le \frac{1}{\beta} \big\| \|\nabla(\boldsymbol{W}-\widetilde{\boldsymbol{W}})\|_2 \big\|_{L^\infty} = \frac{1}{\beta} \big\| \tfrac{1}{2}(\tilde{x}-x,\tilde{x}-x) \big\|_2 = \frac{1}{\sqrt{2}\,\beta}|x-\tilde{x}|,$$

and by the Cauchy–Schwarz inequality we also have the $\ell_1$ estimate

$$W_{\infty,1}(f,g) \le \sqrt{2}W_{\infty,2}(f,g) \le \frac{1}{\beta}|x-\tilde{x}|.$$

In particular, we note that such an estimate only provides contraction as long as $\beta > 1$ and, in addition, the contraction factor is worse than the one claimed in Lemma 2.6 as depicted in Figure 3.

We refer to [Khudiakova et al. 2024] for a nice and fruitful anisotropic version of (3-1) which enables us to obtain directly the claimed contraction factor.

Thus, we need to improve our proof and avoid using the $\ell_2$ norm. This was done, formally, in the previous section, but we need a rigorous proof which also takes care of the boundary. Let us focus on the observation made in [Ferrari and Santambrogio 2021, Lemma 3.1] that, for generic $f$ and $g$ smooth on a $\ell_2$ ball and bounded away from zero on it, the maximal $\ell_2$ displacement of the Brenier map must be attained at some interior point in the ball. Apparently, the use of $\ell_2$ norms to quantify the size of the displacement proved extremely well-suited to control the boundary information on $\ell_2$ balls. Interestingly, in the sequel we show that in order to find precise information about the maximizers for the $\ell_1$ displacement, we need densities $f$ and $g$ to be supported over $\ell_\infty$ balls $\overline{B}_R$ (see (1-22)). This is the content of the following.

**Lemma 3.2** (maximizers in the $\ell_1$ setting). *Consider two densities $f, g \in L^1_+(\mathbb{R}^2) \cap \mathcal{P}(\mathbb{R}^2)$, assume that,*

$$\{z \in \mathbb{R}^2 : f(z) > 0\} = \{z \in \mathbb{R}^2 : g(z) > 0\} = \overline{Q}_R,$$

*where $Q_R$ is the $\ell_\infty$ ball (see (1-22)), and suppose that $f, g \in C^{1,\delta}(\overline{Q}_R)$ for some $\delta > 0$. Let $T = \nabla\phi : \overline{Q}_R \to \overline{Q}_R$ be the Brenier map from $f$ to $g$, define the displacement potential $\psi(z) := \phi(z) - \frac{1}{2}\|z\|_2^2$ and the displacement function quantified in the $\ell_1$ norm*

$$H(z) := \|T(z) - z\|_1 = |\partial_{x_1}\psi(z)| + |\partial_{x_2}\psi(z)|, \quad z \in \overline{Q}_R. \tag{3-2}$$

*Then, $T \in C^{2,\delta}(\overline{Q}_R)$ and we have the optimality conditions*

$$\nabla\widetilde{H}(z^*) = 0, \quad D^2\widetilde{H}(z^*) \le 0 \tag{3-3}$$

*for any maximizer $z^* = (z_1^*, z_2^*) \in \overline{Q}_R$ of $H$, where $\widetilde{H}$ is the auxiliary function*

$$\widetilde{H}(z) := \operatorname{sgn}(\partial_{x_1}\psi(z^*))\,\partial_{x_1}\psi(z) + \operatorname{sgn}(\partial_{x_2}\psi(z^*))\,\partial_{x_2}\psi(z), \quad z \in \overline{Q}_R. \tag{3-4}$$

In contrast with the standard regularity theory for optimal transport, $Q_R$ is not uniformly convex. Then, the regularity theory of the Monge–Ampère equation is not directly applicable in full generality. Specifically, since $f, g \in C^{1,\delta}(\overline{Q}_R)$ are bounded away from zero on $\overline{Q}_R$, we have $T \in C^{0,\delta}(\overline{Q}_R)$ by [Caffarelli 1992a]. However, the lack of uniform convexity may prevent the full elliptic regularity [Caffarelli 1996], which claims that $T$ is a diffeomorphism of class $C^{2,\delta}(\overline{Q}_R)$. Fortunately, we can proceed as in [Jhaveri 2019, Theorem 3.3] which, thanks to a clever symmetrization argument around each corner of $Q_R$ and the classical interior regularity in [Caffarelli 1992b], shows that $T$ is indeed a diffeomorphism of class $C^{2,\delta}(\overline{Q}_R)$. Moreover, it fixes the corners and sends each segment of the boundary to itself. This guarantees in particular that $\widetilde{H} \in C^2(\overline{Q}_R)$ and the optimality conditions above make sense, as shown below.

*Proof of Lemma 3.2.* We remark that $z^* \in \overline{Q}_R$ must also be a maximizer of $\widetilde{H}$ since we have

$$\widetilde{H}(z) \le H(z) \le H(z^*) = \widetilde{H}(z^*)$$

for every $z^* \in \overline{Q}_R$ by the definitions of $H$ and $\widetilde{H}$ in (3-2) and (3-4). Since the maximizer $z^*$ may lie in principle in all $\overline{Q}_R$, two possible options arise, either $z^* \in Q_R$ or $z^* \in \partial Q_R$. In the first case, the usual optimality conditions at interior points yield (3-3). In the second case, namely $z^* \in \partial Q_R$, note that the result is trivial if $z^*$ is one of the four corners since those are fixed points of $T$ and therefore $\widetilde{H} \equiv 0$.

Hence, from here on we will assume that $z^* \in \partial Q_R$ is not at a corner, but it lies in the interior of some of the four segments. Note that at those points we only have to prove that $\nabla \widetilde{H}(z^*) = 0$. In fact, we remark that those $z^*$ can be approached by interior points from any direction, and then the above readily implies the second-order optimality condition $D^2 \widetilde{H}(z^*) \leq 0$. To show that $\nabla \widetilde{H}(z^*) = 0$, note that the boundary $\partial Q_R$ contains four segments:

$$S_1^+ := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = R,\ x_2 \in [-R, R]\}, \qquad S_2^+ := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \in [-R, R],\ x_2 = R\},$$
$$S_1^- := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = -R,\ x_2 \in [-R, R]\}, \quad S_2^- := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \in [-R, R],\ x_2 = -R\}.$$

Since $T(\partial Q_R) = \partial Q_R$ and each segment is mapped to itself, we have

$$\partial_{x_1} \psi(z) = 0 \quad \text{if } z \in S_1^+ \cup S_1^-, \tag{3-5}$$
$$\partial_{x_2} \psi(z) = 0 \quad \text{if } z \in S_2^+ \cup S_2^-. \tag{3-6}$$

By differentiation it is clear that we also have

$$\partial_{x_1 x_2} \psi(z) = 0 \quad \text{if } z \in \partial Q_R. \tag{3-7}$$

Now, we argue according to the four possible segments of $\partial Q_R$ that $z^*$ may belong to.

**Case 1:** $z^* \in S_1^+ \cup S_1^-$. In this case, by (3-5) we have $\partial_{x_1} \psi(z^*) = 0$ and therefore we have

$$\widetilde{H}(z) = \operatorname{sgn}(\partial_{x_2} \psi(z^*))\, \partial_{x_2} \psi(z), \quad z \in \overline{Q}_R.$$

Since $z^*$ is a maximizer of $\widetilde{H}$, there exists $\lambda \in \mathbb{R}$ (indeed $\lambda \geq 0$ if $z^* \in S_1^+$ and $\lambda \leq 0$ if $z^* \in S_1^-$) such that its gradient at $z^*$ equals the multiple $\lambda(1, 0)$ of the outer normal vector, that is,

$$\nabla \widetilde{H}(z^*) = \operatorname{sgn}(\partial_{x_2} \psi(z^*)) \begin{pmatrix} \partial_{x_1 x_2} \psi(z^*) \\ \partial_{x_2 x_2} \psi(z^*) \end{pmatrix} = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}.$$

This implies that the second component of the gradient must vanish, but the first one also vanishes by the condition (3-7) on the crossed derivative. Then, we have $\nabla \widetilde{H}(z^*) = 0$.

**Case 2:** $z^* \in S_2^+ \cup S_2^-$. In this case, by (3-6) we have $\partial_{x_2} \psi(z^*) = 0$ and therefore we have

$$\widetilde{H}(z) = \operatorname{sgn}(\partial_{x_1} \psi(z^*))\, \partial_{x_1} \psi(z), \quad z \in \overline{Q}_R.$$

Since $z^*$ is a maximizer of $\widetilde{H}$, there exists $\lambda \in \mathbb{R}$ (indeed $\lambda \geq 0$ if $z^* \in S_2^+$ and $\lambda \leq 0$ if $z^* \in S_2^-$) such that its gradient at $z^*$ equals the multiple $\lambda(0, 1)$ of the outer normal vector, that is,

$$\nabla \widetilde{H}(z^*) = \operatorname{sgn}(\partial_{x_1} \psi(z^*)) \begin{pmatrix} \partial_{x_1 x_1} \psi(z^*) \\ \partial_{x_1 x_2} \psi(z^*) \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda \end{pmatrix}.$$

This implies that the first component of the gradient must vanish, but the second one also vanishes by the condition (3-7) on the crossed derivative. Then, we have $\nabla \widetilde{H}(z^*) = 0$. □

We remark that the unique formal point of the sketch of the proof of Lemma 2.6 in Section 2 which could break down is the fact that for the global densities $f = \boldsymbol{p}$ and $g = \tilde{\boldsymbol{p}}$ in (2-9)–(2-10) the $\ell_1$ displacement of their Brenier map does not necessarily attain its maximum. In particular, we may be deprived of the optimality condition (2-18), which was crucially used throughout the maximum-type principle sketched in Section 2. However, Lemma 3.2 does guarantee that the maximum must be attained and the optimality conditions (3-3) must hold in particular when $f$ and $g$ are set to be the truncation of the densities $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ on $\ell_\infty$ balls. In fact, the result does not exploit the special potential $\boldsymbol{V}$ in the definition (2-9)–(2-10) of $\boldsymbol{p}$, $\tilde{\boldsymbol{p}}$, which corresponds to the potential of the eigenfunction $\boldsymbol{F} = e^{-\boldsymbol{V}}$ in Theorem 1.2(i), but it can actually be replaced by any strongly convex function supported on $\overline{Q}_R$. Since we shall use this more general version later in Section 4, we state in full generality below.

**Lemma 3.3** (maximum principle on $\ell_\infty$ balls). *For any $\gamma$-convex potential $V \in C^{1,\delta}_{\mathrm{loc}}(\mathbb{R})$ with $\gamma > 0$, any $x, \tilde{x} \in \mathbb{R}$ with $x \neq \tilde{x}$, and any $R > 0$ we define $f, g \in L^1_+(\mathbb{R}^d) \cap \mathcal{P}(\mathbb{R}^2)$ given by*

$$f(z) = \frac{1}{Z} e^{-W(z)}, \quad g(z) = \frac{1}{\widetilde{Z}} e^{-\widetilde{W}(z)}, \quad z \in \mathbb{R}^2,$$

*where the potentials $W$ and $\widetilde{W}$, and the normalizing constants $Z$ and $\widetilde{Z}$ are*

$$\begin{aligned} W(z) &:= \tfrac{1}{2}\left|x - \tfrac{1}{2}(x_1 + x_2)\right|^2 + V(x_1) + V(x_2) + \chi_{\overline{Q}_R}(z), \\ \widetilde{W}(z) &:= \tfrac{1}{2}\left|\tilde{x} - \tfrac{1}{2}(x_1 + x_2)\right|^2 + V(x_1) + V(x_2) + \chi_{\overline{Q}_R}(z), \\ Z &:= \iint_{\mathbb{R}^2} e^{-W(z)}\, dz, \quad \widetilde{Z} := \iint_{\mathbb{R}^2} e^{-\widetilde{W}(z)}\, dz, \end{aligned}$$

*and $\chi_{\overline{Q}_R}$ is the characteristic function associated to the $\ell_\infty$ ball $\overline{Q}_R$; see* (1-23). *Then, the Brenier map $T = \nabla\phi : \overline{Q}_R \to \overline{Q}_R$ from $f$ to $g$ satisfies*

$$W_{\infty,1}(f, g) \le \big\| \|T - I\|_1 \big\|_{L^\infty(\overline{Q}_R)} \le \frac{2}{1+2\gamma} |x - \tilde{x}|.$$

As explained above, we omit the proof since it follows the formal proof of Lemma 2.6 in Section 2 and the optimality conditions in Lemma 3.2. In particular, by setting $V = \boldsymbol{V}$ (and therefore $\gamma = \beta$) we have that Lemma 3.3 is directly applicable to the truncations to $\overline{Q}_R$ of the densities $\boldsymbol{p}$, $\tilde{\boldsymbol{p}}$ in (2-9)–(2-10).

**Definition 3.4** (truncation to $\overline{Q}_R$). For the probability densities $\boldsymbol{p}, \tilde{\boldsymbol{p}} \in L^1_+(\mathbb{R}^2) \cap \mathcal{P}(\mathbb{R}^2)$ given in (2-9)–(2-10), we define their truncations to the $\ell_\infty$ ball $\overline{Q}_R$ (see (1-22)) as

$$\begin{aligned} \boldsymbol{p}_R(z) &:= \frac{1}{\boldsymbol{Z}_R} e^{-\boldsymbol{W}_R(z)}, & \tilde{\boldsymbol{p}}_R(z) &:= \frac{1}{\widetilde{\boldsymbol{Z}}_R} e^{-\widetilde{\boldsymbol{W}}_R(z)}, \\ \boldsymbol{W}_R(z) &:= \boldsymbol{W}(z) + \chi_{\overline{Q}_R}(z), & \widetilde{\boldsymbol{W}}_R(z) &:= \widetilde{\boldsymbol{W}}(z) + \chi_{\overline{Q}_R}(z), \\ \boldsymbol{Z}_R &:= \int_{\mathbb{R}^2} e^{-\boldsymbol{W}_R(z)}\, dz, & \widetilde{\boldsymbol{Z}}_R &:= \int_{\mathbb{R}^2} e^{-\widetilde{\boldsymbol{W}}_R(z)}\, dz, \end{aligned}$$

for any $R > 0$, where $\chi_{\overline{Q}_R}$ is the characteristic function associated to the $\ell_\infty$ ball $\overline{Q}_R$; see (1-23).

Then, we are in position to rigorously prove Lemma 2.6 by taking limits $R \to \infty$ and noting that Lemma 3.3 yields a uniform bound of the displacement independent of $R$.

*Rigorous proof of Lemma 2.6.* Consider $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ given in (2-9)–(2-10) and set the associated Brenier map $T : \mathbb{R}^2 \to \mathbb{R}^2$ from $\boldsymbol{p}$ to $\tilde{\boldsymbol{p}}$. Similarly, we consider the family of truncations $\boldsymbol{p}_R$ and $\tilde{\boldsymbol{p}}_R$ in Definition 3.4

and we set the associated Brenier maps $T_R : \mathbb{R}^2 \to \mathbb{R}^2$. By the above Lemma 3.3 we have

$$\big\| \|T_R - I\|_1 \big\|_{L^\infty(\bar{Q}_R)} \le \frac{2}{1+2\beta}|x-\tilde{x}| \tag{3-8}$$

for every $R > 0$. We set the optimal transference plans $\gamma \in \Gamma_o(\boldsymbol{p}, \tilde{\boldsymbol{p}})$ and $\gamma_R \in \Gamma_o(\boldsymbol{p}_R, \tilde{\boldsymbol{p}}_R)$ associated with the $W_{2,2}$ distance, which are known to be supported on the graph of the above Brenier maps, i.e.,

$$\gamma := (I, T)_\# \boldsymbol{p}, \quad \gamma_R := (I, T_R)_\# \boldsymbol{p}_R .$$

Since the involved potentials $\boldsymbol{W}$ and $\widetilde{\boldsymbol{W}}$ are $\beta$-convex, we have enough integrability on $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ to ensure that $\boldsymbol{p}, \tilde{\boldsymbol{p}} \in \mathcal{P}_2(\mathbb{R}^2)$. Hence, the dominated convergence theorem applies and we have indeed

$$\boldsymbol{p}_R \to \boldsymbol{p}, \quad \tilde{\boldsymbol{p}}_R \to \tilde{\boldsymbol{p}} \quad \text{in } (\mathcal{P}_2(\mathbb{R}^2), W_{2,2}).$$

By stability of optimal transference plans, the sequence $\gamma_R$ must converge narrowly to some optimal transference plan (up to a subsequence); see [Ambrosio et al. 2008, Proposition 7.1.3]. Since the unique optimal transference plan between $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ is precisely the above $\gamma$ supported on the graph of $T$, we obtain

$$\gamma_R \to \gamma \quad \text{narrowly in } \mathcal{P}(\mathbb{R}^2).$$

Now we use the Kuratowski convergence of the supports under the narrow convergence of measures; see [Ambrosio et al. 2008, Proposition 5.1.8]. Namely, consider any $z \in \mathbb{R}^2$. Since $(z, T(z)) \in \operatorname{supp}\gamma$, there exists $(z^R, w^R) \in \operatorname{supp}\gamma_R$ such that $(z^R, w^R) \to (z, T(z))$. Since $\gamma_R$ is supported on the graph of $T_R$, we have $z^R \in \bar{Q}_R$ and $w^R = T_R(z^R)$. In particular, we have $T_R(z^R) - z^R \to T(z) - z$ as $R \to \infty$ and by the above uniform bound (3-8) the same bound is preserved in the limit, that is,

$$W_{\infty,1}(\boldsymbol{p}, \tilde{\boldsymbol{p}}) \le \big\| \|T - I\|_1 \big\|_{L^\infty} \le \frac{2}{1+2\beta}|x-\tilde{x}|. \qquad \square$$

**Remark 3.5** (replacing $\ell_\infty$ balls by $\ell_1$ balls). We note that in Lemmas 3.2 and 3.3 the choice of $\ell_\infty$ is crucial. However, this is not the only possible choice and a similar proof could be obtained if replacing $\ell_\infty$ balls with $\ell_1$ balls. It is clear anyway that the shape of the boundary and the norm to be optimized should satisfy some form of compatibility conditions.

## 4. Analysis of a truncated problem

In this part, we study an auxiliary version of the original time marching problem (1-1) restricted to the bounded interval $I_R := (-R, R)$ with $R > 0$, namely,

$$F_n^R = \mathcal{T}_R[F_{n-1}^R], \quad n \in \mathbb{N},\ x \in \mathbb{R}. \tag{4-1}$$

Here, we truncate the selection function $m_R$ as

$$m_R(x) := m(x) + \chi_{\bar{I}_R}(x), \quad x \in \mathbb{R}, \tag{4-2}$$

where $\chi_{\bar{I}_R}$ is the characteristic function associated to the interval $\bar{I}_R$ (see (1-23)), so that the truncated integral operator $\mathcal{T}_R$ takes the form

$$\mathcal{T}_R[F](x) := e^{-m_R(x)} \iint_{\mathbb{R}^2} G\Big(x - \frac{x_1+x_2}{2}\Big) F(x_1) \frac{F(x_2)}{\|F\|_{L^1}}\, dx_1\, dx_2, \quad x \in \mathbb{R}. \tag{4-3}$$

Again, solutions of the form $F_n^R(x) = (\lambda^R)^n F^R(x)$ come as eigenpairs of the nonlinear eigenproblem

$$\begin{aligned} &\lambda^R F^R = \mathcal{T}_R[F^R], \quad x \in \mathbb{R},\\ &F^R \geq 0, \quad \int_{\mathbb{R}} F^R(x)\,dx = 1. \end{aligned} \tag{4-4}$$

The goal of this section is to derive an analogous truncated version of Theorem 1.2. More specifically, we study (i) existence of a unique strongly log-concave solution $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ to (4-4), and (ii) quantitative relaxation of the solutions to (4-1) towards the quasiequilibrium $(\boldsymbol{\lambda}^R)^n\,\boldsymbol{F}^R$.

**Theorem 4.1** (truncated problem). *Consider any $m \in C^2(\mathbb{R})$ satisfying* (H1)–(H2) *in Theorem 1.2. Set any $R > 0$ and define the truncation $m_R$ according to* (4-2). *Then, the following statements hold true*:

(i) (existence of quasiequilibrium) *There is a unique solution $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ to* (4-4). *In addition, $\boldsymbol{F}^R = e^{-\boldsymbol{V}^R} \in L^1_+(\mathbb{R}) \cap C^\infty(\bar{I}_R)$ is compactly supported on $\bar{I}_R$ and bounded away from zero on it and $\beta$-log-concave with parameter $\beta > 0$ given in* (1-7) *in Theorem 1.2.*

(ii) (one-step contraction) *Consider any $F_0^R \in L^1_+(\mathbb{R}) \cap C^1(\bar{I}_R)$ compactly supported on $\bar{I}_R$ and bounded away from zero on it, and let $\{F_n^R\}_{n\in\mathbb{N}}$ be the solution to* (4-1) *issued at $F_0^R$. Then, we have*

$$\left\| \frac{d}{dx}\left(\log \frac{F_n^R}{\boldsymbol{F}^R}\right)\right\|_{L^\infty(\bar{I}_R)} \leq \frac{2}{1+2\beta} \left\| \frac{d}{dx}\left(\log \frac{F_{n-1}^R}{\boldsymbol{F}^R}\right)\right\|_{L^\infty(\bar{I}_R)}$$

*for any $n \in \mathbb{N}$.*

(iii) (asynchronous exponential growth) *Consider any $F_0^R \in L^1_+(\mathbb{R}) \cap C^1(\bar{I}_R)$ compactly supported on $\bar{I}_R$ and bounded away from zero on it, and let $\{F_n^R\}_{n\in\mathbb{N}}$ be the solution to* (4-1) *issued at $F_0^R$. Then, we have*

$$\begin{aligned} \left| \frac{\|F_n^R\|_{L^1}}{\|F_{n-1}^R\|_{L^1}} - \boldsymbol{\lambda}^R \right| &\leq C_R \left(\frac{2}{1+2\beta}\right)^n,\\ \left\| \frac{F_n^R}{\|F_n^R\|_{L^1}} - \boldsymbol{F}^R \right\|_{C^1} &\leq C_R' \left(\frac{2}{1+2\beta}\right)^n \end{aligned}$$

*for any $n \in \mathbb{N}$ and some constants $C_R$, $C_R'$ depending on $R$ and $F_0^R$.*

As we show below, our proof exploits the overarching local contraction result, Lemma 3.3, to answer simultaneously both questions. More specifically, our main observation is the following type of contraction which holds true providing that the initial data $F_0^R$ is strongly log-concave.

**Lemma 4.2** (Cauchy-type property). *Let $m \in C^2(\mathbb{R})$ satisfy* (H1)–(H2) *in Theorem 1.2. Consider a $\beta_0$-log-concave density $F_0^R \in L^1_+(\mathbb{R}) \cap C^{1,\delta}(\bar{I}_R)$ with $\beta_0 > 0$ and $0 < \delta < 1$, compactly supported on $\bar{I}_R$ and bounded away from zero on it. Let $\{F_n^R\}_{n\in\mathbb{N}}$ be the solution to* (4-1) *issued at $F_0^R$. Then, we have*

$$\left\| \frac{d}{dx}\left(\log \frac{F_n^R}{F_{n-1}^R}\right)\right\|_{L^\infty(\bar{I}_R)} \leq \frac{2}{1+2\beta_{n-2}} \left\| \frac{d}{dx}\left(\log \frac{F_{n-1}^R}{F_{n-2}^R}\right)\right\|_{L^\infty(\bar{I}_R)}, \quad n \geq 2,$$

*where the sequence $\{\beta_n\}_{n\in\mathbb{N}}$ is defined by recurrence as in* (2-2).

*Proof.* For any $n \in \mathbb{N}$, we define

$$u_n^R(x) := \frac{F_n^R(x)}{F_{n-1}^R(x)}, \quad x \in \bar{I}_R,$$

and note that, arguing as in (2-3), we have that $\{u_n\}_{n\in\mathbb{N}}$ must solve the following analog of (2-4):

$$u_n^R(x) = \frac{\|F_{n-2}^R\|_{L^1}}{\|F_{n-1}^R\|_{L^1}} \iint_{\bar{Q}_R} P_n^R(x; x_1, x_2) u_{n-1}^R(x_1) u_{n-1}^R(x_2)\, dx_1\, dx_2$$

for any $x \in \bar{I}_R$ and $n \geq 2$. We remark that the system above holds only on $\bar{I}_R$ and the one-step transition probability $P_n^R(x; \cdot) \in L_+^1(\bar{Q}_R) \cap \mathcal{P}(\bar{Q}_R)$ is not time-homogeneous but it depends explicitly on $n$, namely

$$\begin{aligned}
P_n^R(x; x_1, x_2) &:= \frac{1}{Z_n^R(x)} e^{-W_n^R(x; x_1, x_2)}, \quad x \in \bar{I}_R,\ (x_1, x_2) \in \bar{Q}_R,\\
W_n^R(x; x_1, x_2) &:= \tfrac{1}{2}\left|x - \tfrac{1}{2}(x_1 + x_2)\right|^2 + V_{n-2}^R(x_1) + V_{n-2}^R(x_2),\\
Z_n^R(x) &:= \iint_{\bar{Q}_R} e^{-W_n^R(x; x_1, x_2)}\, dx_1\, dx_2,
\end{aligned}$$

where we let $V_n^R : \bar{I}_R \to \mathbb{R}$ so that $F_n^R = e^{-V_n^R}$. By Lemma 2.2, $V_{n-2}^R$ is $\beta_{n-2}$-convex and therefore the contractivity Lemma 3.3 applies to $f = P_n^R(x; \cdot)$ and $g = P_n^R(\tilde{x}; \cdot)$ with $x, \tilde{x} \in \bar{I}_R$ leading to

$$W_{\infty,1}(P_n^R(x; \cdot), P_n^R(\tilde{x}; \cdot)) \leq \frac{2}{1 + 2\beta_{n-2}} |x - \tilde{x}|.$$

Therefore, arguing as in Lemma 2.4 we end the proof. □

*Proof of Theorem 4.1.* **Step 1:** Proof of (i). Under appropriate assumptions on $F_0^R$ we shall prove that $\|F_n^R\|_{L^1}/\|F_{n-1}^R\|_{L^1}$ and $F_n^R/\|F_n^R\|_{L^1}$ must converge as in (iii), and their limit $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ solves (4-4). We set a $\beta_0$-log-concave density $F_0^R \in L_+^1(\mathbb{R}) \cap C^{1,\delta}(\bar{I}_R)$ with $\beta_0 > \beta$ and $0 < \delta < 1$, compactly supported on $\bar{I}_R$ and bounded away from zero on it. Let $\{F_n^R\}_{n\in\mathbb{N}}$ be the solution to (4-1). Since the initial datum has been chosen strongly log-concave, Lemma 4.2 implies

$$\left\| \frac{d}{dx}\left(\log \frac{F_n^R}{F_{n-1}^R}\right) \right\|_{L^\infty(\bar{I}_R)} \leq \left(\frac{2}{1 + 2\beta}\right)^{n-1} \left\| \frac{d}{dx}\left(\log \frac{F_1^R}{F_0^R}\right) \right\|_{L^\infty(\bar{I}_R)}$$

for all $n \geq 1$ because $F_n^R$ are $\beta_n$-log-concave with $\beta_n > \beta$ for all $n \in \mathbb{N}$ by Lemma 2.2. Setting $V_n^R : \bar{I}_R \to \mathbb{R}$ as before so that $F_n^R = e^{-V_n^R}$ we obtain

$$\left\| \frac{d}{dx}(V_n^R - V_m^R) \right\|_{L^\infty(\bar{I}_R)} \leq \sum_{k=m+1}^{n} \left\| \frac{d}{dx}(V_k^R - V_{k-1}^R) \right\|_{L^\infty(\bar{I}_R)} \leq \sum_{k=m}^{n-1} \left(\frac{2}{1 + 2\beta}\right)^k \left\| \frac{d}{dx}(V_1^R - V_0^R) \right\|_{L^\infty(\bar{I}_R)}$$

for all $n \geq m \geq 1$. Since $\frac{2}{1+2\beta} < 1$ by Remark 1.7, $\left\{\frac{d}{dx}(V_n^R)\right\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $C(\bar{I}_R)$ and therefore it must converge uniformly to some limit $D^R \in C(\bar{I}_R)$. In particular, we have

$$\frac{d}{dx}(\log F_n^R) \to D^R \quad \text{in } C(\bar{I}_R). \tag{4-5}$$

Now, we show that $F_n^R/\|F_n^R\|_{L^1}$ must also converge when evaluated at least at one point, and we choose $x=0$ for instance. To this purpose, we note that $F_n^R(0)/\|F_n^R\|_{L^1}$ can be restated as

$$\frac{\iint_{\overline{Q}_R} G\big(\frac{1}{2}(x_1+x_2)\big)\exp\big(-(V_{n-1}^R(x_1)-V_{n-1}^R(0))-(V_{n-1}^R(x_2)-V_{n-1}^R(0))\big)dx_1\,dx_1}{\int_{\bar{I}_R}\iint_{\overline{Q}_R} G\big(x'-\frac{1}{2}(x_1+x_2)\big)\exp\big(-m(x')-(V_{n-1}^R(x_1)-V_{n-1}^R(0))-(V_{n-1}^R(x_2)-V_{n-1}^R(0))\big)\,dx'\,dx_1\,dx_2},$$

and, by the fundamental theory of calculus, $V_{n-1}^R(x)-V_{n-1}^R(0)$ in the integrand can be represented by

$$V_{n-1}^R(x)-V_{n-1}^R(0)=\int_0^1\frac{dV_{n-1}^R}{dx}(\theta x)x\,d\theta,\quad x\in\bar{I}_R,$$

which converges uniformly to some limit. Therefore, there exists $L^R\in\mathbb{R}$ such that

$$\log\frac{F_n^R(0)}{\|F_n^R\|_{L^1}}\to L^R. \tag{4-6}$$

Putting (4-5)–(4-6) together and using the fundamental theorem of calculus gives

$$\log\frac{F_n^R(x)}{\|F_n^R\|_{L^1}}=\log\frac{F_n^R(0)}{\|F_n^R\|_{L^1}}+\int_0^1\frac{d}{dx}(\log F_n^R)(\theta x)x\,d\theta\to L^R+\int_0^1 D^R(\theta x)x\,d\theta\quad\text{in } C^1(\bar{I}_R).$$

We define $\boldsymbol{F}^R(x):=\exp\big(L^R+\int_0^1 D^R(\theta x)x\,d\theta+\chi_{\bar{I}_R}(x)\big)\in L^1_+(\mathbb{R})\cap\mathcal{P}(\mathbb{R})$ and therefore we achieve

$$\frac{F_n^R}{\|F_n^R\|_{L^1}}\to\boldsymbol{F}^R\quad\text{in } C^1(\bar{I}_R). \tag{4-7}$$

Our second step is to prove the convergence of $\|F_n^R\|_{L^1}/\|F_{n-1}^R\|_{L^1}$. Note that we have

$$\frac{\|F_n^R\|_{L^1}}{\|F_{n-1}^R\|_{L^1}}=\iint_{\mathbb{R}^2}H_R(x_1,x_2)\frac{F_{n-1}^R(x_1)}{\|F_{n-1}^R\|_{L^1}}\frac{F_{n-1}^R(x_2)}{\|F_{n-1}^R\|_{L^1}}\,dx_1\,dx_2, \tag{4-8}$$

where we have defined

$$H_R(x_1,x_2):=\int_{\bar{I}_R}e^{-m(x)}G\big(x-\tfrac{1}{2}(x_1+x_2)\big)\,dx,\quad (x_1,x_2)\in\mathbb{R}^2.$$

Since $H_R$ is a bounded function, we have $H_R\in L^1(\overline{Q}_R)$ and, consequently, the above uniform convergence (4-7) of the normalized profiles, along with (4-8), implies that there must exists $\boldsymbol{\lambda}^R$ with

$$\frac{\|F_n^R\|_{L^1}}{\|F_{n-1}^R\|_{L^1}}\to\boldsymbol{\lambda}^R. \tag{4-9}$$

The last step is to show that $(\boldsymbol{\lambda}^R,\boldsymbol{F}^R)$ must solve (4-4). This is actually clear because we have

$$\frac{\|F_n^R\|_{L^1}}{\|F_{n-1}^R\|_{L^1}}\frac{F_n^R}{\|F_n^R\|_{L^1}}=\mathcal{T}_R\bigg[\frac{F_{n-1}^R}{\|F_{n-1}^R\|_{L^1}}\bigg]$$

for all $n\in\mathbb{N}$, and $\|F_n^R\|_{L^1}/\|F_{n-1}^R\|_{L^1}$ and $F_n^R/\|F_n^R\|_{L^1}$ converge in the above sense (4-7)–(4-9). We note that $\boldsymbol{F}^R$ must be $\beta$-log-concave because so is $F_n^R$ for all $n\in\mathbb{N}$. The uniqueness of the solution to (4-4) will not be analyzed here, but it will hold as a consequence of the next contraction property in Step 2.

**Step 2:** Proof of (ii). Once a strongly log-concave solution $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ of the truncated nonlinear eigenproblem (4-4) exists, the one-step contraction property follows the same ideas as in the global version in Theorem 1.2(ii) sketched in Section 2. More specifically, we shall argue like in the proof of Lemma 4.2 where again we replace $u_n$ by the normalization of $F_n^R$ by the quasiequilibrium $(\boldsymbol{\lambda}^R)^n \boldsymbol{F}^R$. That is, for any $n \in \mathbb{N}$, we define

$$u_n^R(x) := \frac{F_n^R(x)}{(\boldsymbol{\lambda}^R)^n \boldsymbol{F}^R}, \quad x \in \bar{I}_R,$$

which must solve

$$u_n^R(x) = \frac{1}{\|u_{n-1}^R \boldsymbol{F}^R\|_{L^1}} \iint_{\bar{Q}_R} \boldsymbol{P}^R(x; x_1, x_2) u_{n-1}^R(x_1) u_{n-1}^R(x_2)\, dx_1\, dx_2$$

for any $x \in \bar{I}_R$ and $n \in \mathbb{N}$, where $\boldsymbol{P}^R(x; \cdot) \in L^1_+(\bar{Q}_R) \cap \mathcal{P}(\bar{Q}_R)$ is the one-step transition probability

$$\begin{aligned} \boldsymbol{P}^R(x; x_1, x_2) &:= \frac{1}{\boldsymbol{Z}^R(x)} e^{-\boldsymbol{W}^R(x; x_1, x_2)}, \quad x \in \bar{I}_R,\ (x_1, x_2) \in \bar{Q}_R, \\ \boldsymbol{W}^R(x; x_1, x_2) &:= \tfrac{1}{2}\big|x - \tfrac{1}{2}(x_1 + x_2)\big|^2 + \boldsymbol{V}^R(x_1) + \boldsymbol{V}^R(x_2), \\ \boldsymbol{Z}^R(x) &:= \iint_{\bar{Q}_R} e^{-\boldsymbol{W}^R(x; x_1, x_2)}\, dx_1\, dx_2. \end{aligned}$$

Again, we let $\boldsymbol{V}^R : \bar{I}_R \to \mathbb{R}$ so that $\boldsymbol{F}^R = e^{-\boldsymbol{V}^R}$. By Step 1 we have that $\boldsymbol{V}^R$ is $\beta$-convex and therefore the contractivity result, Lemma 3.3, applies to $\boldsymbol{P}^R(x; \cdot)$ and $\boldsymbol{P}^R(\tilde{x}; \cdot)$ with $x, \tilde{x} \in \bar{I}_R$ leading to

$$W_{\infty,1}(\boldsymbol{P}^R(x; \cdot), \boldsymbol{P}^R(\tilde{x}; \cdot)) \le \frac{2}{1 + 2\beta} |x - \tilde{x}|.$$

Therefore, arguing as in Lemma 2.4 we end the proof.

In particular, the above implies that $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ must be the unique solution to the truncated nonlinear eigenequation (4-4). Indeed, if a second solution $(\lambda^R, F^R)$ exists, one can always define the special solution $F_n^R(x) = (\lambda^R)^n F^R(x)$ of (4-1) and therefore the above one-step contraction implies

$$\left\| \frac{d}{dx}\left(\log \frac{F^R}{\boldsymbol{F}^R}\right) \right\|_{L^\infty(\bar{I}_R)} \le \frac{2}{1 + 2\beta} \left\| \frac{d}{dx}\left(\log \frac{F^R}{\boldsymbol{F}^R}\right) \right\|_{L^\infty(\bar{I}_R)}.$$

Since $\frac{2}{1+2\beta} < 1$ by Remark 1.7, we have $F^R = \boldsymbol{F}^R$ (and therefore $\lambda^R = \boldsymbol{\lambda}^R$) because both $F^R$ and $\boldsymbol{F}^R$ are probability densities by definition.

**Step 3:** Proof of (iii). We prove that the convergence in Step 1 holds for generic initial data $F_0^R \in L^1_+(\mathbb{R}) \cap C^1(\bar{I}_R)$ compactly supported on $\bar{I}_R$ and bounded away from zero on it, and not necessarily strongly log-concave. Note that by the above one-step contractivity property we have again

$$\left\| \frac{d}{dx}(V_n^R - \boldsymbol{V}^R) \right\|_{L^\infty(\bar{I}_R)} \le \left(\frac{2}{1 + 2\beta}\right)^n \left\| \frac{d}{dx}(V_0^R - \boldsymbol{V}^R) \right\|_{L^\infty(\bar{I}_R)},$$

for all $n \in \mathbb{N}$. Then, the same argument as in Step 1 can be applied with explicit convergence rates and equal to $\left(\frac{2}{1+2\beta}\right)^n$ at each step: first $\frac{d}{dx}(\log F_n^R)$, then $\log(F_n^R(0)/\|F_n^R\|_{L^1})$, hence $\log(F_n^R/\|F_n^R\|_{L^1})$, and finally also $\|F_n^R\|_{L^1}/\|F_{n-1}^R\|_{L^1}$. Therefore, we readily obtain the claimed convergence rates for the rates of growth and the normalized profiles. □

## 5. Existence and uniqueness of strongly log-concave quasiequilibria

In this section, we employ the truncated quasiequilibria in the above Theorem 4.1 to build a globally defined quasiequilibrium of the nontruncated model (1-1), thus proving Theorem 1.2(i). In the following, we show that the probability densities in the family $\{\boldsymbol{F}^R\}_{R>0}$ are uniformly tight, and therefore weak limits cannot lose mass at infinity, which will be useful in the sequel in order to pass to the limit with $R \to \infty$.

**Proposition 5.1** (bounded second-order moments). *Under the assumptions in Theorem 4.1, let us consider the unique eigenpair* $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ *of* (4-4) *for any* $R > 0$ *according to Theorem 4.1*(i). *Then,*

$$\sup_{R>0} \int_{\mathbb{R}} x^2 \, \boldsymbol{F}^R(x)\, dx < \infty. \tag{5-1}$$

We recall that a similar result was necessary in [Calvez et al. 2024]. Indeed, a general strategy was developed therein to propagate second-order moments along any solution $\{F_n\}_{n\in\mathbb{N}}$ under the a priori knowledge that the centers of mass stay uniformly bounded. However, such a condition proved difficult to verify unless the initial datum $F_0$ is centered at the origin, and $m$ is an even function, which would leave the center of mass fixed at the origin (and thus bounded) for all times. To overcome this problem, an alternative approach was developed in [Calvez et al. 2024, Lemma 4.5] in order to control the convergence to zero of the center of mass in the case of quadratic selection. Unfortunately, the proof exploits the Gaussian structure in a crucial way and cannot be easily adapted to more general selection functions. Here, we propose an alternative strategy based on the extra knowledge that $\boldsymbol{F}^R$ are $\beta$-log-concave.

*Proof of Proposition 5.1.* **Step 1:** Uniform bound of the variance. Let us define the center of mass and the variance

$$\boldsymbol{\mu}_R := \int_{\mathbb{R}} x \, \boldsymbol{F}^R(x)\, dx \quad \text{and} \quad \boldsymbol{\sigma}_R^2 := \int_{\mathbb{R}} (x-\boldsymbol{\mu}_R)^2 \, \boldsymbol{F}^R(x)\, dx,$$

for any $R > 0$. Since each eigenfunction $\boldsymbol{F}^R$ is $\beta$-log-concave, a straightforward application of the Brascamp–Lieb inequality shows that variances $\boldsymbol{\sigma}_R^2$ satisfy

$$\boldsymbol{\sigma}_R^2 \le \frac{1}{\beta} \tag{5-2}$$

for any $R > 0$; see [Brascamp and Lieb 1976, Theorem 4.1]. Then, in order to control the (noncentered) second-order moments, we actually need to find a bound of the center of mass $\boldsymbol{\mu}_R$.

**Step 2:** Uniform bound of the center of mass. Assume that $\{\boldsymbol{\mu}_R\}_{R>0}$ is unbounded by contradiction. Changing variables $x$ with $-x$ if necessary, we may assume without loss of generality that $\boldsymbol{\mu}_R \nearrow +\infty$ as $R \nearrow +\infty$ up to an appropriate subsequence, which we denote in the same way for simplicity of notation. Note that integrating (4-4) against $e^{m_R(x)}$ and remarking that $\int_{\mathbb{R}} \mathcal{B}[\boldsymbol{F}^R](x)\, dx = \int_{\mathbb{R}} \boldsymbol{F}^R(x)\, dx = 1$ (where $\mathcal{B}$ is given in (1-3)) we obtain

$$A_R B_R = 1 \tag{5-3}$$

for every $R > 0$, where each factor reads

$$A_R := \int_{\mathbb{R}} e^{m_R(x)} \, \boldsymbol{F}^R(x)\, dx, \qquad B_R := \int_{\mathbb{R}^2} \phi^R\big(\tfrac{1}{2}(x_1+x_2)\big) \, \boldsymbol{F}^R(x_1)\, \boldsymbol{F}^R(x_2)\, dx_1\, dx_2,$$

and $\phi^R := G * e^{-m_R}$. By Chebyshev's inequality we know that

$$\int_{|x-\boldsymbol{\mu}_R|\leq\sqrt{2}\,\boldsymbol{\sigma}_R} \boldsymbol{F}^R(x)\,dx \geq \tfrac{1}{2} \tag{5-4}$$

for all $R > 0$. Therefore, noting that $m$ is nondecreasing in $\mathbb{R}_+$ by virtue of the hypotheses (H1)–(H2) we obtain the lower bound

$$A_R \geq \int_{|x-\boldsymbol{\mu}_R|\leq\sqrt{2}\,\boldsymbol{\sigma}_R} e^{m_R(x)}\,\boldsymbol{F}^R(x)\,dx \geq \tfrac{1}{2} \min_{|x-\boldsymbol{\mu}_R|\leq\sqrt{2}\,\boldsymbol{\sigma}_R} e^{m(x)} = \tfrac{1}{2} e^{m(\boldsymbol{\mu}_R-\sqrt{2}\,\boldsymbol{\sigma}_R)} \tag{5-5}$$

for large enough $R > 0$ so that $[\boldsymbol{\mu}_R - \sqrt{2}\,\boldsymbol{\sigma}_R, \boldsymbol{\mu}_R + \sqrt{2}\,\boldsymbol{\sigma}_R] \subset \mathbb{R}_+$. Similarly, using (5-4) and noting that $\phi^R$ is nonincreasing at the right of its maximizer (by strong log-concavity, see Lemma 2.2) we obtain

$$\begin{aligned} B_R &\geq \iint_{|x_i-\boldsymbol{\mu}_R|\leq\sqrt{2}\sigma_R} \phi^R\big(\tfrac{1}{2}(x_1+x_2)\big)\,\boldsymbol{F}^R(x_1)\,\boldsymbol{F}^R(x_2)\,dx_1\,dx_2 \\ &\geq \tfrac{1}{4} \min_{|x-\boldsymbol{\mu}_R|\leq\sqrt{2}\,\boldsymbol{\sigma}_R} \phi^R(x) \geq \tfrac{1}{4}\phi^R(\boldsymbol{\mu}_R + \sqrt{2}\,\boldsymbol{\sigma}_R) \end{aligned} \tag{5-6}$$

for large enough $R > 0$ so that $[\boldsymbol{\mu}_R - \sqrt{2}\,\boldsymbol{\sigma}_R, \boldsymbol{\mu}_R + \sqrt{2}\,\boldsymbol{\sigma}_R]$ lies in that region of the domain. Note that the above can be obtained if $R > 0$ is large enough since $\boldsymbol{\mu}^R - \sqrt{2}\,\boldsymbol{\sigma}_R \to \infty$ by assumptions, but the maximizers of $\phi^R$ must converge to the maximizer of $\phi$, which is a fixed number in the real line. Multiplying (5-5) and (5-6) yields the lower bound

$$A_R B_R \geq \tfrac{1}{8} e^{m_R(\boldsymbol{\mu}_R-\sqrt{2}\,\boldsymbol{\sigma}_R)} (G * e^{-m_R})(\boldsymbol{\mu}_R + \sqrt{2}\,\boldsymbol{\sigma}_R) \tag{5-7}$$

for large enough $R > 0$. Lemma B.2 provides an explicit lower bound (B-6) on Gaussian convolutions. Therefore, applying it to the second factor in (5-7) with the choices

$$f = e^{-m}, \quad \gamma = \alpha, \quad x_0 = \boldsymbol{\mu}_R, \quad \delta = \sqrt{2}\,\boldsymbol{\sigma}_R$$

implies the lower bound

$$\begin{aligned} A_R B_R &\geq G(2\sqrt{2}\,\boldsymbol{\sigma}_R) \int_0^{\frac{\alpha}{\alpha+1}\boldsymbol{\mu}_R - \frac{\sqrt{2}\boldsymbol{\sigma}_R}{\alpha+1}} \exp\big(\tfrac{1}{2}(\alpha+1)z^2\big)\,dz \\ &\geq G\bigg(\frac{2\sqrt{2}}{\sqrt{\beta}}\bigg) \int_0^{\frac{\alpha}{\alpha+1}\boldsymbol{\mu}_R - \frac{\sqrt{2}}{\sqrt{\beta}(\alpha+1)}} \exp\big(\tfrac{1}{2}(\alpha+1)z^2\big)\,dz, \end{aligned} \tag{5-8}$$

where in the last line we have used the bound (5-2) of variances. Since the left-hand side in (5-8) diverges as $R \to \infty$ because $\boldsymbol{\mu}_R \to +\infty$, we reach a contradiction with (5-3), and this ends the proof. □

**Theorem 5.2** (existence of quasiequilibria). *Under the assumptions in Theorem 4.1, let us consider the unique eigenpair* $(\boldsymbol{\lambda}^R, \boldsymbol{F}^R)$ *of* (4-4) *for any* $R > 0$. *Then, there exist* $\boldsymbol{\lambda} \in \mathbb{R}$ *and* $\boldsymbol{F} \in L^1_+(\mathbb{R}) \cap C^\infty(\mathbb{R})$ *which is* $\beta$*-log-concave* (*with* $\beta$ *given in* (1-7)) *such that*

$$\boldsymbol{\lambda}^R \to \boldsymbol{\lambda}, \quad \boldsymbol{F}^R \to \boldsymbol{F}, \quad \text{as } R \to \infty,$$

*up to subsequence, both pointwise and in any space* $(\mathcal{P}_p(\mathbb{R}), W_p)$ *with* $1 \leq p < 2$. *Moreover, the pair* $(\boldsymbol{\lambda}, \boldsymbol{F})$ *is the unique solution to* (1-5) *among all pairs* $(\lambda, F)$ *satisfying* (1-8).

*Proof.* **Step 1:** Existence via limit as $R \to \infty$. Let us notice that by (5-1) in Proposition 5.1 we have that $\{\boldsymbol{F}^R\}_{R>0}$ is a uniformly tight sequence of probability measures. Therefore, by Prokhorov's theorem there must exist $R_n \nearrow \infty$ and some limiting probability measure $\boldsymbol{F} \in \mathcal{P}(\mathbb{R})$ such that

$$\boldsymbol{F}^{R_n} \to \boldsymbol{F} \quad \text{narrowly in } \mathcal{P}(\mathbb{R}). \tag{5-9}$$

By integration on (4-4) we also obtain that

$$\boldsymbol{\lambda}^{R_n} = \iint_{\mathbb{R}^2} (e^{-m_{R_n}} * G)\big(\tfrac{1}{2}(x_1 + x_2)\big)\, \boldsymbol{F}^{R_n}(x_1)\, \boldsymbol{F}^{R_n}(x_2)\, dx_1\, dx_2,$$

and then we can pass to the limit as $n \to \infty$ in the eigenvalues too. Specifically, since $e^{-m_R} \to e^{-m}$ in $L^\infty(\mathbb{R})$, we have $e^{-m_R} * G \to e^{-m} * G$ in $C_b(\mathbb{R})$, and therefore by (5-9) we obtain

$$\boldsymbol{\lambda}^{R_n} \to \boldsymbol{\lambda} \tag{5-10}$$

as $n \to \infty$, where $\boldsymbol{\lambda}$ is given by

$$\boldsymbol{\lambda} := \iint_{\mathbb{R}^2} (e^{-m} * G)\big(\tfrac{1}{2}(x_1 + x_2)\big)\, \boldsymbol{F}(x_1)\, \boldsymbol{F}(x_2)\, dx_1\, dx_2 = \int_{\mathbb{R}} \mathcal{T}[\boldsymbol{F}](x)\, dx. \tag{5-11}$$

Putting (5-9) and (5-10) together and taking limits as $n \to \infty$ in (4-4) implies that $\{\boldsymbol{F}^{R_n}\}_{n\in\mathbb{N}}$ must also converge pointwise to some other limit $\widetilde{\boldsymbol{F}} \in L^1_+(\mathbb{R})$ by Fatou's lemma. Note that since $\boldsymbol{F}^R$ are all $\beta$-log-concave, their pointwise limit $\widetilde{\boldsymbol{F}}$ must be also. Indeed, note that we further have

$$\boldsymbol{\lambda}\, \widetilde{\boldsymbol{F}}(x) = \mathcal{T}[\boldsymbol{F}](x), \quad x \in \mathbb{R}, \tag{5-12}$$

and therefore, $\widetilde{\boldsymbol{F}} \in L^1_+(\mathbb{R}) \cap \mathcal{P}(\mathbb{R})$, in view of (5-11). Then, we actually have $\boldsymbol{F}^{R_n} \to \widetilde{\boldsymbol{F}}$ in $L^1(\mathbb{R})$ (thus narrowly in $\mathcal{P}(\mathbb{R})$) by Scheffé's lemma. Since $\boldsymbol{F}$ is a narrow limit of the same sequence, we have $\widetilde{\boldsymbol{F}} = \boldsymbol{F}$ and by (5-12) we obtain that $(\boldsymbol{\lambda}, \boldsymbol{F})$ must satisfy the initial problem (1-5). Let us also emphasize that we indeed have convergence in any $L^p$ Wasserstein space with $1 \le p < 2$ because all the $p$-th order moments with $1 \le p < 2$ are uniformly integrable by (5-1); see [Ambrosio et al. 2008, Proposition 7.1.5].

**Step 2:** Uniqueness of quasiequilibria. Note that several different convergent subsequences of $\{\boldsymbol{F}^R\}_{R>0}$ in Step 1 could give rise to various eigenpairs $(\boldsymbol{\lambda}, \boldsymbol{F})$ of (1-5). Whilst the global uniqueness is unclear with this method, we prove that there can only exist one solution to (1-5) among the pairs $(\lambda, F)$ satisfying (1-8). We exploit the one-step contraction property in Theorem 1.2(ii). Specifically, assume that $(\lambda, F)$ is any other solution to (1-5) and define $F_n(x) = \lambda^n F(x)$, which is clearly a solution to the evolution problem (1-1) with initial datum $F_0 \in L^1_+(\mathbb{R}) \cap C^1(\mathbb{R})$ satisfying the hypothesis (H3) by virtue of the assumption (1-8). Then, (1-9) implies

$$\left\| \frac{d}{dx}\left(\log \frac{F}{\boldsymbol{F}}\right) \right\|_{L^\infty} \le \frac{2}{1 + 2\beta} \left\| \frac{d}{dx}\left(\log \frac{F}{\boldsymbol{F}}\right) \right\|_{L^\infty}.$$

Again, since $\frac{2}{1+2\beta} < 1$ by Remark 1.7, we obtain that $F/\boldsymbol{F}$ must be constant. Since both $F$ and $\boldsymbol{F}$ are normalized probability densities, then we necessarily have that $F = \boldsymbol{F}$ (and therefore $\lambda = \boldsymbol{\lambda}$). □

## 6. Convergence to equilibrium for restricted initial data

In this section, we prove asynchronous exponential growth as claimed in Theorem 1.2(iii). More specifically, we show that for restricted initial data the asymptotic behavior of the rate of growth of the mass $\|F_n\|_{L^1}/\|F_{n-1}\|_{L^1}$ and the normalized profiles $F_n/\|F_n\|_{L^1}$ is dictated by the solution $(\boldsymbol{\lambda}, \boldsymbol{F})$ of the eigenproblem (1-5) obtained in Theorem 1.2(i). We derive the relaxation of the normalized profiles under the relative entropy metric. Our starting point is the one-step contraction property of the $L^\infty$ relative Fisher information in Theorem 1.2(ii) and the following version of the logarithmic-Sobolev inequality with respect to strongly log-concave densities, which relate the $(L^2)$ relative Fisher information and the relative entropy.

**Proposition 6.1** (logarithmic-Sobolev inequality). *Consider any pair $P, Q \in L^1_+(\mathbb{R}) \cap \mathcal{P}(\mathbb{R})$ such that $Q$ is $\gamma$-log-concave for some $\gamma > 0$. Then, we have*

$$\mathcal{D}_{KL}(P\,\|\,Q) \le \frac{1}{2\gamma}\,\mathcal{I}_2(P\,\|\,Q) \le \frac{1}{2\gamma}\,\mathcal{I}^2_\infty(P\,\|\,Q), \tag{6-1}$$

*where $\mathcal{D}_{\mathrm{KL}}$ is the relative entropy* (1-12), *$\mathcal{I}_2$ is the usual* (*or $L^2$*) *relative Fisher information* (1-18), *and $\mathcal{I}_\infty$ is the $L^\infty$ relative Fisher information* (1-6).

On the one hand, the first part of the inequality (6-1) amounts to the usual logarithmic-Sobolev inequality with respect to a strongly log-concave measure; see Corollary 5.7.2 and Section 9.3.1 in [Bakry et al. 2014] for details. On the other hand, the second part of the inequality readily holds by definition. Therefore, putting Theorem 1.2(ii) and Proposition 6.1 together, we end the proof of Theorem 1.2(iii).

*Proof of Theorem 1.2*(iii). By iterating $n$ times the one-step contraction property in Theorem 1.2(ii) and using the logarithmic-Sobolev inequality (6-1) in Proposition 6.1 we obtain

$$\mathcal{D}_{\mathrm{KL}}\left(\frac{F_n}{\|F_n\|_{L^1}}\,\middle\|\,\boldsymbol{F}\right) \le C_1\left(\frac{2}{1+2\beta}\right)^{2n} \tag{6-2}$$

for every $n \in \mathbb{N}$, where the constant $C_1$ reads

$$C_1 := \frac{1}{2\gamma}\mathcal{I}^2_\infty(F_0\,\|\,\boldsymbol{F}),$$

and it is finite by the assumption (H3). This proves the relaxation of the normalized profiles towards $\boldsymbol{F}$ in the relative entropy sense. Regarding the rate of growth, we note that

$$\frac{\|F_n\|_{L^1}}{\|F_{n-1}\|_{L^1}} = \iint_{\mathbb{R}^2} \phi\left(\frac{x_1+x_2}{2}\right)\frac{F_{n-1}(x_1)}{\|F_{n-1}\|_{L^1}}\,\frac{F_{n-1}(x_2)}{\|F_{n-1}\|_{L^1}}\,dx_1\,dx_2 \tag{6-3}$$

$$\boldsymbol{\lambda} = \iint_{\mathbb{R}^2} \phi\left(\frac{x_1+x_2}{2}\right)\boldsymbol{F}(x_1)\,\boldsymbol{F}(x_2)\,dx_1\,dx_2. \tag{6-4}$$

where $(\boldsymbol{\lambda}, \boldsymbol{F})$ is the solution to (1-5) in Theorem 1.2(i), and $\phi := G * e^{-m}$ again. Taking the difference of the two identities (6-3) and (6-4) above, we achieve

$$\begin{aligned}\left|\frac{\|F_n\|_{L^1}}{\|F_{n-1}\|_{L^1}} - \boldsymbol{\lambda}\right| &\le \|\phi\|_{L^\infty}\left\|\frac{F_{n-1}}{\|F_{n-1}\|_{L^1}} \otimes \frac{F_{n-1}}{\|F_{n-1}\|_{L^1}} - \boldsymbol{F}\otimes\boldsymbol{F}\right\|_{L^1}\\ &\le \|\phi\|_{L^\infty}\sqrt{\frac{1}{2}\mathcal{D}_{\mathrm{KL}}\left(\frac{F_{n-1}}{\|F_{n-1}\|_{L^1}} \otimes \frac{F_{n-1}}{\|F_{n-1}\|_{L^1}}\,\middle\|\,\boldsymbol{F}\otimes\boldsymbol{F}\right)}\\ &= \|\phi\|_{L^\infty}\sqrt{\mathcal{D}_{\mathrm{KL}}\left(\frac{F_{n-1}}{\|F_{n-1}\|_{L^1}}\,\middle\|\,\boldsymbol{F}\right)} \le C_2\left(\frac{2}{1+2\beta}\right)^n,\end{aligned}$$

with a explicit constant $C_2 > 0$ taking the form

$$C_2 := \|\phi\|_{L^\infty}\sqrt{C_1}.$$

Note that above, we have used successively Hölder's inequality, Pinsker's inequality, the tensorization property of the relative entropy, and (6-2) to reach the conclusion. □

## Appendix A: Intermediate dualities

For simplicity of the discussion, we do not present here the intermediate Kantorovich-type dualities in the case of nonlinear transition semigroups as in (2-4), but we rather focus on linear semigroups. More specifically, we have the following intermediate result which is reminiscent of the natural interpolation of Kantorovich duality for $L^1$ Wasserstein distance, and Lemma 2.4 for $L^\infty$ Wasserstein metric.

**Proposition A.1.** *Consider any $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ for some $1 \le p \le \infty$, and set any function $u \in C^1(\mathbb{R}^d)$ such that $u > 0$ and $\nabla(u^{1/p}) \in L^\infty(\mathbb{R}^d, \mathbb{R}^d)$. Then, the following inequality holds true for any $1 \le q \le \infty$, and $q'$ given by $\frac{1}{q} + \frac{1}{q'} = 1$:*

$$\left| \left( \int_{\mathbb{R}^d} u(x)\, \mu(dx) \right)^{1/p} - \left( \int_{\mathbb{R}^d} u(x)\, \nu(dx) \right)^{1/p} \right| \le \big\| \|\nabla(u^{1/p})\|_{q'} \big\|_{L^\infty} W_{p,q}(\mu, \nu).$$

*Here, $W_{p,q}$ denotes the $L^p$ Wasserstein distance associated with $\ell_q$ norm of $\mathbb{R}^d$, see* (1-24), *and we use the convention that $u^{1/\infty} = \log u$ for all $u > 0$.*

*Proof.* Let us consider any constant-speed geodesic $t \in [0, 1] \mapsto \rho_t \in \mathcal{P}_p(\mathbb{R}^d)$ in the Wasserstein space $(\mathcal{P}_p(\mathbb{R}^d), W_{p,q})$ joining $\mu$ to $\nu$. Specifically, $\rho$ satisfies the continuity equation

$$\begin{aligned} &\partial_t \rho_t + \operatorname{div}(\rho_t v_t) = 0, \quad t \in [0, 1],\ x \in \mathbb{R}^d, \\ &\rho_0 = \mu, \quad \rho_1 = \nu, \end{aligned} \tag{A-1}$$

in the distributional sense and, in addition, we have

$$\big\| \|v_t\|_q \big\|_{L^p(\rho_t)} = W_{p,q}(\mu, \nu), \quad t \in [0, 1]. \tag{A-2}$$

Let us also define the function

$$E(t) := \int_{\mathbb{R}^d} u(y)\, \rho_t(dy), \quad t \in [0, 1].$$

Since $\rho \in \operatorname{Lip}([0, 1], \mathcal{P}_p(\mathbb{R}^d))$, we have $E \in \operatorname{AC}([0, 1])$ and by the continuity equation (A-1) we have

$$\frac{dE}{dt}(t) = \int_{\mathbb{R}^d} \nabla u(y) \cdot v_t(y)\, \rho_t(dy) = p \int_{\mathbb{R}^d} \nabla(u^{1/p})(y) \cdot v_t(y) u^{1/p'}(y)\, \rho_t(dy) \tag{A-3}$$

for a.e. $t \in [0, 1]$, where we have used the identity $\nabla u = p \nabla(u^{1/p}) u^{1/p'}$. Therefore, we obtain

$$\begin{aligned} \left| \frac{dE}{dt}(t) \right| &\le p \int_{\mathbb{R}^d} \|\nabla(u^{1/p})(y)\|_{q'} \|v_t(y)\|_q u^{1/p'}(y)\, \rho_t(dy) \\ &\le p \big\| \|\nabla(u^{1/p})\|_{q'} \big\|_{L^\infty} \int_{\mathbb{R}^d} \|v_t(y)\|_q u^{1/p'}(y)\, \rho_t(dy) \\ &\le p \big\| \|\nabla(u^{1/p})\|_{q'} \big\|_{L^\infty} \big\| \|v_t\|_q \big\|_{L^p(\rho_t)} \|u^{1/p'}\|_{L^{p'}(\rho_t)} \end{aligned}$$

for a.e. $t \in [0,1]$, where in the first step we have used Hölder's inequality with the exponent $q$ applied to the inner product in the integrand of (A-3), and in the last step we have used Hölder's inequality with exponent $p$ applied to the integral of the second line. Using the constant-speed condition (A-2) in the second factor, and $\|u^{1/p'}\|_{L^{p'}(\rho_t)} = E(t)^{1/p'}$ in the last one, we obtain the relation

$$\left|\frac{dE}{dt}(t)\right| \leq p \big\| \|\nabla(u^{1/p})\|_{q'} \big\|_{L^\infty} W_{p,q}(\mu,\nu) E(t)^{1/p'}$$

for a.e. $t \in [0,1]$, which amounts to

$$\left|\frac{dE^{1/p}}{dt}(t)\right| \leq \big\| \|\nabla(u^{1/p})\|_{q'} \big\|_{L^\infty} W_{p,q}(\mu,\nu)$$

for a.e. $t \in [0,1]$. Integrating between 0 and 1 implies

$$|E(0)^{1/p} - E(1)^{1/p}| \leq \big\| \|\nabla(u^{1/p})\|_{q'} \big\|_{L^\infty} W_{p,q}(\mu,\nu).$$

Then, noting that $E(0) = \int_{\mathbb{R}^d} u(x)\,\mu(dx)$ and $E(1) = \int_{\mathbb{R}^d} u(x)\,\nu(dx)$ ends the proof. □

As a consequence, we obtain the following result, which allows identifying the Lipschitz constant of a function with the Lipschitz constant of an associated nonlinear functional over $\mathcal{P}_p(\mathbb{R}^d)$.

**Corollary A.2.** *Consider any $1 \leq p \leq \infty$, set any $v \in C^1(\mathbb{R}^d)$ with $\nabla v \in L^\infty(\mathbb{R}^d,\mathbb{R}^d)$, and assume that $v > 0$ when $p < \infty$ but not necessarily when $p = \infty$. Define the functional $\Phi_{p,v} : \mathcal{P}_p(\mathbb{R}^d) \to \mathbb{R}$ by*

$$\Phi_{p,v}[\mu] := \begin{cases} \left(\int_{\mathbb{R}^d} v(x)^p\,\mu(dx)\right)^{1/p} & \text{if } p < \infty, \\ \log\left(\int_{\mathbb{R}^d} e^{v(x)}\,\mu(dx)\right) & \text{if } p = \infty, \end{cases}$$

*for any $\mu \in \mathcal{P}_p(\mathbb{R}^d)$. Then, for any $1 \leq q \leq \infty$ the following identify holds true*:

$$\big\| \|\nabla v\|_{q'} \big\|_{L^\infty} = \sup_{\mu,\nu \in \mathcal{P}_p(\mathbb{R}^d)} \frac{\Phi_{p,v}[\mu] - \Phi_{p,v}[\nu]}{W_{p,q}(\mu,\nu)}.$$

*Proof.* First, note that the change of variable $v = u^{1/p}$ and Proposition A.1 readily imply

$$\big\| \|\nabla v\|_{q'} \big\|_{L^\infty} \geq \sup_{\mu,\nu \in \mathcal{P}_p(\mathbb{R}^d)} \frac{\Phi_{p,v}[\mu] - \Phi_{p,v}[\nu]}{W_{p,q}(\mu,\nu)}.$$

On the other hand, also note that by particularizing the measures $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ to be Dirac masses at respective points $x, x' \in \mathbb{R}^d$ we obtain

$$\sup_{\mu,\nu \in \mathcal{P}_p(\mathbb{R}^d)} \frac{\Phi_{p,v}[\mu] - \Phi_{p,v}[\nu]}{W_{p,q}(\mu,\nu)} \geq \sup_{x,x' \in \mathbb{R}^d} \frac{\Phi_{p,v}[\delta_x] - \Phi_{p,v}[\delta_{x'}]}{W_{p,q}(\delta_x,\delta_{x'})} = \sup_{x,x' \in \mathbb{R}^d} \frac{v(x) - v(x')}{\|x - x'\|_q} = \big\| \|\nabla v\|_{q'} \big\|_{L^\infty}.$$

This proves the converse inequality and then the above identity holds. □

## Appendix B: Lower bound of Gaussian convolution of log-concave densities

We present a technical result which computes an explicit lower bound on the convolution of a Gaussian density and any strongly log-concave probability density.

**Lemma B.1** (lower bound I). *Consider any* $f = e^{-V} \in L^1_+(\mathbb{R}) \cap \mathcal{P}(\mathbb{R})$, *such that* $V \in C^1(\mathbb{R})$ *with* $V'(0) = 0$, *and* $f$ *is* $\gamma$*-log-concave for some* $\gamma > 0$. *Then, we have*

$$(G * f)(x_0 + \delta) \geq G(2\delta) f(x_0 - \delta) \int_0^{\frac{\gamma}{\gamma+1}x_0 - \frac{\delta}{\gamma+1}} \exp\left(\frac{\gamma+1}{2} z^2\right) dz \tag{B-1}$$

*for any* $\delta > 0$ *and each* $x_0 > \frac{\gamma+2}{\gamma}\delta$, *where G denotes the standard Gaussian distribution* (1-4).

*Proof.* For simplicity of notation, we define $x_\pm := x_0 \pm \delta$ and we note that we can write

$$(G * f)(x_+) = \frac{1}{(2\pi)^{1/2}} f(x_-) \int_{\mathbb{R}} e^{V(x_-) - U(x)} \, dx, \tag{B-2}$$

where the function $U : \mathbb{R} \to \mathbb{R}$ is defined by

$$U(x) := V(x) + \tfrac{1}{2}(x - x_+)^2, \quad x \in \mathbb{R}.$$

Since the potential $V$ is $\gamma$ convex, we have that the potential $U$ is $(\gamma+1)$-convex. By the convexity inequality applied to the pair of points $(x, x_-)$ we then obtain

$$U(x_-) \geq U(x) + U'(x)(x_- - x) + \frac{\gamma+1}{2}(x_- - x)^2 \tag{B-3}$$

for any $x \in \mathbb{R}$. Consider the unique minimizer $x_* \in \mathbb{R}$ of the potential $U$. Since in particular $x_*$ is a critical point of $U$, we have

$$0 = U'(x_*) = V'(x_*) + (x_* - x_+).$$

Multiplying above by $x^*$, using that $V'(0) = 0$ by hypothesis along with the convexity inequality of $V$ applied at the pair $(x_*, 0)$, we infer $\gamma x_*^2 \leq (x_+ - x_*)x_*$, and therefore

$$|x_*| \leq \frac{1}{\gamma+1} x_+. \tag{B-4}$$

Since $U'(x) > 0$ for $x > x_*$ and $x_- - x > 0$ for $x < x_-$, (B-3) implies

$$U(x_-) \geq U(x) + \frac{\gamma+1}{2}(x_- - x)^2$$

for any $x \in (x_*, x_-)$. Let us note that indeed we have the appropriate ordering $x_* < x_-$ since by (B-4) and the assumption $x_0 > \frac{\gamma+2}{\gamma}\delta$ we obtain

$$x_* \leq \frac{1}{\gamma+1} x_+ = \frac{1}{\gamma+1}(x_0 + \delta) \leq x_0 - \delta = x_-.$$

Writing everything in terms of $V$ implies

$$V(x_-) - U(x) \geq -\frac{1}{2}(x_- - x_+)^2 + \frac{\gamma+1}{2}(x_- - x)^2 \tag{B-5}$$

for any $x \in (x_*, x_-)$. Injecting (B-5) into (B-2) we obtain

$$(G * f)(x_+) \geq G(x_+ - x_-) f(x_-) \int_{x_*}^{x_-} \exp\left(\frac{\gamma+1}{2}(x_- - x)^2\right) dx.$$

Of course, the above implies (B-1) by a simple change of variables $z = x_- - x$, and noting again that

$$x_- - x_* \geq x_- - \frac{1}{\gamma+1} x_+ = (x_0 - \delta) - \frac{1}{\gamma+1}(x_0 + \delta) = \frac{\gamma}{\gamma+1} x_0 - \frac{\gamma+2}{\gamma+1}\delta,$$

thanks to (B-4), which yields again positive a positive upper bound by the assumption $x_0 > \frac{\gamma+2}{\gamma}\delta$. □

Note that arguing along the same lines, we can prove an analogous result where the above positive strongly log-concave density $f$ is replaced by its truncation $f_R$ to intervals $I_R := (-R, R)$. Specifically, anything that we need to guarantee is that $[x_*, x_-] \subset I_R$. First, note that $x_- < R$ amounts to the condition $x_0 < R + \delta$. Second, by (B-4) we obtain that $x_* > -R$ as long as $\frac{1}{\gamma+1} x_+ < R$, which amounts to the condition $x_0 < (\gamma+1)R - \delta$. If we take $R$ large enough (namely $R > 2\delta/\gamma$) then we have that the former condition on $x_0$ is the most restrictive. Therefore, we have the following result.

**Lemma B.2** (lower bound II). *Under the assumptions in Lemma B.1, let us define*

$$f_R(x) := e^{-V_R(x)}, \qquad x \in \mathbb{R},$$
$$V_R(x) := V(x) + \chi_{\bar{I}_R}(x), \quad x \in \mathbb{R},$$

*for any $R > 0$, where $\chi_{\bar{I}_R}$ is the characteristic function associated to $\bar{I}_R$ (see (1-23)). Then, we have*

$$(G * f_R)(x_0 + \delta) \geq G(2\delta) f_R(x_0 - \delta) \int_0^{\frac{\gamma}{\gamma+1}x_0 - \frac{\delta}{\gamma+1}} \exp\left(\frac{\gamma+1}{2} z^2\right) dz \tag{B-6}$$

*for any $\delta > 0$, each $\frac{\gamma+2}{\gamma}\delta < x_0 < R + \delta$, and every $R > \frac{2\delta}{\gamma}$.*

## Appendix C: Euclidean estimates on the displacement of the Brenier map between perturbations of log-concave measures

In this section we present a proof of the uniform bound of the $\ell_2$ norm on the displacement of the Brenier map between perturbations of log-concave measures.

**Lemma C.1.** *Consider two densities $f, g \in L^1_+(\mathbb{R}^d) \cap \mathcal{P}(\mathbb{R}^d)$, assume that*

$$\{z \in \mathbb{R}^d : f(z) > 0\} = \{z \in \mathbb{R}^d : g(z) > 0\} = \bar{B}_R,$$

*where $B_R$ is the Euclidean ball, and suppose that $f = e^{-W}$, $g = e^{-\widetilde{W}}$ are $\gamma$-log-concave for some $\gamma > 0$ and $f, g \in C^{1,\delta}(\bar{B}_R)$ for some $\delta > 0$. Let $T = \nabla\phi : \bar{B}_R \to \bar{B}_R$ be the Brenier map from $f$ to $g$. Then,*

$$W_{\infty,2}(f, g) \leq \big\| \|T - I\|_2 \big\|_{L^\infty(\bar{B}_R)} \leq \frac{1}{\gamma} \big\| \|\nabla(W - \widetilde{W})\|_2 \big\|_{L^\infty(\bar{B}_R)}.$$

As mentioned in Remark 3.1, this result is not enough for the sake of this paper, but was the starting point to prove Lemma 2.6. The technique to prove it is essentially based on the computations in [Ferrari and Santambrogio 2021], but we provide the proof here since the statement is not a direct consequence of it. On the other hand, this very result has its own interest, as one can see from the recent paper [Khudiakova et al. 2024].

*Proof of Lemma C.1.* Since $f, g \in C^{1,\delta}(\bar{B}_R)$ are bounded below on $B_R$ by a positive constant, $f = g = 0$ outside $B_R$, and $B_R$ is uniformly convex, Caffarelli's theory [1996] proves that $T \in C^{2,\delta}(\bar{B}_R)$. We consider $T(z) - z = \nabla\psi(z)$, where $\psi(z) = \phi(z) - \frac{1}{2}\|z\|_2^2$. The function $\psi$ solves the Monge–Ampère equation, which we write in logarithmic form:

$$\log\det(D^2\psi(z) + I) = \widetilde{W}(\nabla\psi(z) + z) - W(z), \quad z \in \mathbb{R}^d. \tag{C-1}$$

Taking partial derivatives $\partial_{x_k}$ in (C-1) we have

$$\operatorname{tr}((D^2\phi)^{-1}\partial_{x_k} D^2\psi) = \nabla\widetilde{W}(\nabla\psi + z)\cdot\partial_{x_k}\nabla\psi + (\nabla\widetilde{W}(\nabla\psi + z) - \nabla W)\cdot e_k, \quad z \in \mathbb{R}^d,$$

for $1 \le k \le d$. We then multiply by $\partial_{x_k}\psi$ and sum over $k$, so that we obtain

$$\operatorname{tr}\biggl((D^2\phi)^{-1}\sum_k \partial_{x_k} D^2\psi\,\partial_{x_k}\psi\biggr) = \nabla\widetilde{W}(\nabla\psi+z)\cdot\partial_{x_k}\bigl(\tfrac{1}{2}\|\nabla\psi\|_2^2\bigr) + (\nabla\widetilde{W}(\nabla\psi+z) - \nabla W)\cdot\nabla\psi(z), \quad z \in \mathbb{R}^d.$$

We now consider the point $z^* \in \bar{B}_R$ which maximizes $\frac{1}{2}\|\nabla\psi\|_2^2$, which is also the maximum point for the displacement $\|T - I\|_2$. Such a point exists since the ball $\bar{B}_R$ is compact. Moreover, [Ferrari and Santambrogio 2021, Lemma 3.1] shows that such a maximum cannot be attained on the boundary $\partial B_R$. Hence, we can apply first- and second-order optimality conditions. In particular, we have $\partial_{x_k}\bigl(\frac{1}{2}\|\nabla\psi\|_2^2\bigr)(z^*) = 0$ and the Hessian matrix $D^2\bigl(\frac{1}{2}\|\nabla\psi\|_2^2\bigr)(z^*)$ has to be negative-definite, i.e.,

$$\sum_k \partial_{x_k} D^2\psi(z^*)\partial_{x_k}\psi(z^*) + (D^2\psi(z^*))^2 \le 0.$$

Using the fact that $(D^2\psi(z^*))^2$ is the square of a symmetric matrix, and hence is negative, we obtain that $\sum_k \partial_{x_k} D^2\psi(z^*)\partial_{x_k}\psi(z^*)$ is itself negative definite, and the trace of its product times $(D^2\phi)^{-1}$ is also negative. This allows to obtain

$$(\nabla\widetilde{W}(\nabla\psi(z^*) + z^*) - \nabla W(z^*)\cdot\nabla\psi(z^*) \le 0,$$

which implies

$$\bigl(\nabla W(\nabla\psi(z^*) + z^*) - \nabla W(z^*)\bigr)\cdot\nabla\psi(z^*) \le \|\nabla(\widetilde{W} - W)\|_{L^\infty}\|\nabla\psi(z^*)\|_2,$$

and hence by $\gamma$-convexity of $W$ we have

$$\gamma\|\nabla\psi(z^*)\|_2^2 \le \|\nabla(\widetilde{W} - W)\|_{L^\infty}\|\nabla\psi(z^*)\|_2,$$

which ends the proof. □

Similarly to Lemma 2.6 for the $\ell_1$ norm of the displacement of the Brenier map, a more general result holds for strictly positive densities $f, g \in C^{1,\delta}_{\mathrm{loc}}(\mathbb{R}^d)$ supported in the full space $\mathbb{R}^d$.

**Corollary C.2.** *Consider two densities* $f, g \in L^1_+(\mathbb{R}^d) \cap \mathcal{P}(\mathbb{R}^d)$, *assume that* $f, g > 0$, *and suppose that* $f = e^{-W}$, $g = e^{-\widetilde{W}}$ *are* $\gamma$*-log-concave for some* $\gamma > 0$ *and* $f, g \in C^{1,\delta}_{\mathrm{loc}}(\mathbb{R}^d)$ *for some* $\delta > 0$. *Let* $T = \nabla\phi : \mathbb{R}^d \to \mathbb{R}^d$ *be the Brenier map from* $f$ *to* $g$. *Then*,

$$W_{\infty,2}(f, g) \leq \big\| \|T - I\|_2 \big\|_{L^\infty(\mathbb{R}^d} \leq \frac{1}{\gamma} \big\| \|\nabla(W - \widetilde{W})\|_2 \big\|_{L^\infty(\mathbb{R}^d)}.$$

The proof is similar to the one of Lemma 2.6 arguing by a truncation argument and applying the local version in Lemma C.1. Specifically, we truncate $W$ and $\widetilde{W}$ and accordingly $f$ and $g$ to an increasing sequence $B_R$ of Euclidean balls preserving the Lipschitz and convexity bounds. We obtain a sequence of optimal transport maps $T_R$ transporting the associated truncations $f_R$ onto $g_R$ and satisfying

$$\big\| \|T_R - I\|_2 \big\|_{L^\infty(\bar{B}_R)} \leq \frac{1}{\gamma} \|\nabla(W - \widetilde{W})\|_{L^\infty(\mathbb{R}^d)},$$

for all $R > 0$. Finally, we pass to the limit in the above estimate as $R \to \infty$.

## Acknowledgments

The authors are indebted to Thomas Lepoutre for stimulating discussions and valuable suggestions on the paper's presentation. We also thank Laurent Lafleche for pointing us to the identity in Corollary A.2. Vincent Calvez and David Poyato have received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 865711). David Poyato has received funding from the European Union's Horizon Europe research and innovation program under the Marie Skłodowska-Curie grant agreement no. 101064402, and partially from grant C-EXP-265-UGR23 funded by Consejería de Universidad, Investigación e Innovación & ERDF/EU Andalusia Program, from grant PID2022-137228OB-I00 funded by the Spanish Ministerio de Ciencia, Innovación y Universidades, MICIU/AEI/10.13039/501100011033 & "ERDF/EU A way of making Europe", and from Modeling Nature (MNAT) project QUAL21-011 funded by Junta de Andalucía. Vincent Calvez thanks the France 2030 framework programme Centre Henri Lebesgue ANR-11-LABX-0020-01 for creating an attractive mathematical environment. Filippo Santambrogio acknowledges the support of the Lagrange Mathematics and Computation Research Center project on Optimal Transportation.

## References

[Ambrosio et al. 2008] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd ed., Birkhäuser, Basel, 2008. MR Zbl

[Arnold et al. 2001] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter, "On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker–Planck type equations", *Comm. Partial Differential Equations* **26**:1-2 (2001), 43–100. MR Zbl

[Bakry 1994] D. Bakry, "L'hypercontractivité et son utilisation en théorie des semigroupes", pp. 1–114 in *Lectures on probability theory* (Saint-Flour, 1992), edited by P. Bernard, Lecture Notes in Math. **1581**, Springer, 1994. MR Zbl

[Bakry et al. 2014] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundl. Math. Wissen. **348**, Springer, 2014. MR Zbl

[Barles et al. 2009] G. Barles, S. Mirrahimi, and B. Perthame, "Concentration in Lotka–Volterra parabolic or integral equations: a general convergence result", *Methods Appl. Anal.* **16**:3 (2009), 321–340. MR Zbl

[Barton et al. 2017] N. H. Barton, A. M. Etheridge, and A. Véber, "The infinitesimal model: definition, derivation, and implications", *Theoret. Popul. Biol.* **118** (2017), 50–73. Zbl

[Berestycki et al. 2016] H. Berestycki, J. Coville, and H.-H. Vo, "Persistence criteria for populations with non-local dispersion", *J. Math. Biol.* **72**:7 (2016), 1693–1745. MR Zbl

[Brascamp and Lieb 1976] H. J. Brascamp and E. H. Lieb, "On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation", *J. Functional Analysis* **22**:4 (1976), 366–389. MR Zbl

[Brenier 1991] Y. Brenier, "Polar factorization and monotone rearrangement of vector-valued functions", *Comm. Pure Appl. Math.* **44**:4 (1991), 375–417. MR Zbl

[Caffarelli 1992a] L. A. Caffarelli, "Boundary regularity of maps with convex potentials", *Comm. Pure Appl. Math.* **45**:9 (1992), 1141–1151. MR Zbl

[Caffarelli 1992b] L. A. Caffarelli, "The regularity of mappings with a convex potential", *J. Amer. Math. Soc.* **5**:1 (1992), 99–104. MR Zbl

[Caffarelli 1996] L. A. Caffarelli, "Boundary regularity of maps with convex potentials, II", *Ann. of Math.* (2) **144**:3 (1996), 453–496. MR Zbl

[Caffarelli 2000] L. A. Caffarelli, "Monotonicity properties of optimal transportation and the FKG and related inequalities", *Comm. Math. Phys.* **214**:3 (2000), 547–563. Correction in **225**:2 (2002), 449–450. MR Zbl

[Calvez et al. 2019] V. Calvez, J. Garnier, and F. Patout, "Asymptotic analysis of a quantitative genetics model with nonlinear integral operator", *J. Éc. polytech. Math.* **6** (2019), 537–579. MR Zbl

[Calvez et al. 2024] V. Calvez, T. Lepoutre, and D. Poyato, "Ergodicity of the Fisher infinitesimal model with quadratic selection", *Nonlinear Anal.* **238** (2024), art. id. 113392. MR Zbl

[Colombo and Fathi 2021] M. Colombo and M. Fathi, "Bounds on optimal transport maps onto log-concave measures", *J. Differential Equations* **271** (2021), 1007–1022. MR Zbl

[Colombo et al. 2017] M. Colombo, A. Figalli, and Y. Jhaveri, "Lipschitz changes of variables between perturbations of log-concave measures", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **17**:4 (2017), 1491–1519. MR Zbl

[Dekens 2022] L. Dekens, "Evolutionary dynamics of complex traits in sexual populations in a heterogeneous environment: how normal?", *J. Math. Biol.* **84**:3 (2022), art. id. 15. MR Zbl

[Diekmann et al. 2005] O. Diekmann, P.-E. Jabin, S. Mischler, and B. Perthame, "The dynamics of adaptation: an illuminating example and a Hamilton–Jacobi approach", *Theoret. Popul. Biol.* **67**:4 (2005), 257–271. Zbl

[Ferrari and Santambrogio 2021] V. Ferrari and F. Santambrogio, "Lipschitz estimates on the JKO scheme for the Fokker–Planck equation on bounded convex domains", *Appl. Math. Lett.* **112** (2021), art. id. 106806. MR Zbl

[Fisher 1919] R. A. Fisher, "The correlation between relatives on the supposition of mendelian inheritance", *Trans. Royal Soc. Edinburgh* **52**:2 (1919), 399–433.

[Fisher 1922] R. A. Fisher, "On the mathematical foundations of theoretical statistics", *Philos. Trans. Royal Soc. London Ser. A* **222**:594-604 (1922), 309–368. Zbl

[Frouvelle and Taing 2025] A. Frouvelle and C. Taing, "On the Fisher infinitesimal model without variability", *J. Stat. Phys.* **192**:1 (2025), art. id. 9. MR Zbl

[Garnier et al. 2023] J. Garnier, O. Cotto, E. Bouin, T. Bourgeron, T. Lepoutre, O. Ronce, and V. Calvez, "Adaptation of a quantitative trait to a changing environment: new analytical insights on the asexual and infinitesimal sexual models", *Theoret. Popul. Biol.* **152** (2023), 1–22. Zbl

[Guerand et al. 2025] J. Guerand, M. Hillairet, and S. Mirrahimi, "A moment-based approach for the analysis of the infinitesimal model in the regime of small variance", *Kinet. Relat. Models* **18**:3 (2025), 389–425. MR Zbl

[Jhaveri 2019] Y. Jhaveri, "On the (in)stability of the identity map in optimal transportation", *Calc. Var. Partial Differential Equations* **58**:3 (2019), art.id. 96. MR Zbl

[Khudiakova et al. 2024] K. A. Khudiakova, J. Maas, and F. Pedrotti, "$L^\infty$-optimal transport of anisotropic log-concave measures and exponential convergence in Fisher's infinitesimal model", preprint, 2024. arXiv 2402.04151

[Kuwada 2010] K. Kuwada, "Duality on gradient estimates and Wasserstein controls", *J. Funct. Anal.* **258**:11 (2010), 3758–3774. MR Zbl

[Li et al. 2017] F. Li, J. Coville, and X. Wang, "On eigenvalue problems arising from nonlocal diffusion models", *Discrete Contin. Dyn. Syst.* **37**:2 (2017), 879–903. MR Zbl

[Mahadevan 2007] R. Mahadevan, "A note on a non-linear Krein–Rutman theorem", *Nonlinear Anal.* **67**:11 (2007), 3084–3090. MR Zbl

[Mirrahimi and Raoul 2013] S. Mirrahimi and G. Raoul, "Dynamics of sexual populations structured by a space variable and a phenotypical trait", *Theoret. Popul. Biol.* **84** (2013), 87–103. Zbl

[Nussbaum 1988] R. D. Nussbaum, *Hilbert's projective metric and iterated nonlinear maps*, Mem. Amer. Math. Soc. **391**, Amer. Math. Soc., Providence, RI, 1988. MR Zbl

[Nussbaum 1994] R. D. Nussbaum, "Finsler structures for the part metric and Hilbert's projective metric and applications to ordinary differential equations", *Differential Integral Equations* **7**:5-6 (1994), 1649–1707. MR Zbl

[Ollivier 2007] Y. Ollivier, "Ricci curvature of metric spaces", *C. R. Math. Acad. Sci. Paris* **345**:11 (2007), 643–646. MR Zbl

[Ollivier 2009] Y. Ollivier, "Ricci curvature of Markov chains on metric spaces", *J. Funct. Anal.* **256**:3 (2009), 810–864. MR Zbl

[Patout 2023] F. Patout, "The Cauchy problem for the infinitesimal model in the regime of small variance", *Anal. PDE* **16**:6 (2023), 1289–1350. MR Zbl

[Perthame and Barles 2008] B. Perthame and G. Barles, "Dirac concentrations in Lotka–Volterra parabolic PDEs", *Indiana Univ. Math. J.* **57**:7 (2008), 3275–3301. MR Zbl

[Raoul 2017] G. Raoul, "Macroscopic limit from a structured population model to the Kirkpatrick–Barton model", preprint, 2017. arXiv 1706.04094

[Raoul 2021] G. Raoul, "Exponential convergence to a steady-state for a population genetics model with sexual reproduction and selection", preprint, 2021. arXiv 2104.06089

[von Renesse and Sturm 2005] M.-K. von Renesse and K.-T. Sturm, "Transport inequalities, gradient estimates, entropy, and Ricci curvature", *Comm. Pure Appl. Math.* **58**:7 (2005), 923–940. MR Zbl

[Saumard and Wellner 2014] A. Saumard and J. A. Wellner, "Log-concavity and strong log-concavity: a review", *Stat. Surv.* **8** (2014), 45–114. MR Zbl

[Stigler 2005] S. Stigler, "Fisher in 1921", *Statist. Sci.* **20**:1 (2005), 32–49. MR Zbl

[Villani 2003] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics **58**, Amer. Math. Soc., Providence, RI, 2003. MR Zbl

[Webb 1987] G. F. Webb, "An operator-theoretic formulation of asynchronous exponential growth", *Trans. Amer. Math. Soc.* **303**:2 (1987), 751–763. MR Zbl

Received 28 Apr 2023. Revised 15 May 2024. Accepted 20 Sep 2024.

VINCENT CALVEZ: vincent.calvez@math.cnrs.fr
*CNRS, Université de Bretagne Occidentale, UMR 6205, Laboratoire de Mathématiques de Bretagne Atlantique, Brest, France*

DAVID POYATO: davidpoyato@ugr.es
*Departamento de Matemática Aplicada and Research Unit "Modeling Nature" (MNat), Facultad de Ciencias, Universidad de Granada, Granada, Spain*

FILIPPO SANTAMBROGIO: santambrogio@math.univ-lyon1.fr
*Université Claude Bernard Lyon 1, CNRS, Ecole Centrale de Lyon, INSA Lyon, Université Jean Monnet, Institut Camille Jordan UMR5208, Villeurbanne, France*

ANALYSIS AND PDE
Vol. 18 (2025), No. 8, pp. 1875–1896
DOI: 10.2140/apde.2025.18.1875

# THE $L^\infty$ ESTIMATE FOR PARABOLIC COMPLEX MONGE–AMPÈRE EQUATIONS

QIZHI ZHAO

Following the recent developments in Chen and Cheng (2023) and Guo et al. (2023), we derive the $L^\infty$ estimate for Kähler–Ricci flows under certain integral assumptions. The technique also extends to some other parabolic Monge–Ampère equations derived from Kähler geometry and $G_2$ geometry.

## 1. Introduction

We will derive the $L^\infty$ estimate for the Kähler–Ricci flow

$$\begin{cases} \partial_t \varphi = \log\Big(\dfrac{{\omega_\varphi}^n}{e^{nF}{\omega_0}^n}\Big), \\ \varphi(\,\cdot\,,0) = \varphi_0(\,\cdot\,) \end{cases} \tag{1-1}$$

under the assumption that the $p$-entropy $\mathrm{Ent}_p(F) = \int_M |F|^p e^{nF}{\omega_0}^n$ is bounded and $\int_M F{\omega_0}^n$ has a lower bound. Here is our main theorem.

**Theorem 1.1.** *Let us consider the flow equation* (1-1) *on* $M \times [0,T)$, *where* $M$ *is an* $n$*-dimensional compact Kähler manifold. Let* $F$ *be a space function, i.e.,* $F : M \to \mathbb{R}$. *Assume, for some* $p > n+1$, *the* $p$*-entropy* $\mathrm{Ent}_p(F) = \int_M |F|^p e^{nF}{\omega_0}^n$ *is bounded and* $\int_M nF{\omega_0}^n \geq -K$. *Moreover, suppose* $\varphi$ *is a* $C^2$ *solution, and let* $\tilde\varphi = \varphi - \fint \varphi{\omega_0}^n$ *be a normalization which has the zero integral. Then we have the* $L^\infty$ *estimate*

$$\|\tilde\varphi\|_{L^\infty(M\times[0,T))} \leq C,$$

*where* $C$ *depends on* $n,\ \omega_0,\ \varphi_0,\ p,\ K$, *and* $\mathrm{Ent}_p(F)$. *Most importantly, such a* $C$ *does not depend on* $T$.

Yau [1978] applied the method of Moser iteration to derive the $L^\infty$ estimate for Monge–Ampère equations when $\|e^{nF}\|_{L^p}$ is bounded for some $p > n$. Later, Kołodziej [2003] gave another proof by using the pluripotential theory under a weaker assumption that $\|e^{nF}\|_{L^p}$ is bounded for some $p > 1$. More recently, Guo, Phong, and Tong [Guo et al. 2023] recovered Kołodziej's estimate by a PDE method which was partly motivated by the breakthrough on the cscK metric of Chen and Cheng [2021].

The Kähler–Ricci flow was firstly studied by Cao [1985] when he gave an alternative proof of Calabi's conjecture for $c_1(M) = 0$ and $c_1(M) < 0$, which investigated the estimates for the Kähler–Ricci flow instead of Monge–Ampère equations. There are abundant results on Kähler–Ricci flow, see [Eyssidieux et al. 2015; 2016; Guedj et al. 2021; Jian and Shi 2024]. Our result requires a weaker regularity on the right-hand side than Cao's $L^\infty$ estimate and can be viewed as a parabolic analogue of [Guo et al. 2023; Wang et al. 2021].

*MSC2020:* 53E30, 58J90.
*Keywords:* auxiliary equations, energy estimates, parabolic complex Monge–Ampère equations.

© 2025 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

There are some technical improvements in our paper compared with previous results in [Chen and Cheng 2023; Guo and Phong 2023; 2024; Guo et al. 2023]. The difficulty for the flow problem is that we want to derive an $L^\infty$ estimate independent of $T$. But the original auxiliary equation may not serve as a good choice. Our approach is to consider a local version of auxiliary flows instead, see Section 2. Compared with the elliptic version of $L^\infty$ estimates, our theorem requires an extra integral condition. Rewriting (1-1) by $\omega_\varphi^n = e^{\dot\varphi + nF}{\omega_0}^n$, we could see that the $L^\infty$ estimate of $\varphi$ comes from some $p$-entropy bounds on $e^{\dot\varphi+nF}$. Roughly speaking, we need not only the $p$-entropy bound controls on $e^{nF}$ but also some upper bounds on $\dot\varphi$. In Kähler–Ricci flow, Cao proved the supremum of $\dot\varphi$ can be controlled by the infimum of $e^{nF}$ when $F$ is smooth. Indeed such estimates can be generalized to general parabolic Monge–Ampère flows. However, in our theorem, we can improve the pointwise condition by some integral condition on $F$.

There are two directions to generalize Theorem 1.1. As in [Chen and Cheng 2023], we can replace the Monge–Ampère operator on the right-hand side by a more general nonlinear operator $\mathcal{F}$. Write $\omega_0 = \sqrt{-1}g_{j\bar m}\,\mathrm{d}z^j \wedge \mathrm{d}\bar z^m$ in local coordinate; then the corresponding endomorphism $h_\varphi$, which is relative to $\omega_\varphi$, can be expressed by $(h_\varphi)^j_k = g^{j\bar m}(\omega_\varphi)_{\bar m k}$ in local coordinate. Let $\lambda[h_\varphi]$ be the vector of eigenvalues of $h_\varphi$, and consider the nonlinear operator $\mathcal{F} : \Gamma \subset \mathbb{R}^n \to \mathbb{R}_+$ with the following four conditions:

(1) The domain $\Gamma$ is a symmetric cone with $\Gamma_n \subset \Gamma \subset \Gamma_1$, where $\Gamma_k$ is defined to be the cone of vectors $\lambda$ with $\sigma_j(\lambda) > 0$ for $1 \le j \le k$, where $\sigma_j$ is the $j$-th symmetric polynomial in $\lambda$.

(2) $\mathcal{F}(\lambda)$ is symmetric in $\lambda \in \Gamma$ and it is of homogeneous degree $r$.

(3) $\frac{\partial \mathcal{F}}{\partial \lambda_j} > 0$ for each $j = 1, \dots, n$ and $\lambda \in \Gamma$.

(4) There is a $\gamma > 0$ such that

$$\prod_{j=1}^n \frac{\partial \mathcal{F}(\lambda)}{\partial \lambda_j} \ge \gamma \mathcal{F}^{n(1-1/r)} \quad \text{for all } \lambda \in \Gamma. \tag{1-2}$$

The above requirements come from [Guo et al. 2023], and there is a slight modification on the last condition since the homogeneous degree of $\mathcal{F}$ is $r$ under our assumption. The complex Hessian operators and $p$-Monge–Ampère operators are examples. More examples can be found in [Harvey and Lawson 2023]. Here is our first generalization.

**Theorem 1.2.** *Let $\varphi$ be a $C^2$ solution of the flow*

$$\begin{cases} \partial_t \varphi = \log\Big(\dfrac{\mathcal{F}(\lambda[h_\varphi])}{e^{rF}}\Big), \\ \varphi(\,\cdot\,, 0) = \varphi_0 \end{cases} \tag{1-3}$$

*on $M \times [0, T)$, where $M$ is an $n$-dimensional compact Kähler manifold. Let $F : M \to \mathbb{R}$ be a space function. Assume, for some $p > n+1$, the $p$-entropy* $\mathrm{Ent}_p(F) = \int_M |F|^p e^{nF} {\omega_0}^n$ *is bounded and $\int_M F{\omega_0}^n \ge -K$. Then we have the $L^\infty$ estimate on the normalization $\tilde\varphi$*

$$\|\tilde\varphi\|_{L^\infty(M\times[0,T))} \le C,$$

*where $C$ depends on $n$, $\omega_0$, $\varphi_0$, $p$, $K$, $\gamma$, $r$, and* $\mathrm{Ent}_p(F)$.

Another direction of generalization is motivated by Chen and Cheng [2023], who considered the $L^\infty$ estimate for the inverse Monge–Ampère flow

$$\begin{cases}(-\partial_t u)\omega_\varphi^n = e^{nF}{\omega_0}^n,\\ \varphi(\,\cdot\,,0)=\varphi_0.\end{cases} \tag{1-4}$$

Indeed, we can consider the general complex Monge–Ampère flow

$$\begin{cases}\partial_t\varphi = \Theta\Big(\dfrac{{\omega_\varphi}^n}{e^{nF}{\omega_0}^n}\Big),\\ \varphi(\,\cdot\,,0)=\varphi_0,\end{cases} \tag{1-5}$$

where $\Theta:\mathbb{R}_+\to\mathbb{R}$ is a strictly increasing smooth function. Picard and Zhang [2020] proved the long time existence and convergence of the flow (1-5) under the assumption that $F\in C^\infty(M,\mathbb{R})$. It is the Kähler–Ricci flow when $\Theta(y)=\log y$ and the inverse Monge–Ampère flow (1-4) when $\Theta(y)=-1/y$. The general parabolic Monge–Ampère flow (1-5) also arises from many other geometric problems. For example, when $\Theta(y)=y$, this is the flow reduced from the anomaly flow with conformal Kähler initial data; see [Phong et al. 2019]. When $\Theta(y)=y^{1/3}$, this is the reduction of the $G_2$-Laplacian flow over a seven dimensional manifold [Picard and Suan 2024]. We can apply the new technique to prove the $L^\infty$ estimate of the solution $\varphi$ to the flow (1-5) under an analogue assumption on $F$.

**Theorem 1.3.** *Assume* $\Theta(y)=-1/y,\ y,$ *or* $y^{1/3}$, *and* $\mathrm{Ent}_p(F)$ *is bounded. Moreover, consider the constant $K$ equal to* $\max\big(0,\int_M\Theta(e^{-nF}){\omega_0}^n\big)$. *Then, there exists a constant $C$ depending on $n$, $\omega_0$, $\varphi_0$, $p$, $K$, and* $\mathrm{Ent}_p(F)$ *such that*

$$\|\tilde\varphi\|_{L^\infty(M\times[0,T))}\le C,$$

*where $\tilde\varphi$ is a normalization of a $C^2$ solution of the flow* (1-5).

Since $\Theta<0$ in the inverse Monge–Ampère equation, we have $K\equiv 0$, which means there is no extra condition for this case.

Going forward, a constant is called *universal* if it depends only on $n,\ \omega_0,\ \varphi_0,\ p,\ \gamma,\ r,\ K$, and $\mathrm{Ent}_p(F)$.

## 2. Auxiliary equations

In this section, we want to find suitable auxiliary equations as in [Guo et al. 2023] and [Chen and Cheng 2023]. To motivate what a good auxiliary equation is, we first consider the parabolic Monge–Ampère flow (1-5). To have some monotonicity properties of the auxiliary solutions, we prefer a flow with negative time derivative of $\psi$. Thus the inverse Monge–Ampère flow, see (1-4), is a good candidate.

Let us consider the inverse Monge–Ampère flow

$$\begin{cases}(-\dot\psi_s){\omega_{\psi_s}}^n = f_s e^{nF}{\omega_0}^n,\\ \psi_s(\,\cdot\,,0)=0,\end{cases} \tag{2-1}$$

where

$$f_s=\frac{(-\varphi-s)_+}{A_s},\qquad A_s=\int_{\Omega_s}(-\varphi-s)e^{nF}{\omega_0}^n\,\mathrm{d}t,\qquad \Omega_s=\{(z,t)\mid -\varphi(z,t)-s>0\}.$$

But such a flow has singularities, since the factor $(-\varphi - s)_+/A_s$ is not smooth in a neighborhood of $\partial\Omega_s$. Thus we need to consider a sequence of smooth functions $\tau_k(x)$ which converges uniformly to $x \cdot \chi_{\mathbb{R}^+}(x)$ and replace $f_s$ on the right-hand side by

$$\frac{\tau_k(-\varphi - s)}{\int_\Omega \tau_k(-\varphi - s)e^{nF}}.$$

By the dominated convergence theorem, $\psi_{s,k}$ converges to $\psi_s$ uniformly, which means we can always take a limit in the inequalities to get the desired estimates as in [Guo et al. 2023] and [Chen and Cheng 2023]. To simplify our computations, we will keep using (2-1) as our auxiliary equation.

Another crucial modification of our auxiliary flow is that we must restrict the integration over the time slices. To express our idea more clearly, we need Lemma 2.1 as well as Corollary 2.2 in [Chen and Cheng 2023], which will be stated below. For the reader's convenience, we will also include the proof from [Chen and Cheng 2023] here.

**Lemma 2.1.** *Consider the inverse Monge–Ampère flow*

$$\begin{cases}(-\dot\varphi)\omega_\varphi{}^n = e^{nF}\omega_0{}^n,\\ \varphi|_{t=0} = \varphi_0.\end{cases}$$

*Assuming*

$$\int_{M\times[0,T]} e^{nF}\omega_0{}^n \,\mathrm{d}t = C_1 < \infty,$$

*we have* $|\sup_M \varphi| \le C$, *where* $C$ *depends on* $n$, $\omega_0$, $C_1$, *and* $\|\varphi_0\|_{L^\infty}$.

*Proof.* Since $\dot\varphi < 0$, we can get the upper bound by

$$\sup_M \varphi \le \sup_M \varphi_0 \le \|\varphi_0\|_{L^\infty}.$$

To estimate the lower bound of $\sup_M \varphi$, let us consider the $I$-functional

$$I(\varphi) = \frac{1}{n+1}\int_M \varphi \sum_{j=0}^{n} \omega_0{}^{n-j} \wedge \omega_\varphi{}^j$$

and its derivative

$$\frac{\mathrm{d}}{\mathrm{d}t} I(\varphi) = \int_M \partial_t\varphi\,\omega_\varphi{}^n = -\int_M e^{nF}\omega_0{}^n.$$

Therefore, for any $t' \in [0, T]$, we have

$$\begin{aligned} I(\varphi) - I(\varphi_0) = \int_0^{t'} \frac{\mathrm{d}}{\mathrm{d}t} I(\varphi)\,\mathrm{d}t &= -\int_{M\times[0,t']} e^{nF}\omega_0{}^n\,\mathrm{d}t \\ &\ge -\int_{M\times[0,T]} e^{nF}\omega_0{}^n\,\mathrm{d}t = -C_1, \end{aligned}$$

which implies $I(\varphi)$ is bounded from below on $[0, T]$.

The lower bound estimate of $\int_M \varphi\omega_0{}^n$ comes from integration by parts:

$$\begin{aligned}\int_M \varphi\omega_0{}^n - I(\varphi) &= \int_M \varphi\frac{1}{n+1}\sum_{j=0}^{n}\omega_0{}^{n-j}\wedge(\omega_0{}^j-\omega_\varphi{}^j)\\ &= \frac{1}{n+1}\int_M \varphi\sum_{j=0}^{n}\omega_0{}^{n-j}\wedge\sqrt{-1}\partial\bar\partial(-\varphi)\sum_{l=0}^{j-1}\omega_0{}^{j-1-l}\wedge\omega_\varphi{}^l\\ &= \frac{1}{n+1}\int_M \sqrt{-1}\partial\varphi\wedge\bar\partial\varphi\wedge\sum_{j=0}^{n}\sum_{l=0}^{j-1}\omega_0{}^{n-1-l}\wedge\omega_\varphi{}^l \geq 0.\end{aligned}$$

Therefore, we have $\int_M \varphi\omega_0{}^n \geq -C_1$ and

$$\sup_M \varphi \geq \frac{1}{\mathrm{Vol}(M,\omega_0)}\int_M \varphi\omega_0{}^n \geq -\frac{C_1}{\mathrm{Vol}(M,\omega_0)}. \qquad \square$$

**Corollary 2.2.** *There exists a constant* $\alpha > 0$ *depending only on* $\omega_0$ *such that*

$$\sup_{t\in[0,T]}\int_M e^{-\alpha\varphi}\omega_0{}^n \leq C_2,$$

*where* $C_2$ *depends on* $M,\ \omega_0,\ C_1$, *and* $\|\varphi_0\|_{L^\infty}$.

This is a flow version of Hörmander's result; see [Hörmander 1973, Lemma 4.4] and [Tian 1987] for local and global version of such integral estimate, respectively.

*Proof.* Since $\varphi$ is a $\omega_0$-psh function for every $t \in [0, T]$, we have

$$\sup_{t\in[0,T]}\int_M e^{-\alpha(\varphi-\sup_M\varphi)}\omega_0{}^n \leq C.$$

From Lemma 2.1, the uniform bound of $\sup_M \varphi$ gives us the desired inequality. $\square$

The above Corollary 2.2 gives us a uniform bound on each time slice. If we apply this corollary on the space time $M \times [0, T)$, then a factor $T$ seems unavoidable on the right-hand side. Therefore it is better to divide the whole space-time into several pieces $M \times [t_0, t_0+1]$ and try to seek an estimate independent of $t_0$. This idea inspires us to consider such auxiliary equations involving only local information.

To get the $L^\infty$ estimate, we need also to consider the normalization in Theorem 1.1, which follows the same normalization in [Picard and Zhang 2020]. In conclusion, we need to use the domain

$$\widetilde\Omega_s = \{(z,t) \mid -\tilde\varphi(z,t) - s > 0\}$$

as a substitute for $\Omega_s$.

For any $t_0 \in [0, T-1)$, let us consider a family of regions $\Omega_{s,t_0} = \widetilde\Omega_s \cap (M \times [t_0, t_0+1])$ and define a family of auxiliary equations

$$\begin{cases}(-\dot\psi_{s,t_0})\omega_{\psi_{s,t_0}}{}^n = f_{s,t_0}e^{nF}\omega_0{}^n,\\ \psi_{s,t_0}(\,\cdot\,,0) = 0,\end{cases} \tag{2-2}$$

where $A_{s,t_0} = \int_{\Omega_{s,t_0}}(-\tilde\varphi - s)e^{nF}\omega_0{}^n\,\mathrm{d}t$ and $f_{s,t_0} = (-\tilde\varphi - s)\cdot\chi_{\Omega_{s,t_0}}/A_{s,t_0}$.

The benefits appear when we apply Corollary 2.2 to such auxiliary equations. As a result of the choice of $f_{s,t_0}$, the integral on the right-hand side of the equation is 1. This implies

$$\int_{\Omega_{s,t_0}} e^{-\alpha\psi_{s,t_0}} {\omega_0}^n \,\mathrm{d}t \le C_2, \tag{2-3}$$

where the $C_2$ is universal now. The above inequality is crucial, because it integrates against $t$ while it remains independent of $T$ and $t_0$. We will use inequality (2-3) frequently in the following sections.

The family of auxiliary equations (2-2) meets the same problem as (2-1). To be precise, we also need to apply $\tau_k$ to remove the singularities. For the same reason, we will keep using (2-2) in the following sections.

The extra integral condition was chosen to make the normalization $\tilde{\varphi}$ satisfy three properties, as follows.

**Lemma 2.3.** *Let $\tilde{\varphi}$ be given in Theorem 1.1. Then we have*

(1) $\sup_{t\in[0,T)} \int_M \dot{\varphi}{\omega_0}^n \le C_3$,

(2) $\tilde{\varphi} \le C_3$,

(3) $\int_M |\tilde{\varphi}|{\omega_0}^n \le C_3$,

*where $C_3$ is universal.*

*Proof.* For (1), it comes directly from the estimates

$$\begin{aligned} \fint_M \dot{\varphi}{\omega_0}^n &= \fint_M \log\left(\frac{{\omega_\varphi}^n}{e^{nF}{\omega_0}^n}\right){\omega_0}^n = \fint \log\left(\frac{{\omega_\varphi}^n}{{\omega_0}^n}\right){\omega_0}^n - \fint \log(e^{nF}){\omega_0}^n \\ &\le \log\left(\fint {\omega_\varphi}^n\right) - \fint \log(e^{nF}){\omega_0}^n = -\fint \log(e^{nF}){\omega_0}^n \le K. \end{aligned}$$

The first estimate comes from Jensen's inequality while the second one comes from the assumption on $F$. The average integral is chosen with respect to $V = \int_M {\omega_0}^n$.

To prove (2) and (3), let's consider Green's formula

$$\tilde{\varphi} = \fint_M \tilde{\varphi}{\omega_0}^n - \int_M G\Delta\tilde{\varphi}{\omega_0}^n = -\int_M G\Delta\tilde{\varphi}{\omega_0}^n = -\int_M G\Delta\varphi{\omega_0}^n,$$

where $G$ is the Green's function with respect to $\omega_0$.

It is well known that the Green's function $G$ could be shifted to be nonnegative and with $L^1$ norm bound $C'$. Combining with $\operatorname{tr}_{\omega_0}\omega_\varphi = n + \Delta\varphi > 0$ and Green's formula, we have the universal estimate $\tilde{\varphi} \le nC'$.

Let $I_+$ and $I_-$ be the integrals of the positive and negative parts of $\tilde{\varphi}$, respectively. Then we have

$$0 = \int_M \tilde{\varphi} = I_+ - I_- \quad \text{and} \quad I_+ \le nC'V.$$

Thus

$$\int_M |\tilde{\varphi}| = I_+ + I_- = 2I_+ \le 2nC'V.$$

The lemma follows from choosing $C_3 = \max(KV, nC', 2nC'V)$. ☐

## 3. Entropy bounded by energy

From Section 2, we get a good choice of a family of auxiliary equations (2-2). The following lemma is a key to the proof of Theorem 1.1,

**Lemma 3.1.** *Let $\varphi$ be as in Theorem 1.1 and $\psi_{s,t_0}$ be a solution of the auxiliary flow* (2-2). *Then there are constants $\beta$, $\epsilon$, and $\Lambda$, with*

$$\beta=\frac{n+1}{n+2},\quad \epsilon^{n+2}=\Big(\frac{n+2}{n+1}\Big)^{n+2}\Lambda,\quad \epsilon^{n+2}=\Big(\frac{C_4}{(n+1)\beta}\Big)^{n+1}A_{s,t_0},$$

*where $C_4$ is a universal constant defined below, such that*

$$-\epsilon(-\psi_{s,t_0}+\Lambda)^\beta-\tilde{\varphi}-s\le 0 \tag{3-1}$$

*holds on $M\times[t_0,t_0+1]$.*

In the following estimates we will use $\psi$ and $f$ to denote $\psi_{s,t_0}$ and $f_{s,t_0}$, respectively. Let us consider the test function $H=-\epsilon(-\psi+\Lambda)^\beta-\tilde{\varphi}-s$ and the linearization operator $L=-\partial/\partial t+\Delta_{\omega_{\varphi_t}}$ of the Kähler–Ricci flow (1-1). The idea of the following argument comes from [Guo et al. 2023].

Let $\Omega^\circ_{s,t_0}$ denote the interior of $\Omega_{s,t_0}$. Suppose the maximum of $H$ is attained at some point $x_0=(z_0,t_0)$ outside $\Omega^\circ_{s,t_0}$; then we have $H\le H(x_0)\le-\tilde{\varphi}(x_0)-s\le 0$. To complete the proof, we only need to assume the maximal point $x_0$ of $H$ is in $\Omega^\circ_{s,t_0}$ and then apply the maximum principle.

The scheme of the proof is to estimate $LH$ at $x_0$. The constants are chosen to make $0<\beta<1$ and $1-\beta\epsilon\Lambda^{\beta-1}=0$, which imply some cancellations. Moreover, such relations among the constants imply that $LH$ is different from $H$ by a positive coefficient at $x_0$. Roughly speaking, the lemma holds because of the facts that H is proportional to $LH$ and $LH\le 0$. Therefore let us firstly apply the operator $L$ to $H$ at $x_0$:

$$\begin{aligned}
0\ge LH&=-\beta\epsilon(-\psi+\Lambda)^{\beta-1}\dot{\psi}+\dot{\varphi}-⨍\dot{\varphi}\omega_0{}^n\\
&\qquad\qquad+\beta\epsilon(-\psi+\Lambda)^{\beta-1}\Delta_{\omega_\varphi}\psi+\beta(1-\beta)\epsilon(-\psi+\Lambda)^{\beta-2}|\partial\varphi|^2_{\omega_\varphi}-\Delta_{\omega_\varphi}\varphi\\
&\ge\beta\epsilon(-\psi+\Lambda)^{\beta-1}(-\dot{\psi})-(-\dot{\varphi})+\beta\epsilon(-\psi+\Lambda)^{\beta-1}\Delta_{\omega_\varphi}\psi-\Delta_{\omega_\varphi}\varphi-C_3\\
&=\beta\epsilon(-\psi+\Lambda)^{\beta-1}(-\dot{\psi})-(-\dot{\varphi})+\beta\epsilon(-\psi+\Lambda)^{\beta-1}\operatorname{tr}_{\omega_\varphi}\omega_\psi\\
&\qquad\qquad-\operatorname{tr}_{\omega_\varphi}\omega_\varphi+(1-\beta\epsilon(-\psi+\Lambda)^{\beta-1})\operatorname{tr}_{\omega_\varphi}\omega_0-C_3\\
&\ge\beta\epsilon(-\psi+\Lambda)^{\beta-1}(-\dot{\psi}+\operatorname{tr}_{\omega_\varphi}\omega_\psi)-(-\dot{\varphi}+C_3+n).
\end{aligned}$$

The last estimate comes from the choice of auxiliary equations. Since $\psi$ solves some inverse Monge–Ampère flows with initial data being identically 0, we have $\psi\le 0$ on $M\times[0,T)$, and moreover

$$1-\beta\epsilon(-\psi+\Lambda)^{\beta-1}\ge 1-\beta\epsilon\Lambda^{\beta-1}=0.$$

Then we need to deal with the factor $-\dot{\psi}+\operatorname{tr}_{\omega_\varphi}\omega_\psi$, which is the main term of the estimate. By the geometric-arithmetic inequality, we have

$$-\dot{\psi}+\operatorname{tr}_{\omega_\varphi}\omega_\psi\ge-\dot{\psi}+n\Big(\frac{\omega_\psi^n}{\omega_\varphi^n}\Big)^{1/n}. \tag{3-2}$$

Combining (3-2) with the two flow equations (1-1) and (2-2) and using the geometric-arithmetic inequality again, we have

$$\begin{aligned} -\dot{\psi} + \operatorname{tr}_{\omega_\varphi} \omega_\psi &\ge -\dot{\psi} + n\Big(\frac{\omega_\psi^n}{\omega_\varphi^n}\Big)^{1/n} \\ &\ge -\dot{\psi} + n f^{1/n} \exp\Big(-\frac{1}{n}\dot{\varphi}\Big)(-\dot{\psi})^{-1/n} \\ &\ge (n+1) f^{1/(n+1)} \exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big). \end{aligned} \tag{3-3}$$

Thus replacing $-\dot{\psi} + \operatorname{tr}_{\omega_\varphi} \omega_\psi$ by (3-3), we have the following estimate at $x_0$:

$$\begin{aligned} 0 \ge LH &\ge (n+1)\beta\epsilon(-\psi+\Lambda)^{\beta-1} f^{1/(n+1)} \exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big) - (-\dot{\varphi} + C_3 + n) \\ &\ge \Big[(n+1)\beta\epsilon(-\psi+\Lambda)^{\beta-1} f^{1/(n+1)} - (-\dot{\varphi} + C_3 + n)\exp\Big(\frac{1}{n+1}\dot{\varphi}\Big)\Big] \exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big). \end{aligned}$$

Since the exponential function is positive, we can simplify it by

$$0 \ge (n+1)\beta\epsilon(-\psi+\Lambda)^{\beta-1} f^{1/(n+1)} + (\dot{\varphi} - C_3 - n)\exp\Big(\frac{1}{n+1}\dot{\varphi}\Big), \tag{3-4}$$

which looks similar to the desired inequality (3-1) in Lemma 3.1. Let us consider a function

$$h(x) = (x - n - C_3)\exp\Big(\frac{1}{n+1}x\Big).$$

The function $h : \mathbb{R} \to \mathbb{R}$ has a universal lower bound $-C_4$, where

$$C_4 = (n+1)\exp\Big(\frac{C_3 - 1}{n+1}\Big).$$

Thus we have $h(\dot{\varphi}) \ge -C_4$, and moreover

$$(n+1)\beta\epsilon(-\psi+\Lambda)^{\beta-1} f^{1/(n+1)} - C_4 \le 0. \tag{3-5}$$

Since $f = (-\tilde{\varphi} - s)/A_{s,t_0}$ at the maximal point $x_0$, inequality (3-5) is equivalent to

$$\Big(\frac{(n+1)\beta}{C_4}\Big)^{n+1} \epsilon^{n+1} \frac{-\tilde{\varphi} - s}{A_{s,t_0}} \le (-\psi+\Lambda)^{(n+1)(1-\beta)}, \tag{3-6}$$

which is the test function when we chose the constants $\beta$, $\epsilon$, and $\Lambda$ as stated in Lemma 3.1. Thus we have

$$H \le H(x_0) \le 0,$$

which completes the proof.

From the above Lemma 3.1, we have

$$\frac{-\tilde{\varphi} - s}{A_{s,t_0}^{1/(n+2)}} \le (c(-\psi+\Lambda))^{(n+1)/(n+2)}, \tag{3-7}$$

where

$$c = \Big(\frac{(n+1)^2\epsilon}{(n+2)C_4}\Big)^{n+2}.$$

The following estimate comes from (3-7):

$$\int_{\Omega_{s,t_0}} \exp\biggl[\lambda\biggl(\frac{-\tilde{\varphi}-s}{A_{s,t_0}^{1/(n+2)}}\biggr)^{(n+2)/(n+1)}\biggr]\omega_0{}^n \,\mathrm{d}t \le \int_{\Omega_{s,t_0}} \exp\{\lambda c(-\psi_{s,t_0}+\Lambda)\}\omega_0{}^n \,\mathrm{d}t.$$

If the universal constant $\lambda$ is chosen to make $\lambda c=\alpha$, then, by Corollary 2.2 and inequality (2-3), we have

$$\int_{\Omega_{s,t_0}} e^{-\lambda c\psi_{s,t_0}}\omega_0{}^n \,\mathrm{d}t \le \int_{M\times[t_0,t_0+1]} e^{-\alpha\psi_{s,t_0}}\omega_0{}^n \,\mathrm{d}t \le C_2,$$

where $C_2$ is universal. In Lemma 3.1, The constant $\Lambda$ is chosen to be proportional to $A_{s,t_0}$, i.e., $\Lambda=c'A_{s,t_0}$ for some universal constant $c'$. Then we can bound the right integral by $C\exp(CA_{s,t_0})$ for some universal constant $C$.

Let us define $E=\sup_{t_0\in[0,T-1)}\int_{M\times[t_0,t_0+1]}(-\tilde{\varphi}+C_3)e^{nF}\omega_0{}^n\,\mathrm{d}t$. Then we have

$$\begin{aligned} A_{s,t_0} &= \int_{\Omega_{s,t_0}}(-\tilde{\varphi}-s)e^{nF} \le \int_{\Omega_{s,t_0}}(-\tilde{\varphi}+C_3)e^{nF} + \int_{\Omega_{s,t_0}}(-C_3-s)e^{nF} \\ &\le \int_{\Omega_{s,t_0}}(-\tilde{\varphi}+C_3)e^{nF} = \int_{M\times[t_0,t_0+1]}(-\tilde{\varphi}+C_3)e^{nF} + \int_{M\times[t_0,t_0+1]\setminus\Omega_{s,t_0}}(\tilde{\varphi}-C_3)e^{nF} \\ &= E + \int_{M\times[t_0,t_0+1]\setminus\Omega_{s,t_0}}(\tilde{\varphi}-C_3)e^{nF} \le E, \end{aligned}$$

where the last inequality comes from $\tilde{\varphi}\le C_3$ by Lemma 2.3. In summary, we have

$$\int_{\Omega_{s,t_0}} \exp\biggl[\lambda\biggl(\frac{-\tilde{\varphi}-s}{A_{s,t_0}^{1/(n+2)}}\biggr)^{(n+2)/(n+1)}\biggr]\omega_0{}^n \,\mathrm{d}t \le C\exp(CE). \tag{3-8}$$

The $E$ defined above is called the energy. The $C_3$ term inside the integral comes purely from a technical consideration that makes the inside function of the integral positive. There is no significant difference from the elliptic case where $E=\int(-\tilde{\varphi})$ since the extra integral $0<\int e^{nF}\le \mathrm{Vol}(M,\omega_0)+\mathrm{Ent}_p(F)$ is universally bounded from both sides.

To end this section, we will use the De Giorgi iteration method to derive the $C^0$ estimate by assuming $E$ is universally bounded. In the next section, we will apply the ABP estimate to get the universal bound on $E$ which will complete the proof of Theorem 1.1. To prepare for the iteration procedure, we need such an inequality to run the iteration:

$$r\phi_{t_0}(s+r)\le A_{s,t_0}\le B_0\phi_{t_0}(s)^{1+\delta_0}, \tag{3-9}$$

where $\phi_{t_0}(s)=\int_{\Omega_{s,t_0}}e^{nF}\omega_0{}^n\,\mathrm{d}t$.

The following lemma is the De Giorgi iteration mentioned above.

**Lemma 3.2.** *Let $\Phi:\mathbb{R}_+\to\mathbb{R}_+$ be a decreasing right continuous function with $\lim_{s\to\infty}\Phi(s)=0$. Moreover, assume $r\Phi(s+r)\le B_0\Phi(s)^{1+\delta_0}$ for some constant $B_0>0$ and all $s>0$ and $r\in[0,r]$. Then there exists a constant $S_\infty=S_\infty(\delta_0,B_0,s_0)>0$ such that $\Phi(s)=0$ for all $s\ge S_\infty$, where $s_0$ is defined during the proof.*

*Proof.* Fix an $s_0 > 0$ such that $\Phi(s_0)^{\delta_0} < 1/(2B_0)$. Such an $s_0$ exists since $\Phi(s) \to 0$ as $s \to \infty$. Define $\{s_j\}$ by

$$s_{j+1} = \sup\bigl\{s > s_j \mid \phi(s) > \tfrac{1}{2}\phi(s_j)\bigr\}.$$

Thus

$$s_{j+1} - s_j \le B_0 \Phi(s_j)^{1+\delta}/\Phi(s_{j+1}) \le 2B_0 \Phi(s_j)^{\delta} \le 2B_0 2^{-j\delta_0}\Phi^{\delta_0} \le 2^{-j\delta_0}.$$

Letting

$$S_\infty = s_0 + \sum_{j\ge 0}(s_{j+1}-s_j) \le s_0 + \frac{1}{1-2^{-\delta_0}},$$

we complete the proof. □

In our application, $\Phi$ is chosen to be $\phi_{t_0}$. Let us derive the two sides of (3-9) separately. The left-hand side of (3-9) can be derived by definition which is similar to the proof in [Guo et al. 2023]. The following calculations are direct:

$$\begin{aligned} A_{s,t_0} &= \int_{\Omega_{s,t_0}} (-\tilde{\varphi} - s)e^{nF}\omega_0{}^n \,\mathrm{d}t \ge \int_{\Omega_{s+r,t_0}} (-\tilde{\varphi} - s)e^{nF}\omega_0{}^n \,\mathrm{d}t \\ &\ge \int_{\Omega_{s+r,t_0}} (s+r-s)e^{nF}\omega_0{}^n \,\mathrm{d}t = r\phi(s+r). \end{aligned}$$

To get the right-hand side of (3-9), we need to apply the following inequality coming from Young's inequality:

$$\int_{\Omega_{s,t_0}} v^p e^{nF}\omega_0{}^n \,\mathrm{d}t \le \|e^{nF}\|_{L^1(\log L)^p(M\times[t_0,t_0+1])} + C_p \int_{\Omega_{s,t_0}} e^{2v}\omega_0{}^n \,\mathrm{d}t. \tag{3-10}$$

If we choose

$$v = \frac{\lambda}{2}\biggl(\frac{-\tilde{\varphi} - s}{A_{s,t_0}^{1/(n+2)}}\biggr)^{(n+2)/(n+1)},$$

then by the above inequality (3-10) we have

$$\int_{\Omega_{s,t_0}} (-\tilde{\varphi} - s)^{(n+2)p/(n+1)} e^{nF}\omega_0{}^n \,\mathrm{d}t \le C(E) A_{s,t_0}^{p/(n+1)}, \tag{3-11}$$

where the factor $C(E)$ is a constant dependent on $n,\ \omega_0,\ \varphi_0,\ p,\ \gamma,\ K,\ \mathrm{Ent}_p(F)$ and additionally on $E$. The explicit dependence of the constant $C(E)$ on $E$ can be expressed by combining inequalities (3-10) and (3-8).

Thus the right-hand side can be derived by the estimate

$$\begin{aligned} A_{s,t_0} &= \int_{\Omega_{s,t_0}} (-\tilde{\varphi} - s)e^{nF}\omega_0{}^n \,\mathrm{d}t \\ &\le \biggl(\int_{\Omega_{s,t_0}} (-\tilde{\varphi} - s)^{(n+2)p/(n+1)} e^{nF}\biggr)^{(n+1)/((n+2)p)} \cdot \biggl(\int_{\Omega_{s,t_0}} e^{nF}\biggr)^{1/q} \\ &\le C(E)^{(n+1)/((n+2)p)} A_{s,t_0}^{1/(n+2)} \phi_{t_0}^{1-(n+1)/(p(n+2))}, \end{aligned}$$

where the first line is by Hölder's inequality and $p$ is the Hölder coefficient

$$\frac{n+1}{p(n+2)}+\frac{1}{q}=1.$$

In (3-9), we can choose

$$\delta_0=1+\frac{p-n-1}{p(n+1)} \quad \text{and} \quad B_0=C(E)^{1/p}.$$

To complete this section, we need to get an explicit expression on $s_0$. By Chebyshev's inequality

$$\phi_{t_0}(s)\le\frac{1}{s}\int_{\Omega_{s,t_0}}(-\tilde{\varphi})e^{nF}\omega_0^n\le\frac{E}{s},$$

we can choose $s_0=(2B_0)^{1/\delta_0}E$.

In conclusion, we get the following theorem from the above arguments.

**Theorem 3.3.** *Let $\tilde{\varphi}$ be as in Theorem 1.1. Then we have*

$$\sup_{M\times[0,T)}|\tilde{\varphi}|\le C(n,\omega_0,\varphi_0,p,\gamma,K,\mathrm{Ent}_p(F),E).$$

*Moreover, if $E$ can be controlled by a universal constant, then we have Theorem 1.1.*

## 4. Energy bounds by the ABP estimate

In this section we use a parabolic version of the ABP estimate proved by Krylov [1976] and Tso [1985] to give us a uniform energy bound. This approach was introduced in [Chen and Cheng 2023] and is an analogue to the elliptic version in [Guo et al. 2023].

Let $u$ be a function defined on $D=\Omega\times[0,T]$, where $\Omega$ is a bounded domain in $\mathbb{R}^n$. Then the parabolic ABP estimate says that

$$\sup_D u\le\sup_{\partial_P D}u+C_n(\operatorname{diam}\Omega)^{n/(n+1)}\left(\int_\Gamma|\partial_t u\det D_x^2u|\,\mathrm{d}x\,\mathrm{d}t\right)^{1/(n+1)}, \tag{4-1}$$

where $\partial_P D$ is the parabolic boundary of $D$ and $\Gamma=\{(x,t)\mid\partial_t u\ge0,\ D_x^2u\le0\}$.

As mentioned in Section 2, we want to construct a family of local auxiliary equations. The auxiliary equations in this section are chosen to be

$$\begin{cases}(-\partial_t\psi_{t_0}){\omega_{\psi_{t_0}}}^n=\dfrac{(|F|^p+1)\cdot\chi_{M\times[t_0,t_0+1]}}{\int_{M\times[t_0,t_0+1]}(|F|^p+1)e^{nF}{\omega_0}^n\,\mathrm{d}t}e^{nF}{\omega_0}^n,\\ \psi_{t_0}(\,\cdot\,,0)=0.\end{cases} \tag{4-2}$$

We will use $\psi$ to denote $\psi_{t_0}$ for convenience and will skip the computation involved with $\tau_k$ for the same reason we did in Section 2. Moreover, define a universal constant

$$\Psi=\int_{M\times[t_0,t_0+1]}(|F|^p+1)e^{nF}{\omega_0}^n\,\mathrm{d}t.$$

Parallel with Lemma 3.1, the following lemma plays a key role in this section.

**Lemma 4.1.** *Let $\varphi$ be as in Theorem 1.1 and $\psi$ be a solution of* (4-2). *For any $0 < \beta < 1$, there exists a constant $C$ which depends on $n$, $\omega_0$, $\varphi_0$, $p$, $\gamma$, $K$, $\mathrm{Ent}_p(F)$, and additionally on $\beta$, such that the following holds on $M \times [t_0, t_0 + 1]$:*

$$-\epsilon(-\psi + \Lambda)^{\beta} - \tilde{\varphi} \leq C, \tag{4-3}$$

*where the constants $\epsilon > 0$ and $\Lambda > 0$ are defined as*

$$\beta\epsilon\Lambda^{\beta-1} = \tfrac{1}{4}, \quad \Lambda = \left(\frac{n^p(2n+1)^p 2^p 4^{n+1} 9^{n+1}(n+1)^{n+1} C_4^{n+1}\Psi}{10^{n+1}\alpha^p}\right)^{1/((n+1)(1-\beta))}. \tag{4-4}$$

*The constants $\epsilon$ and $\Lambda$ depend additionally on $\beta$.*

Let $\rho$ be the test function defined by

$$\rho = -\epsilon(-\psi + \Lambda)^{\beta} - \tilde{\varphi}$$

and $L$ be the linearization as above. To prove Lemma 4.1, we only need to restrict $\rho$ to its positive part. More precisely, consider $h_s(x) = x + \sqrt{x^2 + s}$ and use $h_s(\rho)$ to approximate $2\rho_+$. Therefore we have

$$2 \sup \rho \leq 2 \sup \rho_+ \leq \sup h_s(\rho),$$

and in addition the upper bounds of $h_s$ imply an upper bound of $\rho$.

Let us consider $h_s(\rho)^b$, where

$$b = 1 + \frac{1}{(2n+2)(2n+1)},$$

and assume $h_s(\rho)^b$ attains its maximal value $Q$ at some point $x_0 \in M \times [0, T)$. Moreover, we can assume $Q > 1$, otherwise there is nothing to prove. Let us apply the parabolic ABP estimate for $H = h_s(\rho)^b \cdot \eta$, where $\eta$ is a cut-off function defined below.

Assume $r_0 = \min\{1, \mathrm{inj}(M, \omega_0)\}$, where $\mathrm{inj}(M, \omega_0)$ is the injectivity radius of $(M, \omega_0)$. The cut-off function $\eta : M \to \mathbb{R}$ is defined in the following way:

$$\eta \equiv 1 \qquad \text{on } B_{\omega_0}\big(x_0, \tfrac{1}{2}r_0\big), \tag{4-5}$$

$$\eta \equiv 1 - \theta \qquad \text{on } \big\{M \backslash B_{\omega_0}\big(x_0, \tfrac{3}{4}r_0\big)\big\}, \tag{4-6}$$

$$1 - \theta \leq \eta \leq 1 \qquad \text{on } \big\{B_{\omega_0}\big(x_0, \tfrac{3}{4}r_0\big) \backslash B_{\omega_0}\big(x_0, \tfrac{1}{2}r_0\big)\big\}, \tag{4-7}$$

$$|\nabla\eta|^2_{\omega_0} \leq 10\theta^2/r_0^2, \tag{4-8}$$

$$|\nabla^2\eta|_{\omega_0} \leq 10\theta/r_0^2, \tag{4-9}$$

where $0 < \theta < 1$ is a small constant defined by

$$\theta = \min\left\{\frac{r_0^2}{100Q^{1/b}}, \frac{1}{2(2n+1)(2n+2)}\right\} \leq \frac{1}{10}.$$

Since $\eta$ is a space function, it has vanishing time derivative, which reduce our later computations.

*Proof of Lemma 4.1.* The following inequality can be derived directly by applying the operator $L$ on $H$:

$$LH \geq bh'h^{b-1}(-\partial_t \rho)\eta + (\Delta_{\omega_\varphi} h^b)\eta + 2\,\mathrm{Re}\langle \nabla h^b, \overline{\nabla}\eta\rangle_{\omega_\varphi} + h^b \Delta_{\omega_\varphi}\eta. \tag{4-10}$$

We will consider each of the terms separately to get good controls. Let us bound the last two terms:

$$h^b \Delta_{\omega_\varphi}\eta \geq -\frac{10\theta}{r_0^2} h^b \,\mathrm{tr}_{\omega_\varphi} \omega_0, \tag{4-11}$$

$$2\,\mathrm{Re}\langle \nabla h^b, \overline{\nabla}\eta\rangle \geq -\frac{b(b-1)}{2}|\nabla h|^2_{\omega_\varphi} h^{b-2} - \frac{2b}{b-1} h^b |\nabla \eta|^2_{\omega_\varphi}. \tag{4-12}$$

Then, we expand the second term and get

$$(\Delta_{\omega_\varphi} h^b)\eta = b(b-1)|\nabla h|^2_{\omega_\varphi} h^{b-2}\eta + bh'h^{b-1}(\Delta_{\omega_\varphi}\rho)\eta + b|\nabla\rho|^2_{\omega_\varphi} h'' h^{b-1}\eta. \tag{4-13}$$

Combining (4-10)–(4-13), and noticing that the first term in (4-13) can be absorbed into the first term in (4-12) and the third term of (4-13) is positive, we have

$$\begin{aligned} LH &\geq bh'h^{b-1}(-\partial_t\rho)\eta + bh'h^{b-1}(\Delta_{\omega_\varphi}\rho)\eta - \frac{2b}{b-1}h^b|\nabla\eta|^2_{\omega_\varphi} - \frac{10\theta}{r_0^2}h^b\,\mathrm{tr}_{\omega_\varphi}\omega_0 \\ &\geq bh'h^{b-1}(L\rho)\eta - \frac{2b}{b-1}\frac{10\theta^2}{r_0^2}h^b\,\mathrm{tr}_{\omega_\varphi}\omega_0 - \frac{10\theta}{r_0^2}h^b\,\mathrm{tr}_{\omega_\varphi}\omega_0. \end{aligned} \tag{4-14}$$

As we mentioned, the derivatives of the cut-off function $\eta$ will produce $\mathrm{tr}_{\omega_\varphi}\omega_0$ terms in (4-14) which will be absorbed in the later estimates. The $L\rho$ term is the main term of (4-14), and it has the same structure as the main term of the test function appearing in Lemma 3.1. This fact motivates the following argument.

Let us compute $L\rho$ and drop the positive term $\beta(1-\beta)\epsilon(-\psi+\Lambda)^{\beta-2}|\nabla\psi|^2$. Then we have

$$\begin{aligned} L\rho &\geq -\beta\epsilon\dot{\psi}(-\psi+\Lambda)^{\beta-1} + \dot{\tilde{\varphi}} + \beta\epsilon(-\psi+\Lambda)^{\beta-1}\Delta_{\omega_\varphi}\psi + \beta(1-\beta)\epsilon(-\psi+\Lambda)^{\beta-2}|\nabla\psi|^2 - \Delta_{\omega_\varphi}\tilde{\varphi} \\ &\geq -\beta\epsilon\dot{\psi}(-\psi+\Lambda)^{\beta-1} + \dot{\varphi} + \beta\epsilon(-\psi+\Lambda)^{\beta-1}\Delta_{\omega_\varphi}\psi - \Delta_{\omega_\varphi}\varphi - C_3. \end{aligned}$$

Since $\Delta_{\omega_\varphi}\varphi + \mathrm{tr}_{\omega_\varphi}\omega_0 = n$ and $\Delta_{\omega_\varphi}\psi + \mathrm{tr}_{\omega_\varphi}\omega_0 = \mathrm{tr}_{\omega_\varphi}\omega_\psi$, we have

$$L\rho \geq \beta\epsilon(-\psi+\Lambda)^{\beta-1}(-\dot{\psi} + \mathrm{tr}_{\omega_\varphi}\omega_\psi) + \dot{\varphi} + (1-\beta\epsilon\Lambda^{\beta-1})\,\mathrm{tr}_{\omega_\varphi}\omega_0 - C_3 - n. \tag{4-15}$$

The $\mathrm{tr}_{\omega_\varphi}\omega_0$ term in (4-15) will serve as a good term to absorb the last two terms in (4-14). The estimate for the rest of the terms in (4-15) follows the same idea in (3-2)–(3-4).

$$\begin{aligned} &\beta\epsilon(-\psi+\Lambda)^{\beta-1}(-\dot{\psi} + \mathrm{tr}_{\omega_\varphi}\omega_\psi) + \dot{\varphi} - C_3 - n \\ &\qquad\geq (n+1)\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)}\exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big) + \dot{\varphi} - C_3 - n \\ &\qquad= \Big[(n+1)\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)} + (\dot{\varphi} - C_3 - n)\exp\Big(\frac{1}{n+1}\dot{\varphi}\Big)\Big]\exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big) \\ &\qquad\geq [(n+1)\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)} - C_4]\exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big), \end{aligned} \tag{4-16}$$

where

$$\tilde{f} = \frac{(|F|^p + 1)\chi_{M\times[t_0,t_0+1]}}{\int_{M\times[t_0,t_0+1]}(|F|^p + 1)e^{nF}\omega_0{}^n\,\mathrm{d}t}.$$

What we have now is the following inequality, noting that the cut-off function satisfies $\frac{9}{10} \le \eta \le 1$ for any points in the space-time $M \times [0, T)$:

$$LH \ge (n+1)bh'h^{b-1}\left[\frac{9}{10}\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)} - \frac{C_4}{n+1}\right]\exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big) + bh^{b-1}\left[\frac{9}{10}h'(1-\beta\epsilon\Lambda^{\beta-1}) - \frac{20\theta^2}{(b-1)r_0^2}h - \frac{10\theta}{br_0^2}h\right]\mathrm{tr}_{\omega_\varphi}\,\omega_0. \quad (4\text{-}17)$$

Although the core of the lemma is to control $\rho_+$, the choice of $h_s$ makes the negative part of $\rho$ involved in the above inequality. So we will estimate on sets $\Omega_+ = \{\rho > 0\}$ and $\Omega_- = \{\rho \le 0\}$ separately.

On $\Omega_-$, we have

$$0 \le h_s(\rho) = \rho + \sqrt{\rho^2+s} = \frac{s}{\sqrt{\rho^2+s}-\rho} \le \sqrt{s}$$

and

$$0 \le h_s'(\rho) = 1 + \frac{\rho}{\sqrt{\rho^2+s}} \le 1.$$

Combining the two bounds on $h$ and $h'$ and inequality (4-17), we have

$$\begin{aligned} LH &\ge bs^{(b-1)/2}\left[-C\exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big) - \frac{20\theta^2}{(b-1)r_0^2}h\,\mathrm{tr}_{\omega_\varphi}\,\omega_0 - \frac{10\theta}{br_0^2}h\,\mathrm{tr}_{\omega_\varphi}\,\omega_0\right] \\ &= bs^{(b-1)/2}\left[-C\exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big) - c(b,\theta,r_0)h\,\mathrm{tr}_{\omega_\varphi}\,\omega_0\right] \end{aligned}$$

on $\Omega_-$, where $C$ is universal and

$$c(b,\theta,r_0) = \frac{20\theta^2}{(b-1)r_0^2} + \frac{10\theta}{br_0^2}.$$

On the other hand, $1 \le h' \le 2$ on $\Omega_+$. By the choice of the constants $\Lambda$ and $\epsilon$ in (4-4), the coefficient of the $\mathrm{tr}_{\omega_\varphi}\,\omega_0$ term in (4-17) is positive.

Therefore, on the set $\Omega_+$, we have

$$LH \ge Cbh^{b-1}\left[\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)} - \frac{10C_4}{9(n+1)}\right]\exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big).$$

Combining the above two estimates, we obtain

$$LH \ge bs^{(b-1)/2}\Big[-C\exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big) - c(b,\theta,r_0)h\,\mathrm{tr}_{\omega_\varphi}\,\omega_0\Big]\chi_{\Omega_-} + Cbh^{b-1}\Big(\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)} - \frac{10C_4}{9(n+1)}\Big)\exp\Big(-\frac{1}{n+1}\dot{\varphi}\Big)\chi_{\Omega_+} =: R. \quad (4\text{-}18)$$

To apply the parabolic ABP estimate for the test function $H$, we define the domain

$$\widetilde{\Gamma} = \{(z,t) \mid \partial_t H \ge 0,\ D_z^2 H \le 0\}.$$

We need also discuss how to control the operator $L$ on $\widetilde{\Gamma}$. The estimate can be derived by

$$LH = -\frac{\partial}{\partial t}H + \Delta_{\omega_\varphi} H \le -(2n+1)\left(\left|\frac{\partial}{\partial t}H \cdot \det D^2 H\right| \cdot \left(\frac{{\omega_0}^n}{{\omega_\varphi}^n}\right)^2\right)^{1/(2n+1)},$$

which connects our operator with the parabolic ABP estimate. In conclusion,

$$\left(|\partial_t H \cdot \det D^2 H| \cdot \left(\frac{{\omega_0}^n}{{\omega_\varphi}^n}\right)^2\right)^{1/(2n+1)} \le -\frac{R}{2n+1} \le \frac{R_-}{2n+1}.$$

Note that the Hessian matrix $D^2 H$ is the real Hessian of $H$ instead of the complex Hessian of $H$. Therefore we have

$$|\partial_t H \cdot \det D^2 H| \le c_n {R_-}^{2n+1}\left(\frac{{\omega_\varphi}^n}{{\omega_0}^n}\right)^2.$$

The factor ${\omega_\varphi}^n/{\omega_0}^n$ is hard to control since it appears as a quadratic term. We can use different strategies to bound such term on $\Omega_+$ and $\Omega_-$. On the domain $\Omega_+$, we have

$$\begin{aligned} h^{(2n+1)(b-1)}&\left(\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)} - \frac{10C_4}{9(n+1)}\right)_-^{2n+1} \exp\left(-\frac{2n+1}{n+1}\dot{\varphi}\right)\left(\frac{{\omega_\varphi}^n}{{\omega_0}^n}\right)^2 \\ &\le h^{(2n+1)(b-1)}\left(\frac{10C_4}{9(n+1)}\right)^{2n+1} e^{(n(2n+1)/(n+1))F}\left(\frac{{\omega_\varphi}^n}{{\omega_0}^n}\right)^{2-(2n+1)/(n+1)} \\ &= \left(\frac{10C_4}{9(n+1)}\right)^{2n+1} h^{(2n+1)(b-1)} e^{(n(2n+1)/(n+1))F}\left(\frac{{\omega_\varphi}^n}{{\omega_0}^n}\right)^{1/(n+1)}, \end{aligned} \tag{4-19}$$

while on $\Omega_-$, we have

$$\begin{aligned} s^{(2n+1)(b-1)/2}&\left(\exp\left(-\frac{1}{n+1}\dot{\varphi}\right) + c(b,\theta,r_0)h\,\mathrm{tr}_{\omega_\varphi}\omega_0\right)^{2n+1}\left(\frac{{\omega_\varphi}^n}{{\omega_0}^n}\right)^2 \\ &\le C_5 s^{(2n+1)(b-1)/2} = C_5 s^{1/(4n+4)}, \end{aligned} \tag{4-20}$$

where $C_5$ is not universal since it depends additionally on $\varphi$, $t$, and $t_0$.

Defining $D = B_r(x_0)\times[t_0, t_0+1]$ and combining (4-19) and (4-20), the parabolic estimate (4-1) tells us that

$$\begin{aligned} &\sup_D(H) - \sup_{\partial_P D}(H) \\ &\le C_6\left(\int_{D\cap\Omega_+} h^{(2n+1)(b-1)} e^{(n(2n+1)/(n+1))F}\left(\frac{{\omega_\varphi}^n}{{\omega_0}^n}\right)^{1/(n+1)} {\omega_0}^n\,\mathrm{d}t + \int_{D\cap\Omega_-} C_5 s^{1/(4n+4)} {\omega_0}^n\,\mathrm{d}t\right)^{1/(2n+1)} \\ &\le C_6\left(\int_{D\cap\Omega_+}\frac{1}{n+1}\left(nh^{(2n+1)(n+1)(b-1)}e^{n(2n+1)F} + \frac{{\omega_\varphi}^n}{{\omega_0}^n}\right){\omega_0}^n\,\mathrm{d}t + C_5 s^{1/(4n+4)}\right)^{1/(2n+1)} \\ &\le C_6\left(\int_{D\cap\Omega_+} h^{1/2}e^{n(2n+1)F} + \int_{M\times[t_0,t_0+1]} {\omega_\varphi}^n\,\mathrm{d}t + C_5 s^{1/(4n+4)}\right)^{1/(2n+1)} \\ &\le C_6\left(\int_{D\cap\Omega_+} h^{1/2}e^{n(2n+1)F} + 1 + C_5 s^{1/(4n+4)}\right)^{1/(2n+1)}, \end{aligned} \tag{4-21}$$

where $C_6$ is universal. Moreover, $C_6$ changes line by line as it absorbs all universal coefficients derived from the estimates. We use the volume of $M$ to absorb the bad factor on $\Omega_+$ and $C_5$ to absorb the same bad factor on $\Omega_-$

The integral over the set $D\cap\Omega_+$ is in fact integrated over the set

$$\{\rho>0\}\cap\left\{\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)}-\frac{10C_4}{9(n+1)}<0\right\}.$$

Over this set we have, from the choice of constants,

$$\begin{aligned} n(2n+1)|F| &\le \left(\frac{10C_4}{9(n+1)}\right)^{(n+1)/p}\Psi^{1/p}(\beta\epsilon)^{-(n+1)/p}(-\psi+\Lambda)^{(1-\beta)(n+1)/p} \\ &= \frac{\alpha}{2}(-\psi+\Lambda)^{(1-\beta)(n+1)/p}. \end{aligned} \tag{4-22}$$

Moreover, $h(\rho)\le 2\rho+\sqrt{s}$. Then by combining with the inequality (4-21), we have

$$\begin{aligned} &\sup_D(H)-\sup_{\partial_P D}(H) \\ &\le C_6\biggl(\int_{D\cap\Omega_+}(-\tilde{\varphi}+\sqrt{s})^{1/2}\exp\Bigl(\frac{\alpha}{2}(-\psi+\Lambda)^{(1-\beta)(n+1)/p}\Bigr)\omega_0{}^n\,dt+1+C_5s^{1/(4n+4)}\biggr)^{1/(2n+1)} \\ &\le C_6\biggl(\int_{D\cap\Omega_+}(-\tilde{\varphi}+\exp(\alpha(-\psi+\Lambda)^{(1-\beta)(n+1)/p}))\omega_0{}^n\,dt+1+s^{1/2}+C_5s^{1/(4n+4)}\biggr)^{1/(2n+1)} \\ &\le C+C_5s^{1/(4n+2)}, \end{aligned} \tag{4-23}$$

where $C$ has the same dependencies as $\Lambda$. Moreover $C_5$ changes line by line.

The last inequality is derived based on the following two inequalities: the integral $\int_M(-\varphi)\omega_0{}^n\,dt$ is uniformly bounded by Lemma 2.3, and the fact that $0<(1-\beta)(n+1)/p<(n+1)/p<1$ since $0<\beta<1$ and $p>n+1$. Therefore the second integral is bounded by Corollary 2.2 and inequality (2-3).

By (4-23) and the definition of $\theta$, we have

$$cQ^{1-1/b}\le\theta Q\le\sup_D(H)-\sup_{\partial_P D}(H)\le C+C_5s^{1/(4n+2)},$$

where $c$ is universal. In addition, $\sup\rho$ can be controlled by

$$2\sup\rho_+\le\sup h_s(\rho)\le Q^{1/b}.$$

The proof of Lemma 4.1 follows from taking the limit $s\to 0^+$. □

Once we have Lemma 4.1, the following theorem is a direct application of Jensen's inequality.

**Theorem 4.2.** *Let $\varphi$ be the $C^2$ solution defined in Theorem 1.1. For any $\beta\in(0,1)$ and $t_0\in[0,T-1)$, we have the energy estimate*

$$\int_{M\times[t_0,t_0+1]}(-\tilde{\varphi}+C_3)^{1/\beta}e^{nF}\omega_0{}^n\,dt\le C,$$

*where the constant $C$ has the same dependencies as $\Lambda$ defined in Lemma 4.1.*

*Proof.* Similar to the application of Lemma 3.1 and (3-7), we have

$$\int_{M\times[t_0,t_0+1]} \exp(c_\beta(-\tilde{\varphi}+C_3)^{1/\beta})\omega_0{}^n \,\mathrm{d}t \le C_\beta,$$

where $c_\beta$ and $C_\beta$ both have the same dependencies as $\Lambda$.

Let $\widetilde{V}=\int_M e^{nF}\omega_0{}^n\,\mathrm{d}t$ be the volume on the weighted volume form $e^{nF}\omega_0{}^n\,\mathrm{d}t$. Taking logarithms of both sides and applying Jensen's inequality, we have

$$\begin{aligned}\widetilde{V}\log\left(\frac{C_\beta}{\widetilde{V}}\right) &\ge \widetilde{V}\log\left(\frac{1}{\widetilde{V}}\int_{M\times[t_0,t_0+1]}\exp(c_\beta(-\tilde{\varphi}+C_3)^{1/\beta}-nF)e^{nF}\omega_0{}^n\,\mathrm{d}t\right)\\ &\ge \int_{M\times[t_0,t_0+1]}(c_\beta(-\tilde{\varphi}+C_3)^{1/\beta}-nF)e^{nF}\omega_0{}^n\,\mathrm{d}t\\ &\ge c_\beta\int_{M\times[t_0,t_0+1]}(-\tilde{\varphi}+C_3)^{1/\beta}e^{nF}\omega_0{}^n\,\mathrm{d}t-\mathrm{Ent}_p(F).\end{aligned}$$

The theorem follows from $\widetilde{V}\le \mathrm{Ent}_p(F)+V(M,\omega_0)$ and the fact $y\log(y)>-1/e$ for $y>0$. □

*Proof of Theorem 1.1.* Theorem 3.3 tells us the result follows directly from a uniform control on $E$. By Hölder's inequality, we have

$$\int_{M\times[t_0,t_0+1]}(-\tilde{\varphi}+C_3)e^{nF}\omega_0{}^n\,\mathrm{d}t \le \widetilde{V}^{1/n}\left(\int_{M\times[t_0,t_0+1]}(-\tilde{\varphi}+C_3)^{n/(n-1)}e^{nF}\omega_0{}^n\,\mathrm{d}t\right)^{(n-1)/n}.$$

If we fix $\beta=1-1/n$ in Theorem 4.2, then the integral estimate is universal and independent of $t_0$. Thus we complete the proof of Theorem 1.1. □

## 5. Some generalizations

In this section, we derive some generalizations, Theorems 1.2 and 1.3, of Theorem 1.1.

The idea of Theorem 1.2 comes from the result of Chen and Cheng [2023] for general parabolic Hessian equations. Recall that $r$ denotes the homogeneous degree of the operator $\mathcal{F}$ and the linearization of the flow (1-3) is

$$Lu=-\partial_t u+G^{i\bar{j}}u_{i\bar{j}},$$

where

$$G^{i\bar{j}}=\frac{1}{\mathcal{F}}\frac{\partial\mathcal{F}(\lambda[h_\varphi])}{\partial h_{i\bar{j}}}.$$

To prove Theorem 1.2, we will use the family of auxiliary equations (2-2) and follow the same argument as in Sections 3 and 4.

The proof of main estimate (3-1) is tedious, and we will only show the essential differences compared to previous sections. When we apply the operator $L$ to the test function

$$-\epsilon(-\psi+\Lambda)^\beta-\tilde{\varphi}-s,$$

the Laplacian operator will be replaced by the trace operator $\mathrm{tr}_G\, v = G^{i\bar{j}} v_{i\bar{j}}$. More precisely, we have the estimates

$$\begin{aligned}
0 &\geq L(-\epsilon(-\psi+\Lambda)^\beta - \tilde{\varphi} - s)\\
&\geq -\beta\epsilon(-\psi+\Lambda)^{\beta-1}\dot{\psi} + \dot{\tilde{\varphi}} + \beta\epsilon(-\psi+\Lambda)^{\beta-1}\,\mathrm{tr}_G(\sqrt{-1}\partial\bar{\partial}\psi)\\
&\qquad\qquad + \beta(1-\beta)\epsilon(-\varphi+\Lambda)^{\beta-2}|\partial\varphi|_G^2 - \mathrm{tr}_G(\sqrt{-1}\partial\bar{\partial}\varphi)\\
&\geq \beta\epsilon(-\psi+\Lambda)^{\beta-1}(-\dot{\psi}) - (-\dot{\varphi}) + \beta\epsilon(-\psi+\Lambda)^{\beta-1}\,\mathrm{tr}_G(\sqrt{-1}\partial\bar{\partial}\psi) - \mathrm{tr}_G(\sqrt{-1}\partial\bar{\partial}\varphi) - C_3\\
&\geq \beta\epsilon(-\psi+\Lambda)^{\beta-1}(-\dot{\psi}) - (-\dot{\varphi}) + \beta\epsilon(-\psi+\Lambda)^{\beta-1}\,\mathrm{tr}_G\,\omega_\psi - \mathrm{tr}_G\,\omega_\varphi - C_3\\
&\geq \beta\epsilon(-\psi+\Lambda)^{\beta-1}(-\dot{\psi} + \mathrm{tr}_G\,\omega_\psi) - (-\dot{\varphi} + r + C_3).
\end{aligned}$$

We also must deal with the factor $-\dot{\psi} + \mathrm{tr}_G\,\omega_\psi$ as in inequalities (3-2)–(3-4). The lower bound of the determinant on the condition of $\mathcal{F}$ will give us the lower bound $\det G^{i\bar{j}} \geq \gamma\mathcal{F}^{-n/r}$. By the flows (2-2) and (1-3) and the homogeneous degree $r$ condition, we have

$$\begin{aligned}
-\dot{\psi} + \mathrm{tr}_G\,\omega_\psi &\geq (n+1)\sqrt[n+1]{\frac{f e^{nF}{\omega_0}^n}{{\omega_\psi}^n}\cdot{\omega_\psi}^n \det G^{i\bar{j}}}\\
&\geq C_7 f^{1/(n+1)}\sqrt[n+1]{\left(\frac{e^{rF}}{\mathcal{F}}\right)^{n/r}} \geq C_7 f^{1/(n+1)}\exp\left(-\frac{n}{r(n+1)}\dot{\varphi}\right),
\end{aligned}\tag{5-1}$$

where $C_7$ is a universal constant.

Since the function

$$h(x) = (x - r - C_3)\exp\left(\frac{n}{r(n+1)}x\right) > c,$$

where

$$c = -\frac{r(n+1)}{n}\exp\left(\frac{nC_3 - r}{r(n+1)}\right),$$

we have the same estimate

$$0 \geq C_7\beta\epsilon(-\psi+\Lambda)^{\beta-1} f^{1/(n+1)} - \frac{r(n+1)}{n}\exp\left(\frac{nC_3 - r}{r(n+1)}\right).$$

To derive the ABP estimate and (4-16) for this case, we need to calculate the estimate of the operator $L$ on $\widetilde{\Gamma}$. We have

$$LH \leq -(2n+1)\left(|H_t\cdot\det D^2 H|\cdot\left(\frac{1}{(\det G)^2}\right)\right)^{1/(2n+1)}.$$

Moreover the bad factor in the integration over $D\cap\Omega_+$ is

$$\begin{aligned}
\exp\left(-\frac{n(2n+1)}{r(n+1)}\dot{\varphi}\right)\frac{1}{(\det G)^2} &= e^{(n^2(2n+1)/(r(n+1)))F}\mathcal{F}^{-n(2n+1)/(r(n+1))}\frac{1}{(\det G)^2}\\
&\leq \frac{1}{\gamma^2} e^{(n^2(2n+1)/(r(n+1)))F}\mathcal{F}^{n/(r(n+1))}.
\end{aligned}$$

The exponent of $\mathcal{F}$ is $n/(r(n+1))$, which is less than or equal to 1 if we assume the degree satisfies $1 \leq r \leq n$, and so we have a similar control as in (4-21). Theorem 1.2 follows from an analogue of Theorem 4.2.

We can also consider the much more general flow equation (1-5). Let us list the linearization operators for different choices of $\Theta$ firstly:

$$Lu = \begin{cases} -\frac{\partial}{\partial t}u + \dot{\varphi}\Delta_\varphi u, & \Theta(x) = x, \\ -\frac{\partial}{\partial t}u + \frac{1}{3}\dot{\varphi}\Delta_\varphi u, & \Theta(x) = x^{1/3}, \\ -\frac{\partial}{\partial t}u - \dot{\varphi}\Delta_\varphi u, & \Theta(x) = -1/x. \end{cases} \tag{5-2}$$

To get Lemmas 3.1 and 4.1 under the new setting, we need to reprove Lemma 2.3 to get the upper bound of the integral. The following arguments are divided into two cases.

When $\Theta(y) = -1/y$, we have

$$\int_M \dot{\varphi} = \int_M -e^{nF}\frac{{\omega_0}^n}{{\omega_\varphi}^n} < 0.$$

When $\Theta(y) = y^a$ for $a > 0$, we have

$$\ddot{\varphi} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{{\omega_\varphi}^n}{e^{nF}{\omega_0}^n}\right)^a = a\Delta_\varphi\dot{\varphi}\frac{{\omega_\varphi}^n}{e^{nF}{\omega_0}^n}\left(\frac{{\omega_\varphi}^n}{e^{nF}{\omega_0}^n}\right)^{a-1} = a\dot{\varphi}\Delta_\varphi\dot{\varphi}.$$

Consider the first variation of the functional $\int_M \dot{\varphi}{\omega_\varphi}^n$, given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\int_M \dot{\varphi}{\omega_\varphi}^n &= \int_M \ddot{\varphi}{\omega_\varphi}^n + \int_M \dot{\varphi}\frac{\mathrm{d}}{\mathrm{d}t}({\omega_\varphi}^n) \\ &= \int_M \ddot{\varphi}{\omega_\varphi}^n + \int_M \dot{\varphi}\Delta_\varphi\dot{\varphi}{\omega_\varphi}^n \\ &= (a+1)\int_M \dot{\varphi}\Delta_\varphi\dot{\varphi}{\omega_\varphi}^n \\ &= -(a+1)\int_M |\nabla\dot{\varphi}|^2_{\omega_\varphi}{\omega_\varphi}^n \le 0. \end{aligned}$$

Then the estimate of $\int_M \dot{\varphi}{\omega_0}^n$ follows from

$$\begin{aligned} \int_M \dot{\varphi}{\omega_0}^n &\le \int_M \dot{\varphi}{\omega_0}^n - \int_M \dot{\varphi}{\omega_\varphi}^n + \int_M \dot{\varphi}(\,\cdot\,,0){\omega_{\varphi_0}}^n \\ &\le \int_M \dot{\varphi}({\omega_0}^n - {\omega_\varphi}^n) + \int_M e^{-anF}\left(\frac{{\omega_{\varphi_0}}^n}{{\omega_0}^n}\right)^a {\omega_{\varphi_0}}^n \\ &\le \int_M \dot{\varphi}(1 - e^{nF}\dot{\varphi}^{1/a}){\omega_0}^n + C\int_M e^{-anF}{\omega_0}^n, \end{aligned}$$

where $C$ is universal.

Consider a function $A(y) = y - ly^{1+1/a}$ defined on $y \in [0,\infty)$, where $l$ is positive. Using calculus, we have

$$A(y) \le A\left(\frac{a^a}{l^a(a+1)^a}\right) = \frac{a^a}{(a+1)^{a+1}}l^{-a}.$$

This yields the estimate

$$\int_M \dot{\varphi}{\omega_0}^n \le \frac{a^a}{(a+1)^{a+1}}\int_M e^{-anF}{\omega_0}^n + C\int_M e^{-anF}{\omega_0}^n \le CK.$$

Once we have the above results, we need to prove Lemmas 3.1 and 4.1 for the flow equation (1-5). However the operator $L$ has an extra factor $-\dot{\varphi}$ in the Laplacian term which requires slightly different calculations. We will only discuss the case when $\Theta = -1/y$; the other two cases can be treated similarly. To start with, we have

$$L(-\epsilon(-\psi+\Lambda)^{\beta} - \tilde{\varphi} - s) \geq \beta\epsilon(-\psi+\Lambda)^{\beta-1}(-\dot{\psi} - \dot{\varphi}\,\mathrm{tr}_{\omega_\psi}\omega_\varphi) + \dot{\varphi} + n\dot{\varphi}. \tag{5-3}$$

There is no constant $C_3$ since $\int \dot{\varphi} < 0$, and there is a term $n\dot{\varphi}$ since we have the extra factor when we compute the second derivative. By applying the geometric-arithmetic inequality, we have

$$-\dot{\psi} - \dot{\varphi}\,\mathrm{tr}_{\omega_\psi}\omega_\varphi \geq (n+1)\left(\frac{f e^{nF}{\omega_0}^n}{{\omega_\psi}^n}(-\dot{\varphi})^n \frac{{\omega_\psi}^n}{{\omega_\varphi}^n}\right)^{1/(n+1)} = (n+1) f^{1/(n+1)}(-\dot{\varphi})$$

and

$$L(-\epsilon(-\psi+\Lambda)^{\beta} - \tilde{\varphi} - s) \geq (n+1)(\beta\epsilon(-\psi+\Lambda)^{\beta-1} f^{1/(n+1)} - 1)(-\dot{\varphi}).$$

Therefore we can drop the positive factor $-\dot{\varphi}$ and evaluate the inequality at the maximal point.

To derive an analogue of Lemma 4.1, we need to consider

$$\rho = -\epsilon(-\psi+\Lambda)^{\beta} - \tilde{\varphi} - (n+1)(t-t_0),$$

where the last two terms $-(t-t_0)$ do not affect the result since we only estimate locally on $M \times [t_0, t_0+1]$. If the new $\rho$ has an upper bound which has the same dependencies as the constants in Lemma 4.1, then $-\epsilon(-\psi+\Lambda)^{\beta} - \tilde{\varphi}$ does as well. Following a similar calculation, we have

$$L\rho \geq (n+1)\left(\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)} - 1 + \frac{1}{-\dot{\varphi}}\right)(-\dot{\varphi}),$$

where $1/-\dot{\varphi}$ comes from $L(t-t_0) = -1$. Applying the ABP estimate, we have

$$\left(|\partial_t H \cdot \det D^2 H| \cdot (-\dot{\varphi})^{2n}\left(\frac{{\omega_0}^n}{{\omega_\varphi}^n}\right)^2\right)^{1/(2n+1)} \leq \frac{R_-}{2n+1}. \tag{5-4}$$

In conclusion, the main term in the ABP estimate is

$$\begin{aligned}\left(\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)} - 1 + \frac{1}{-\dot{\varphi}}\right)_-^{2n+1} &(-\dot{\varphi})^{2n+1-2n}\left(\frac{{\omega_\varphi}^n}{{\omega_0}^n}\right)^2 \\ &= \left(\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)} - 1 + \frac{1}{-\dot{\varphi}}\right)_-^{2n+1} e^{2nF}\frac{1}{-\dot{\varphi}}.\end{aligned} \tag{5-5}$$

The term $1/-\dot{\varphi}$ can be controlled pointwisely on the domain $D \cap \Omega_+$. More specifically,

$$\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)} - 1 + \frac{1}{-\dot{\varphi}} < 0$$

implies both inequalities

$$\beta\epsilon(-\psi+\Lambda)^{\beta-1}\tilde{f}^{1/(n+1)} - 1 < 0 \quad \text{and} \quad \frac{1}{-\dot{\varphi}} - 1 < 0.$$

The rest of the proof for the $\Theta = -1/y$ case follows the same procedure.

## Acknowledgements

The author would like to thank Prof. Xiangwen Zhang for helpful discussions and frequent encouragement throughout the project. The author is also grateful for Prof. Duong H. Phong and Prof. Bin Guo's interest and useful comments. Thanks also to Prof. Freid Tong for some private conversations that were enlightening.

## References

[Cao 1985] H. D. Cao, "Deformation of Kähler metrics to Kähler–Einstein metrics on compact Kähler manifolds", *Invent. Math.* **81**:2 (1985), 359–372. MR Zbl

[Chen and Cheng 2021] X. Chen and J. Cheng, "On the constant scalar curvature Kähler metrics, I: A priori estimates", *J. Amer. Math. Soc.* **34**:4 (2021), 909–936. MR Zbl

[Chen and Cheng 2023] X. Chen and J. Cheng, "The $L^\infty$ estimates for parabolic complex Monge–Ampère and Hessian equations", *Pure Appl. Math. Q.* **19**:6 (2023), 2869–2913. MR Zbl

[Eyssidieux et al. 2015] P. Eyssidieux, V. Guedj, and A. Zeriahi, "Weak solutions to degenerate complex Monge–Ampère flows, I", *Math. Ann.* **362**:3-4 (2015), 931–963. MR Zbl

[Eyssidieux et al. 2016] P. Eyssidieux, V. Guedj, and A. Zeriahi, "Weak solutions to degenerate complex Monge–Ampère flows, II", *Adv. Math.* **293** (2016), 37–80. MR Zbl

[Guedj et al. 2021] V. Guedj, C. H. Lu, and A. Zeriahi, "The pluripotential Cauchy–Dirichlet problem for complex Monge–Ampère flows", *Ann. Sci. Éc. Norm. Supér.* (4) **54**:4 (2021), 889–944. MR Zbl

[Guo and Phong 2023] B. Guo and D. H. Phong, "Auxiliary Monge–Ampère equations in geometric analysis", *ICCM Not.* **11**:1 (2023), 98–135. MR Zbl

[Guo and Phong 2024] B. Guo and D. H. Phong, "On $L^\infty$ estimates for fully non-linear partial differential equations", *Ann. of Math.* (2) **200**:1 (2024), 365–398. MR Zbl

[Guo et al. 2023] B. Guo, D. H. Phong, and F. Tong, "On $L^\infty$ estimates for complex Monge–Ampère equations", *Ann. of Math.* (2) **198**:1 (2023), 393–418. MR Zbl

[Harvey and Lawson 2023] F. R. Harvey and H. B. Lawson, Jr., "Determinant majorization and the work of Guo–Phong–Tong and Abja–Olive", *Calc. Var. Partial Differential Equations* **62**:5 (2023), art. id. 153. MR Zbl

[Hörmander 1973] L. Hörmander, *An introduction to complex analysis in several variables*, North-Holland Mathematical Library **7**, North-Holland, 1973. MR Zbl

[Jian and Shi 2024] W. Jian and Y. Shi, "$L^\infty$ estimates for Kähler–Ricci flow on Kähler–Einstein Fano manifolds: a new derivation", *Proc. Amer. Math. Soc.* **152**:3 (2024), 1279–1286. MR Zbl

[Kołodziej 2003] S. Kołodziej, "The Monge–Ampère equation on compact Kähler manifolds", *Indiana Univ. Math. J.* **52**:3 (2003), 667–686. MR Zbl

[Krylov 1976] N. V. Krylov, "Sequences of convex functions, and estimates of the maximum of the solution of a parabolic equation", *Sibirsk. Mat. Ž.* **17**:2 (1976), 290–303, 478. In Russian; translated in *Sib. Math. J.* **17**:2 (1976), 226–236. MR Zbl

[Phong et al. 2019] D. H. Phong, S. Picard, and X. Zhang, "A flow of conformally balanced metrics with Kähler fixed points", *Math. Ann.* **374**:3-4 (2019), 2005–2040. MR Zbl

[Picard and Suan 2024] S. Picard and C. Suan, "Flows of $G_2$-structures associated to Calabi–Yau manifolds", *Math. Res. Lett.* **31**:6 (2024), 1837–1877. MR Zbl

[Picard and Zhang 2020] S. Picard and X. Zhang, "Parabolic complex Monge–Ampère equations on compact Kähler manifolds", pp. 639–665 in *Proceedings of the International Consortium of Chinese Mathematicians* (Taipei, 2018), edited by L. Ji and S.-T. Yau, International, Boston, MA, 2020. MR Zbl

[Tian 1987] G. Tian, "On Kähler–Einstein metrics on certain Kähler manifolds with $C_1(M) > 0$", *Invent. Math.* **89**:2 (1987), 225–246. MR Zbl

[Tso 1985] K. Tso, “On an Aleksandrov–Bakel’man type maximum principle for second-order parabolic equations”, *Comm. Partial Differential Equations* **10**:5 (1985), 543–553. MR Zbl

[Wang et al. 2021] J. Wang, X.-J. Wang, and B. Zhou, “A priori estimate for the complex Monge–Ampère equation”, *Peking Math. J.* **4**:1 (2021), 143–157. MR Zbl

[Yau 1978] S. T. Yau, “On the Ricci curvature of a compact Kähler manifold and the complex Monge–Ampère equation, I”, *Comm. Pure Appl. Math.* **31**:3 (1978), 339–411. MR Zbl

Received 29 Jun 2023. Revised 12 Jul 2024. Accepted 20 Sep 2024.

QIZHI ZHAO: qizhiz3@uci.edu
*University of California, Irvine, Irvine, CA, United States*

**mathematical sciences publishers** msp

ANALYSIS AND PDE
Vol. 18 (2025), No. 8, pp. 1897–1920
DOI: 10.2140/apde.2025.18.1897

# SPECTRAL ASYMPTOTICS OF THE NEUMANN LAPLACIAN WITH VARIABLE MAGNETIC FIELD ON A SMOOTH BOUNDED DOMAIN IN THREE DIMENSIONS

MAHA AAFARANI, KHALED ABOU ALFA, FRÉDÉRIC HÉRAU AND NICOLAS RAYMOND

This article is devoted to the semiclassical spectral analysis of the Neumann magnetic Laplacian on a smooth bounded domain in three dimensions. Under a generic assumption on the variable magnetic field (involving a localization of the eigenfunctions near the boundary), we establish a semiclassical expansion of the lowest eigenvalues. In particular, we prove that the eigenvalues become simple in the semiclassical limit.

## 1. Motivation and main result

**1.1.** ***The operator.*** Let $\Omega \subset \mathbb{R}^3$ be a smooth connected open bounded domain. We consider $\boldsymbol{A} : \overline{\Omega} \to \mathbb{R}^3$, a smooth magnetic vector potential. The associated magnetic field is given by

$$\boldsymbol{B}(x) = \nabla \times \boldsymbol{A}(x)$$

and assumed to be nonvanishing on $\overline{\Omega}$. For $h > 0$, we consider the self-adjoint operator

$$\mathscr{L}_h = (-ih\nabla - \boldsymbol{A})^2 \tag{1-1}$$

with domain

$$\operatorname{Dom}(\mathscr{L}_h) = \{\psi \in H^2(\Omega) : \boldsymbol{n} \cdot (-ih\nabla - \boldsymbol{A})\psi = 0 \text{ on } \partial\Omega\},$$

where $\boldsymbol{n}$ is the outward pointing normal to the boundary.

The associated quadratic form is defined, for all $\psi \in H^1(\Omega)$, by

$$\mathcal{Q}_h(\psi) = \int_\Omega |(-ih\nabla - \boldsymbol{A})\psi|^2 \,\mathrm{d}x.$$

Since $\Omega$ is smooth and bounded, the operator $\mathscr{L}_h$ has compact resolvent and we can consider the nondecreasing sequence of its eigenvalues $(\lambda_n(h))_{n\geqslant 1}$ (repeated according to their multiplicities). The aim of this article is to describe the behavior of the eigenvalues $\lambda_n(h)$ in the semiclassical limit $h \to 0$.

**1.2.** ***The operator on a half-space with constant magnetic field.*** The boundary of $\Omega$ has an important influence on the spectral asymptotics. Let us consider $x_0 \in \partial\Omega$ and the angle $\theta(x_0) \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ given by

$$\boldsymbol{B}(x_0) \cdot \boldsymbol{n}(x_0) = \|\boldsymbol{B}(x_0)\| \sin(\theta(x_0)),$$

where $\boldsymbol{n}(x_0)$ is the outward pointing normal at $x_0$.

*MSC2020:* 35Pxx, 81Q10.
*Keywords:* semiclassical analysis, magnetic Laplacian, boundary.

© 2025 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY).
Open Access made possible by subscribing institutions via Subscribe to Open.

Near $x_0$, one will approximate $\Omega$ by the half-space $\mathbb{R}^3_+ = \{(r, s, t) \in \mathbb{R}^3 : t > 0\}$ (the variable $t$ playing the role of the distance to the boundary). Then, this will lead us to consider the Neumann realization of

$$\mathfrak{L}_\theta = (D_r - t\cos\theta + s\sin\theta)^2 + D_s^2 + D_t^2$$

in the ambient space $L^2(\mathbb{R}^3_+)$, which already appeared in [Lu and Pan 2000] in the context of Ginzburg–Landau theory. We use the notation $D = -i\,\partial$. The corresponding magnetic field is $\boldsymbol{b}(\theta) = (0, \cos\theta, \sin\theta)$. We let

$$\boldsymbol{e}(\theta) = \inf \operatorname{sp}(\mathfrak{L}_\theta).$$

It is well known (see [Helffer and Morame 2002; Lu and Pan 2000] and also [Raymond 2017, Section 2.5.2]) that $\boldsymbol{e}$ is even, continuous and increasing on $\left[0, \frac{\pi}{2}\right]$ (from $\Theta_0 := \boldsymbol{e}(0) \in (0, 1)$ to 1) and analytic on $\left(0, \frac{\pi}{2}\right)$. Moreover, we can prove that, for all $\theta \in \left(0, \frac{\pi}{2}\right)$, $\boldsymbol{e}(\theta)$ is also the groundstate energy of the Neumann realization of the "Lu–Pan" operator, acting on $L^2(\mathbb{R}^2_+)$,

$$\mathcal{L}_\theta = (t\cos\theta - s\sin\theta)^2 + D_s^2 + D_t^2; \tag{1-2}$$

see [Raymond 2017, Section 0.1.5.4]. In this case, the groundstate energy belongs to the discrete spectrum and it is a simple eigenvalue.

These considerations lead us to introduce the function $\beta$ on the boundary.

**Definition 1.1.** We let, for all $x \in \partial\Omega$,

$$\beta(x) = \|\boldsymbol{B}(x)\| \boldsymbol{e}(\theta(x)).$$

**1.3.** ***Context, known results, and main theorem.*** The function $\beta$ plays a central role in the semiclassical spectral asymptotics. The one-term asymptotics of $\lambda_1(h)$ are established in [Lu and Pan 2000] (see also [Raymond 2010a] and [Fournais and Helffer 2010], where additional details are provided).

**Theorem 1.2** [Lu and Pan 2000]. *We have*

$$\lambda_1(h) = h\min(b_{\min}, \beta_{\min}) + o(h),$$

*where* $b_{\min} = \min_{x\in\overline{\Omega}} \|\boldsymbol{B}(x)\|$ *and* $\beta_{\min} = \min_{x\in\partial\Omega} \beta(x)$.

When $\boldsymbol{B}$ is constant (or with constant norm), more accurate estimates of the groundstate energy have been obtained in [Helffer and Kachmar 2023; Helffer and Morame 2004; Raymond 2010b]. When looking at Theorem 1.2, natural questions can be asked. Can we describe more than the groundstate energy? Is the groundstate energy a simple eigenvalue? In three dimensions, most of the results in this direction have been obtained rather recently:

- When $b_{\min} < \beta_{\min}$, we can prove that the boundary is essentially not seen by the eigenfunctions with low eigenvalues and that they are localized near the minima of $\|\boldsymbol{B}\|$. Then, if the minimum is unique and nondegenerate, the analysis of [Helffer et al. 2016] applies and it can be established that

$$\lambda_n(h) = b_{\min}h + C_0h^{3/2} + (C_1(2n-1) + C_2)h^2 + o(h^2),$$

where the constants $(C_0, C_1, C_2) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}$ reflect the classical dynamics in a magnetic field.

• When $\boldsymbol{B}$ is constant (or with constant norm), we can prove that $\beta_{\min} < b_{\min}$ and that $\beta_{\min} = \Theta_0 \|\boldsymbol{B}\|$. In this case, the eigenfunctions with low eigenvalues are localized near the points of the boundary where the magnetic field is tangent, that is, where $\boldsymbol{e}(\theta(x))$ is minimal. Assuming that the magnetic field becomes generically tangent to the boundary along a nice closed curve and assuming also a nondegeneracy assumption, we have, from [Hérau and Raymond 2024],

$$\lambda_n(h) = \beta_{\min} h + C_0 h^{4/3} + C_1 h^{3/2} + (C_2(2n-1) + C_3) h^{5/3} + o(h^2)$$

for some constants $(C_0, C_1, C_2, C_3) \in \mathbb{R}^2 \times \mathbb{R}_+ \times \mathbb{R}$.

The result in [Hérau and Raymond 2024] is stated in the case of a constant magnetic field, but only the fact that its norm is constant is actually used in the analysis; see Section 3.2.1 in that same work. Note that without the additional nondegeneracy assumption and stopping the analysis before Section 5.6 in that same work provides us with the two-term expansion. This observation is motivated by [Helffer and Kachmar 2023], where the two-term expansion of the groundstate energy has been obtained independently and where examples are also analyzed in detail.

When $\beta_{\min} < b_{\min}$ and when $\|\boldsymbol{B}\|$ is variable, it seems that less is known. The first estimates of the low-lying eigenvalues, and not only of the first one, are done in [Raymond 2010a] (see also [Raymond 2009]), where an upper bound is obtained under a generic assumption (see Assumption 1.3 below):

$$\lambda_n(h) \leqslant \beta_{\min} h + C_0 h^{3/2} + (C_1(2n-1) + C_2) h^2 + o(h^2) \tag{1-3}$$

for some constants $(C_0, C_1, C_2) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}$ and where $C_1$ is explicitly given by

$$C_1 = \frac{\sqrt{\det \operatorname{Hess}_{x_0} \beta}}{2\|\boldsymbol{B}(x_0)\| \sin\theta(x_0)}.$$

The upper bound (1-3) is obtained by means of a construction of quasimodes in local coordinates near the minimum of $\beta$ and involves a number of rather subtle algebraic cancellations. At a conference in Dijon in March 2010, S. Vũ Ngọc suggested to the last author that these algebraic cancellations were the signs of a hidden normal form. At the same conference, J. Sjöstrand also suggested that a dimensional reduction in the Grushin spirit (see the remarkable survey [Sjöstrand and Zworski 2007]) could provide us with the lower bound. Retrospectively, we will see that both of them were somewhat right, but that some microlocal techniques needed to be developed further in order to tackle the problem in an efficient way.

Until now, the matching lower bound to (1-3) has only been obtained for a toy model in the case of a flat boundary with an explicit polynomial magnetic field; see [Raymond 2012]. The aim of this article is to establish a lower bound that matches (1-3) in the general case. To do so, we will, of course, work under the same assumption as in [Raymond 2010a].

**Assumption 1.3.** *The function $\beta$ has a unique minimum, which is nondegenerate. It is attained at $x_0 \in \partial\Omega$, and we have*

$$\theta(x_0) \in \left(0, \tfrac{\pi}{2}\right). \tag{1-4}$$

*Moreover, we have*

$$\beta_{\min} = \beta(x_0) = \min_{x \in \partial\Omega} \beta(x) < \min_{x \in \overline{\Omega}} \|\boldsymbol{B}(x)\| = b_{\min}.$$

The main result of this article is a three-term expansion of the $n$-th eigenvalue of $\mathscr{L}_h$. Thereby, it completes the picture described above.

**Theorem 1.4.** *Under Assumption 1.3, there exist $C_0, C_1 \in \mathbb{R}$ such that, for all $n \geqslant 1$, we have*

$$\lambda_n(h) \underset{h\to 0}{=} \beta_{\min} h + C_0 h^{3/2} + \left( \frac{\sqrt{\det \operatorname{Hess}_{x_0} \beta}}{\|\boldsymbol{B}(x_0)\| \sin\theta(x_0)} \left(n - \frac{1}{2}\right) + C_1 \right) h^2 + o(h^2).$$

*In particular, for all $n \geqslant 1$, $\lambda_n(h)$ becomes a simple eigenvalue as soon as $h$ is small enough.*

**1.4.** ***Organization and strategy of the proof.*** In Section 2, we recall the already known results of localization of the eigenfunctions near $x_0$. This formally reduces the spectral analysis to a neighborhood of $x_0$. This suggests that we should introduce local coordinates near $x_0$. These coordinates $(r, s, t)$ are adapted to the geometry of the magnetic field: the coordinate $s$ is the curvilinear coordinate along the projection of the magnetic field on the boundary (we use here that $\theta(x_0) < \frac{\pi}{2}$), the coordinate $r$ is the geodesic coordinate transverse to $s$, and $t$ is the distance to the boundary. A rather similar coordinate system has been used and described in [Hérau and Raymond 2024] (inspired from [Helffer and Morame 2004]). Then, the local action of the operator is described in Section 2.3, where we perform a Taylor expansion with respect to the normal variable $t$ only. After a local change of gauge, this makes an approximate magnetic vector potential appear, see (2-10). In Section 2.3.2, we define a new operator on $L^2(\mathbb{R}^3_+)$ by extending the coefficients, seen as functions of $(r, s)$ defined near $(0, 0)$, to functions on $\mathbb{R}^2$. Since this extension occurs away from the localization zone of the eigenfunctions, we get a new operator $\mathscr{L}_h^{\mathrm{app}}$ whose spectrum is close to that of $\mathscr{L}_h$, see Proposition 2.11.

In Section 3, we perform the analysis of $\mathscr{L}_h^{\mathrm{app}}$ with the help of the change of coordinates $(r, s) \mapsto \mathscr{J}(r, s) = (u_1, u_2)$, whose geometric role is to make the normal component of the magnetic field constant (here, we use $\theta(x_0) > 0$). This idea is reminiscent of [Morin et al. 2023] in two dimensions; see Proposition 2.2 in that work. We are reduced to the spectral analysis of the operator $\mathscr{N}_h$, see (3-1). Then, we conjugate $\mathscr{N}_h$ by a tangential Fourier transform (in the direction $u_1$) and a translation/dilation $T$ (after these transforms, the variable $u_1$ becomes $z$). After these explicit transforms, we get a new operator $\mathscr{N}_\hbar^{\sharp}$, which can be seen as a differential operator of order 2 in the variables $(z, t)$ with coefficients that are $h$-pseudodifferential operators (with an expansion in powers of $\hbar = h^{1/2}$) in the variable $u_2$ only, see (3-10). Its eigenfunctions are localized in $(z, t)$, see Proposition 3.3 and Remark 3.4.

In Section 4, this localization with respect to $z$ suggests that we should insert cutoff functions in the coefficients of our operator. By doing this, we get the operator $\mathscr{N}_\hbar^{\flat}$, see (4-1). The advantage of $\mathscr{N}_\hbar^{\flat}$ is that it can be considered as a pseudodifferential operator with operator-valued symbol in a reasonable class $S(\mathbb{R}^2, N)$, see Proposition 4.2. The principal operator symbol $n_0(u, \upsilon)$ is unitarily equivalent to the Lu–Pan operator $\|\boldsymbol{B}(\upsilon, -u)\| \mathcal{L}_{\theta(\upsilon, -u)}$ (where we make a slight abuse of notation by forgetting the reference to the local coordinates on the boundary), see Proposition 4.4. Then, we may construct an inverse for $n_0 - \Lambda$ by means of the so-called Grushin formalism as soon as $\Lambda$ is close to $\beta_{\min}$, see Lemma 4.5. This is the first step in the approximate parametrix construction for $\mathscr{N}_\hbar^{\flat} - \Lambda$ given in Proposition 4.7, which is the key of the proof of Theorem 1.4. Let us emphasize that this parametrix construction is inspired by [Keraval 2018] and based on ideas developed by A. Martinez and J. Sjöstrand. This formalism has recently been used in [Hérau and Raymond 2024] in three dimensions (see also [Bonnaillie-Noël et al. 2022; Fahs et al. 2024; Fournais et al. 2023] in the case of two dimensions). At a formal level, this parametrix construction relates the kernel of $\mathscr{N}_\hbar^\flat - \Lambda$ to that of an effective pseudodifferential operator $Q_\hbar^\pm(\Lambda)$, see (4-7).

In Section 5 we relate the spectrum of $\mathscr{N}_\hbar^\sharp$ to that of the effective operator $(p_\hbar^{\mathrm{eff}})^W$, see (5-1). Note: the effective operator is an operator in one dimension. This contrasts with [Hérau and Raymond 2024], where a double Grushin reduction is used: here this reduction is done in one step with the help of the Lu–Pan operator. The quasi-parametrix in Proposition 4.7 is the bridge between the spectra of $\mathscr{N}_\hbar^\sharp$ and $(p_\hbar^{\mathrm{eff}})^W$.

We emphasize that we have to be very careful when studying this connection since the symbol of the effective operator is not necessarily real-valued (only its principal symbol $p_0$ is a priori real). This again contrasts with [Hérau and Raymond 2024] and all the previous works on the subject. This non-self-adjointness comes from the fact that $\mathscr{N}_h$ is not self-adjoint on the canonical $L^2$-space but on a weighted $L^2$-space. That is why a short detour into the world of non-self-adjoint operators is used in Section 5. In fact, one will not need the operator $(p_\hbar^{\mathrm{eff}})^W$ more than its approximation $(p_\hbar^{\mathrm{mod}})^W$ near the minimum of $p_0$, see Section 5.1. This approximation is a complex perturbation of the harmonic oscillator. Its spectrum is well known as well as the behavior of its resolvent.

In Section 5.2.1, we use rescaled Hermite functions to construct quasimodes for $\mathscr{N}_\hbar^\sharp$. This shows that the spectrum of the model operator is in fact real, and we get an accurate upper bound of $\lambda_n(\mathscr{N}_\hbar^\sharp)$ in (5-5). This reproves in a much shorter way (1-3) (see [Raymond 2010a, Theorem 1.5], where the convention $\|\boldsymbol{B}(x_0)\| = 1$ is used). Section 5.2.2 is devoted to establishing the corresponding lower bound (by using in particular that the eigenvalues of the non-self-adjoint operator $(p_\hbar^{\mathrm{mod}})^W$ have algebraic multiplicity 1).

**Remark 1.5.** The above analysis explains the presence of $\beta_{\min}$, attached to the lowest eigenvalue of the Lu–Pan operator, as the leading term in the semiclassical asymptotics. Similarly, the constant

$$\frac{\sqrt{\det \operatorname{Hess}_{x_0} \beta}}{\|\boldsymbol{B}(x_0)\| \sin\theta(x_0)}$$

appears as the uncertainty constant attached to the effective harmonic oscillator $(p_\hbar^{\mathrm{mod}})^W$ after the Grushin reduction to a one-dimensional problem. This spectral gap combines the normal component of the magnetic field with the spectrum of the Lu–Pan operator (and thus it has no obvious dynamical interpretation). The latter is deeply related to the nondegeneracy assumption on $\beta$. However, the geometric interpretation of the constants $C_0$ and $C_1$ is not clear since they come from the non-self-adjoint linear part of $(p_\hbar^{\mathrm{mod}})^W$.

## 2. Localization near $x_0$ and consequences

**2.1.** ***Localization estimates.*** In this section, we gather some already-known localization properties of the eigenfunctions; see [Raymond 2009].

**Proposition 2.1** (localization near the boundary). *Again under Assumption 1.3, for all $\epsilon > 0$ such that $\beta_{\min} + \epsilon < b_{\min}$, there exist $\alpha$, $C$, $h_0 > 0$ such that, for all $h \in (0, h_0)$ and all eigenfunctions $\psi$ of $\mathscr{L}_h$ associated with an eigenvalue $\lambda \leqslant (\beta_{\min} + \epsilon)h$, we have*

$$\int_\Omega e^{2\alpha \operatorname{dist}(x,\partial\Omega)/\sqrt{h}} |\psi|^2 \, dx \leqslant C\|\psi\|^2. \tag{2-1}$$

For $\delta > 0$, we consider the $\delta$-neighborhood of the boundary given by

$$\Omega_\delta := \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) < \delta\}.$$

Due to Proposition 2.1, in the following, we take

$$\delta = h^{1/2-\eta}$$

for $\eta \in \left(0, \frac{1}{2}\right)$. We consider $\mathscr{L}_{h,\delta} = (-ih\nabla - \boldsymbol{A})^2$, the operator with magnetic Neumann condition on $\partial\Omega$ and Dirichlet condition on $\partial\Omega_\delta \setminus \partial\Omega$.

**Corollary 2.2.** *Let $n \geqslant 1$. There exist $C, h_0 > 0$ such that, for all $h \in (0, h_0)$,*

$$\lambda_n(\mathscr{L}_{h,\delta}) - Ce^{-Ch^{-\eta}} \leqslant \lambda_n(\mathscr{L}_h) \leqslant \lambda_n(\mathscr{L}_{h,\delta}).$$

Note that the upper bound in Corollary 2.2 easily follows from the min-max theorem, whereas the lower bound is obtained by using Proposition 2.1.

Thanks to Corollary 2.2, we may focus on the spectral analysis of $\mathscr{L}_{h,\delta}$. The following proposition can be found in [Fournais and Helffer 2010, Chapter 9] and [Helffer and Morame 2002, Theorem 4.3] (see also the proof of [Hérau and Raymond 2024, Proposition 2.9]).

**Proposition 2.3** (localization near $x_0$). *Let $M > 0$. There exist $C, h_0 > 0$ and $\alpha > 0$ such that, for all $h \in (0, h_0)$ and all eigenfunctions $\psi$ of $\mathscr{L}_{h,\delta}$ associated with an eigenvalue $\lambda$ such that $\lambda \leqslant \beta_{\min} h + Mh^{3/2}$, we have*

$$\int_{\Omega_\delta} e^{2\alpha \operatorname{dist}(x,\partial\Omega)/\sqrt{h}} |\psi(x)|^2 \,\mathrm{d}x + \int_{\Omega_\delta} e^{2\alpha \|x-x_0\|^2/h^{1/4}} |\psi(x)|^2 \,\mathrm{d}x \leqslant C\|\psi\|^2. \tag{2-2}$$

Proposition 2.3 invites us to consider a local chart near $x_0$ and to write the operator in the corresponding coordinates. In order to simplify our analysis, we construct below a system of coordinates compatible with the geometry of the magnetic field.

**2.2.** ***Adapted coordinates near $x_0$.*** This section is devoted to introducing coordinates adapted to the magnetic field. Most of the properties of our coordinates system have been established in [Hérau and Raymond 2024].

**2.2.1.** *Coordinate in the direction of the magnetic field on the boundary.* We set

$$\boldsymbol{b}(x) = \frac{\boldsymbol{B}(x)}{\|\boldsymbol{B}(x)\|},$$

and we consider its projection on the tangent plane at $x \in \partial\Omega$:

$$\boldsymbol{b}^{\|}(x) = \boldsymbol{b}(x) - \langle \boldsymbol{b}(x), \boldsymbol{n}(x)\rangle \boldsymbol{n}(x),$$

where $\boldsymbol{n}$ is the outward pointing normal.

Due to Assumption 1.3, near $x_0$, the vector field $\boldsymbol{b}^{\|}$ does not vanish. This allows us to consider the unit vector field

$$\boldsymbol{f}(x) = \frac{\boldsymbol{b}^{\|}(x)}{\|\boldsymbol{b}^{\|}(x)\|}$$

and the associated integral curve $\gamma$ given by

$$\gamma'(s) = \boldsymbol{f}(\gamma(s)), \quad \gamma(0) = x_0,$$

which is well-defined on $(-s_0, s_0)$ for some $s_0 > 0$. Clearly, $\gamma$ is smooth and with values in $\partial\Omega$.

**2.2.2.** *Coordinates on the boundary.* Denoting by $K$ the second fundamental form of $\partial\Omega$ associated to the Weingarten map defined by,

$$\text{for all } U, V \in T_x\partial\Omega, \quad K_x(U, V) = \langle \mathrm{d}\boldsymbol{n}_x(U), V\rangle,$$

we can consider the ODE with parameter $s$ of unknown $r \mapsto \gamma(r, s)$,

$$\partial_r^2\gamma(r, s) = -K(\partial_r\gamma(r, s), \partial_r\gamma(r, s))\boldsymbol{n}(\gamma(r, s)),$$

with initial conditions

$$\gamma(0, s) = \gamma(s), \quad \partial_r\gamma(0, s) = -\gamma'(s)^{\perp},$$

where $\perp$ is taken in the tangent space and such that $(\gamma', \gamma'^{\perp}, \boldsymbol{n})$ is a direct orthonormal basis. The minus is here so that $(\partial_r\gamma, \partial_s\gamma, \boldsymbol{n})$ is also a direct orthonormal basis along $\gamma(\,\cdot\,)$. The curve $\gamma(r, \cdot\,)$ is the image of $\gamma(\,\cdot\,)$ under the geodesic flow on $\partial\Omega$ (with initial velocity orthogonal to $\gamma(\,\cdot\,)$) at time $r$.

This ODE has a unique smooth solution $(-r_0, r_0) \times (-s_0, s_0) \ni (r, s) \mapsto \gamma(r, s)$, where $r_0 > 0$ is chosen small enough. Let us gather the important properties of $(r, s) \mapsto \gamma(r, s)$. Their proofs may be found in [Hérau and Raymond 2024].

**Proposition 2.4.** *The function $(r, s) \mapsto \gamma(r, s)$ is valued in $\partial\Omega$. Moreover, we have*

$$|\partial_r\gamma(r, s)| = 1, \quad \langle \partial_r\gamma, \partial_s\gamma\rangle = 0.$$

*In this chart $\gamma$, the first fundamental form on $\partial\Omega$ is given by the matrix*

$$g(r, s) = \begin{pmatrix} 1 & 0 \\ 0 & \alpha(r, s) \end{pmatrix}, \quad \alpha(r, s) = |\partial_s\gamma(r, s)|^2.$$

*For all $s \in (-s_0, s_0)$, we have $\alpha(0, s) = 1$ and $\partial_s\alpha(0, s) = 0$.*

**2.2.3.** *Coordinates near the boundary.* We consider the tubular coordinates associated with the chart $\gamma$:

$$y = (r, s, t) \mapsto \Gamma(r, s, t) = \gamma(r, s) - t\boldsymbol{n}(\gamma(r, s)) = x. \tag{2-3}$$

The map $\Gamma$ is a smooth diffeomorphism from $Q_0 := (-r_0, r_0) \times (-s_0, s_0) \times (0, t_0)$ to $\Gamma(Q_0)$, as soon as $t_0 > 0$ is chosen small enough. The differential of $\Gamma$ can be written as

$$\mathrm{d}\Gamma_y = [(\mathrm{Id} - t\,\mathrm{d}\boldsymbol{n})(\partial_r\gamma), (\mathrm{Id} - t\,\mathrm{d}\boldsymbol{n})(\partial_s\gamma), -\boldsymbol{n}], \tag{2-4}$$

and the Euclidean metric becomes

$$\boldsymbol{G} = (\mathrm{d}\Gamma)^{\mathrm{T}}\,\mathrm{d}\Gamma = \begin{pmatrix} \boldsymbol{g} & 0 \\ 0 & 1 \end{pmatrix}, \tag{2-5}$$

with

$$\boldsymbol{g}(r,s,t)=\begin{pmatrix}\|(\mathrm{Id}-t\,\mathrm{d}\boldsymbol{n})(\partial_r\gamma)\|^2 & \langle(\mathrm{Id}-t\,\mathrm{d}\boldsymbol{n})(\partial_r\gamma),(\mathrm{Id}-t\,\mathrm{d}\boldsymbol{n})(\partial_s\gamma)\rangle\\ \langle(\mathrm{Id}-t\,\mathrm{d}\boldsymbol{n})(\partial_r\gamma),(\mathrm{Id}-t\,\mathrm{d}\boldsymbol{n})(\partial_s\gamma)\rangle & \|(\mathrm{Id}-t\,\mathrm{d}\boldsymbol{n})(\partial_s\gamma)\|^2\end{pmatrix}.$$

We have $g(r,s)=\boldsymbol{g}(r,s,0)$, where $g$ is defined in Proposition 2.4.

**2.2.4.** *The magnetic form in tubular coordinates.* In this section, we discuss the expression of the magnetic field in the coordinates induced by $\Gamma$. This discussion can be found in [Raymond 2017, Section 0.1.2.2] and [Hérau and Raymond 2024, Section 3.2]. We consider the 1-form

$$\sigma=\boldsymbol{A}\cdot\mathrm{d}x=\sum_{\ell=1}^{3}A_\ell\,\mathrm{d}x_\ell.$$

Its exterior derivative is the magnetic 2-form

$$\omega=\mathrm{d}\sigma=\sum_{1\leqslant k<\ell\leqslant 3}(\partial_k A_\ell-\partial_\ell A_k)\,\mathrm{d}x_k\wedge\mathrm{d}x_\ell,$$

which can also be written as

$$\omega=B_3\,\mathrm{d}x_1\wedge\mathrm{d}x_2-B_2\,\mathrm{d}x_1\wedge\mathrm{d}x_3+B_1\,\mathrm{d}x_2\wedge\mathrm{d}x_3.$$

Note also that,

$$\text{for all } U,V\in\mathbb{R}^3,\quad \omega(U,V)=\det(U,V,\boldsymbol{B})=\langle U\times V,\boldsymbol{B}\rangle.$$

Let us now consider the effect of the change of variables $\Gamma(y)=x$. We have

$$\Gamma^*\sigma=\sum_{j=1}^{3}\tilde{A}_j\,\mathrm{d}y_j,\quad \tilde{\boldsymbol{A}}=(\mathrm{d}\Gamma)^{\mathrm{T}}\circ\boldsymbol{A}\circ\Gamma,\tag{2-6}$$

and

$$\Gamma^*\omega=\Gamma^*\,\mathrm{d}\sigma=\mathrm{d}(\Gamma^*\sigma)=[\,\cdot\,,\cdot\,,\nabla\times\tilde{\boldsymbol{A}}].$$

Here we use the notation $\Gamma^*$ for the pullback by $\Gamma$. This also gives that, for all $U,V\in\mathbb{R}^3$,

$$\det(\mathrm{d}\Gamma(U),\mathrm{d}\Gamma(V),\boldsymbol{B})=\det(U,V,\nabla\times\tilde{\boldsymbol{A}})\quad\text{or}\quad\det\mathrm{d}\Gamma(\,\cdot\,,\cdot\,,\mathrm{d}\Gamma^{-1}(\boldsymbol{B}))=\det(\,\cdot\,,\cdot\,,\nabla\times\tilde{\boldsymbol{A}}),$$

so that,

$$\nabla\times\tilde{\boldsymbol{A}}=(\det\mathrm{d}\Gamma)\,\mathrm{d}\Gamma^{-1}(\boldsymbol{B}).$$

Note then that, using (2-5), we get

$$|\boldsymbol{g}|^{-1/2}\nabla\times\tilde{\boldsymbol{A}}=\mathcal{B},\tag{2-7}$$

where $\mathcal{B}(y):=\mathrm{d}\Gamma_y^{-1}(\boldsymbol{B}(x))$ corresponds to the coordinates of $\boldsymbol{B}(y)$ in the image of the canonical basis by $\mathrm{d}\Gamma_y$. With our specific change of coordinates (2-3), we have

$$\boldsymbol{B}=\mathrm{d}\Gamma(\mathcal{B})=\mathcal{B}_1(\mathrm{Id}-t\,\mathrm{d}\boldsymbol{n})(\partial_r\gamma)+\mathcal{B}_2(\mathrm{Id}-t\,\mathrm{d}\boldsymbol{n})(\partial_s\gamma)-\mathcal{B}_3\boldsymbol{n}.$$

For all $x \in \partial\Omega$, i.e., $t = 0$, we have

$$\begin{aligned} \boldsymbol{B}(x) &= \mathcal{B}_1(r,s,0)\,\partial_r\gamma + \mathcal{B}_2(r,s,0)\,\partial_s\gamma - \mathcal{B}_3(r,s,0)\boldsymbol{n}(\gamma(r,s)), \\ \|\boldsymbol{B}(x)\|^2 &= \mathcal{B}_1^2(r,s,0) + \alpha(r,s)\mathcal{B}_2^2(r,s,0) + \mathcal{B}_3^2(r,s,0). \end{aligned} \tag{2-8}$$

Moreover, we have

$$\mathcal{B}_1(r,s,0) = \langle \boldsymbol{B}, \partial_r\gamma\rangle, \quad \alpha(r,s)\mathcal{B}_2(r,s,0) = \langle \boldsymbol{B}, \partial_s\gamma\rangle, \quad \mathcal{B}_3(r,s,0) = -\langle \boldsymbol{B}, \boldsymbol{n}\rangle.$$

Note that our choice of coordinate $s$ (along the projection of the magnetic field on the tangent plane) and of transverse coordinate $r$ implies that

$$\mathcal{B}_1(0,s,0) = 0, \quad \mathcal{B}_2(0,s,0) > 0,$$

thanks to Assumption 1.3.

**Definition 2.5.** In a neighborhood of $(0,0)$, we can consider the unique smooth function $\theta$ such that

$$\boldsymbol{B}(\gamma(r,s)) \cdot \boldsymbol{n}(\gamma(r,s)) = \|\boldsymbol{B}(\gamma(r,s))\| \sin\theta(r,s)$$

and satisfying $\theta(r,s) \in \left(0, \frac{\pi}{2}\right)$. With a slight abuse of notation, we let

$$\beta(r,s) = \|\boldsymbol{B}(\gamma(r,s))\|\boldsymbol{e}(\theta(r,s)).$$

**Remark 2.6.** We have

$$\mathcal{B}_3(r,s) = -\|\boldsymbol{B}(\gamma(r,s))\| \sin(\theta(r,s)).$$

Moreover, since $\mathcal{B}_2 > 0$ and $\alpha(0,s) = 1$,

$$\mathcal{B}_2(0,s,0) = \|\boldsymbol{B}(\gamma(0,s))\| \cos\theta(0,s), \quad \mathcal{B}_3(0,s,0) = -\|\boldsymbol{B}(\gamma(0,s))\| \sin\theta(0,s).$$

In fact, we can choose a suitable explicit $\tilde{\boldsymbol{A}}$ such that (2-7) holds in a neighborhood of $(0,0,0)$.

**Lemma 2.7.** *Considering*

$$\begin{aligned} \tilde{A}_1(r,s,t) &= \int_0^t [|\boldsymbol{g}|^{1/2}\mathcal{B}_2](r,s,\tau)\,\mathrm{d}\tau, \\ \tilde{A}_2(r,s,t) &= -\int_0^t [|\boldsymbol{g}|^{1/2}\mathcal{B}_1](r,s,\tau)\,\mathrm{d}\tau + \int_0^r [|\boldsymbol{g}|^{1/2}\mathcal{B}_3](u,s,0)\,\mathrm{d}u, \\ \tilde{A}_3(r,s,t) &= 0, \end{aligned}$$

*we have* $\nabla \times \tilde{\boldsymbol{A}}(r,s,t) = |\boldsymbol{g}|^{1/2}\mathcal{B}(r,s,t)$.

*Proof.* This follows from a straightforward computation and the fact that $|\boldsymbol{g}|^{1/2}\mathcal{B}$ is divergence-free. ☐

**Remark 2.8.** Note that the proof of Lemma 2.7 does not involve global geometric quantities on the boundary as in [Hérau and Raymond 2024, Proposition 3.3], since our analysis is local near $x_0$.

**2.3.** ***First approximation of the magnetic Laplacian in local coordinates.*** If the support of $\psi$ is close enough to $x_0$, we may express $\mathcal{Q}_h(\psi)$ in the local chart given by $\Gamma(y) = x$. Letting $\tilde{\psi}(y) = \psi \circ \Gamma(y)$, we have then

$$\mathcal{Q}_h(\psi) = \int \langle \boldsymbol{G}^{-1}(-ih\nabla_y - \tilde{\boldsymbol{A}}(y))\tilde{\psi}, (-ih\nabla_y - \tilde{\boldsymbol{A}}(y))\tilde{\psi}\rangle |\boldsymbol{g}|^{1/2}\,\mathrm{d}y.$$

In the Hilbert space $L^2(|\boldsymbol{g}|^{1/2}\,\mathrm{d}y)$, the operator locally takes the form

$$|\boldsymbol{g}|^{-1/2}(-ih\nabla_y - \tilde{\boldsymbol{A}}(y)) \cdot |\boldsymbol{g}|^{1/2}\boldsymbol{G}^{-1}(-ih\nabla_y - \tilde{\boldsymbol{A}}(y)), \tag{2-9}$$

where $\boldsymbol{G}$ is defined in (2-5). From now on, the analysis deviates from [Hérau and Raymond 2024].

**2.3.1.** *Expansion with respect to $t$.* Due to the localization near the boundary at the scale $h^{1/2}$, we are led to replace $\tilde{\boldsymbol{A}}$ by its Taylor expansion $\tilde{\boldsymbol{A}}^{[3]}$ at order 3 and $\boldsymbol{g}$ and $\boldsymbol{G}$ by their Taylor expansions at order 2. We let

$$\begin{aligned}
\tilde{A}_1^{[3]}(r,s,t) &= t[|\boldsymbol{g}|^{1/2}\mathcal{B}_2](r,s,0) + C_2\hat{t}^2 + C_3\hat{t}^3,\\
\tilde{A}_2^{[3]}(r,s,t) &= -t[|\boldsymbol{g}|^{1/2}\mathcal{B}_1](r,s,0) + F(r,s) + E_2\hat{t}^2 + E_3\hat{t}^3,\\
\tilde{A}_3^{[3]}(r,s,t) &= 0,
\end{aligned} \tag{2-10}$$

where $\hat{t} = t\chi(h^{-1/2+\eta}t)$ for some smooth cutoff function $\chi$ equal to 1 near 0 and where

$$F(r,s) = \int_0^r [|\boldsymbol{g}|^{1/2}\mathcal{B}_3](\ell,s,0)\,\mathrm{d}\ell, \tag{2-11}$$

and the functions $C_j(r,s)$ and $E_j(r,s)$ are smooth. We emphasize that we only truncate the terms of order at least 2 in $t$ in the above expression.

Due to Assumption 1.3, $(r,s) \mapsto (F(r,s), s)$ is a smooth diffeomorphism on a neighborhood of $(0,0)$.

We also consider the expansions

$$\begin{aligned}
|\boldsymbol{g}|^{1/2}(r,s,t) &= a_0(r,s) + ta_1(r,s) + t^2a_2(r,s) + \mathscr{O}(t^3),\\
\boldsymbol{G}^{-1} &= (M_0(r,s) + tM_1(r,s) + t^2M_2(r,s))^{-1} + \mathscr{O}(t^3),
\end{aligned}$$

and we let

$$m(r,s,t) = a_0(r,s) + \hat{t}a_1(r,s) + \hat{t}^2a_2(r,s), \quad M(r,s,t) = M_0(r,s) + \hat{t}M_1(r,s) + \hat{t}^2M_2(r,s). \tag{2-12}$$

Recall that $|\boldsymbol{g}|(r,s,0) = \alpha(r,s)$.

**2.3.2.** *Extension of the functions of the tangential variables.* It will be convenient to work on the half-space $\mathbb{R}^3_+$ instead of a neighborhood of $(0,0,0)$.

Given $\epsilon_0 > 0$, consider a smooth, odd, and nondecreasing function $\zeta : \mathbb{R} \to \mathbb{R}$ such that $\zeta(x) = x$ on $[0,\epsilon_0]$ and $\zeta(x) = 2\epsilon_0$ for all $x \geqslant 2\epsilon_0$. In particular, $\|\zeta\|_\infty = 2\epsilon_0$. We let

$$Z(r,s) = (\zeta(r), \zeta(s)).$$

The following lemma is a straightforward consequence of Assumption 1.3.

**Lemma 2.9.** *For $\epsilon_0$ small enough, the function $\hat{\beta} = \beta \circ Z : \mathbb{R}^2 \to \mathbb{R}_+$ is smooth and has a unique minimum (at $(0, 0)$), which is nondegenerate and not attained at infinity.*

Let us now replace the function $\mathscr{B} : (r, s) \mapsto \alpha(r, s)^{1/2}\mathcal{B}(r, s, 0)$ by $\mathscr{B} \circ Z$ in (2-10) and (2-11). We replace the other coefficients $C_j$ and $E_j$ by $C_j \circ Z$ and $E_j \circ Z$. Note that we have the following.

**Lemma 2.10.** *For $\epsilon_0$ small enough, the function*

$$\mathscr{J} : \mathbb{R}^2 \ni (r, s) \mapsto \left( \int_0^r [|\boldsymbol{g}|^{1/2}\mathcal{B}_3](Z(\ell, s), 0)\, \mathrm{d}\ell, s \right) = u = (u_1, u_2) \in \mathbb{R}^2$$

*is smooth, and it is a global diffeomorphism.*

This leads us to consider the new vector potential

$$\begin{aligned} \hat{A}_1(r, s, t) &= t\mathring{C}_1 + \mathring{C}_2\hat{t}^2 + \mathring{C}_3\hat{t}^3, \\ \hat{A}_2(r, s, t) &= -t\mathring{E}_1 + \mathscr{J}_1(r, s) + \mathring{E}_2\hat{t}^2 + \mathring{E}_3\hat{t}^3, \\ \hat{A}_3(r, s, t) &= 0, \end{aligned} \tag{2-13}$$

where $C_1 = \alpha^{1/2}\mathcal{B}_2,\ E_1 = \alpha^{1/2}\mathcal{B}_1$ and with the notation $\mathring{f} = f \circ Z$.

The rest of the article will be devoted to the spectral analysis of the operator associated with the new quadratic form

$$\mathcal{Q}_h^{\mathrm{app}}(\varphi) = \int_{\mathbb{R}^3_+} \langle (\mathring{M})^{-1}(-ih\nabla_y - \hat{\boldsymbol{A}}(y))\varphi, (-ih\nabla_y - \hat{\boldsymbol{A}}(y))\varphi \rangle \mathring{m}\, \mathrm{d}y.$$

This self-adjoint operator $\mathscr{L}_h^{\mathrm{app}}$ is acting as

$$\mathring{m}^{-1}(-ih\nabla_y - \hat{\boldsymbol{A}}) \cdot \mathring{m}(\mathring{M})^{-1}(-ih\nabla_y - \hat{\boldsymbol{A}})$$

in the ambient Hilbert space $L^2(\mathbb{R}^3_+, \mathring{m}\, \mathrm{d}y)$. We recall that $m$ and $M$ are given in (2-12). This spectral analysis is motivated by the fact that the low-lying spectra of $\mathscr{L}_h$ and $\mathscr{L}_h^{\mathrm{app}}$ coincide modulo $o(h^2)$ in the sense of the following proposition.

**Proposition 2.11.** *We have, for all $n \geqslant 1$,*

$$\lambda_n(h) = \lambda_n(\mathscr{L}_h^{\mathrm{app}}) + o(h^2).$$

We omit the proof. It follows from Corollary 2.2, the localization estimates given in Proposition 2.3 (which are also true in the coordinates $(r, s, t)$ for the eigenfunctions of $\mathscr{L}_h^{\mathrm{app}}$ by using the same arguments), and the min-max theorem. These localization estimates allow us to remove the cutoff functions up to remainders of order $\mathscr{O}(h^\infty)$ and to control the remainders of the expansion in $t$.

## 3. Change of coordinates and metaplectic transform

In order to perform the spectral analysis of $\mathscr{L}_h^{\mathrm{app}}$, it is convenient to use the change of variable $\mathscr{J}$ given in Lemma 2.10. More precisely, we will use the unitary transform induced by $\mathscr{J}$ defined by

$$U : L^2(\mathbb{R}^3_+, \mathring{m}\, \mathrm{d}y) \to L^2(\mathbb{R}^3_+, \breve{m}|\mathrm{Jac}\,\mathscr{J}^{-1}|\, \mathrm{d}u\, \mathrm{d}t), \quad \varphi \mapsto \breve{\varphi},$$

where we use the notation $\breve{f}(u,t)=f(\mathscr{J}^{-1}(u),t)$ and the slight abuse of notation $\breve{\mathring{f}}=\breve{f}$. Then, we focus on the operator $\mathscr{N}_h=U\mathscr{L}_h^{\mathrm{app}}U^{-1}$ acting in $L^2(\mathbb{R}^3_+,\breve{m}|\mathrm{Jac}\,\mathscr{J}^{-1}|\,\mathrm{d}u\,\mathrm{d}t)$. The operator $\mathscr{N}_h$ is acting as

$$\mathscr{N}_h=U\mathscr{L}_h^{\mathrm{app}}U^{-1}=\breve{m}^{-1}\mathscr{D}_h\cdot\breve{m}(\breve{M})^{-1}\mathscr{D}_h, \tag{3-1}$$

where

$$\mathscr{D}_h=\begin{pmatrix} -ih\breve{C}_0\,\partial_{u_1}-t\breve{C}_1-\hat{t}^2\breve{C}_2-\hat{t}^3\breve{C}_3\\ -ih\,\partial_{u_2}-u_1-ih\breve{E}_0\,\partial_{u_1}+t\breve{E}_1-\hat{t}^2\breve{E}_2-\hat{t}^3\breve{E}_3\\ -ih\,\partial_t\end{pmatrix}$$

and

$$C_0=\partial_r\mathscr{J}_1=\alpha^{1/2}\mathcal{B}_3,\qquad E_0=\partial_s\mathscr{J}_1. \tag{3-2}$$

**Notation 3.1.** We will use the following classical notation for the semiclassical Weyl quantization of a symbol $a=a(u,\upsilon)$. We let

$$a^W\psi(u)=\frac{1}{(2\pi h)^2}\int_{\mathbb{R}^4}e^{i(u-x)\cdot\upsilon/h}a\Big(\frac{u+x}{2},\upsilon\Big)\psi(x)\,\mathrm{d}x\,\mathrm{d}\upsilon.$$

**Proposition 3.2.** *Let $K>0$ and $\eta\in\big(0,\frac12\big)$. Let $\Xi$ be a smooth function of the real variable equal to* 0 *near* 0 *and* 1 *away from a compact neighborhood of* 0*. There exists $h_0>0$ such that, for all $h\in(0,h_0)$ and for all normalized eigenfunctions $\psi$ of $\mathscr{N}_h$ associated with an eigenvalue $\lambda$ such that $\lambda\leqslant Kh$, we have, in $L^2(\mathbb{R}^3_+)$,*

$$\Big[\Xi\Big(\frac{u_1-\upsilon_2}{h^{1/2-\eta}}\Big)\Big]^W\psi=\mathscr{O}(h^\infty).$$

*Proof.* To simplify the notation, we write

$$\Xi_h=\Xi\Big(\frac{u_1-\upsilon_2}{h^{1/2-\eta}}\Big).$$

Note that $\Xi_h^W$ is a bounded operator by virtue of the Calderón–Vaillancourt theorem (see [Zworski 2012, Theorem 4.23]).

Let $\psi$ be a normalized eigenfunction of $\mathscr{N}_h$ associated with an eigenvalue $\lambda$ such that $\lambda\leqslant Kh$. The eigenvalue equation gives us

$$\langle\mathscr{N}_h\Xi_h^W\psi,\Xi_h^W\psi\rangle=\lambda\|\Xi_h^W\psi\|^2+\langle[\mathscr{N}_h,\Xi_h^W]\psi,\Xi_h^W\psi\rangle, \tag{3-3}$$

where $\langle\cdot,\cdot\rangle$ is the scalar product in $L^2(\mathbb{R}^3_+,\breve{m}|\mathrm{Jac}\,\mathscr{J}^{-1}|\,\mathrm{d}u\,\mathrm{d}t)$.

According to the localization at the scale $h^{1/2}$ with respect to $t$, we can insert a cutoff function supported in $\{t\leqslant h^{(1-\eta)/2}\}$, and we obtain, for $j=2,3$,

$$\|t^j\,\Xi_h^W\psi\|\leqslant Ch^{1-\eta}\|\Xi_h^W\psi\|+\mathscr{O}(h^\infty)\|\psi\|. \tag{3-4}$$

Then, we write

$$\begin{aligned}\langle\mathscr{N}_h\Xi_h^W\psi,\Xi_h^W\psi\rangle_{L^2(\mathbb{R}^3_+,\breve{m}|\mathrm{Jac}\,\mathscr{J}^{-1}|\,\mathrm{d}u\,\mathrm{d}t)}&=\langle\mathscr{D}_h\cdot\breve{m}(\breve{M})^{-1}\mathscr{D}_h\Xi_h^W\psi,|\mathrm{Jac}\,\mathscr{J}^{-1}|\Xi_h^W\psi\rangle_{L^2(\mathbb{R}^3_+,\,\mathrm{d}u\,\mathrm{d}t)}\\&=\langle\breve{m}(\breve{M})^{-1}\mathscr{D}_h\Xi_h^W\psi,\mathscr{D}_h(|\mathrm{Jac}\,\mathscr{J}^{-1}|\Xi_h^W\psi)\rangle_{L^2(\mathbb{R}^3_+,\,\mathrm{d}u\,\mathrm{d}t)}.\end{aligned}$$

We notice that $[\mathscr{D}_h, |\mathrm{Jac}\,\mathscr{J}^{-1}|] = \mathscr{O}(h)$ and that $\check{m}(\check{M})^{-1} \geqslant c > 0$. This implies that

$$\begin{aligned}\langle \mathscr{N}_h \Xi_h^W \psi, \Xi_h^W \psi\rangle_{L^2(\mathbb{R}^3_+, \check{m}|\mathrm{Jac}\,\mathscr{J}^{-1}|\,\mathrm{d}u\,\mathrm{d}t)} &\geqslant c\|\mathscr{D}_h \Xi_h^W \psi\|^2 - Ch\|\mathscr{D}_h \Xi_h^W \psi\|\|\Xi_h^W \psi\| \\ &\geqslant \tfrac{1}{2}c\|\mathscr{D}_h \Xi_h^W \psi\|^2 - Ch^2\|\Xi_h^W \psi\|^2,\end{aligned}$$

where we use the Young inequality to get the last estimate.

By using again the Young inequality and (3-4) to deal with the powers $\hat{t}^2$ and $\hat{t}^3$ in $\mathscr{D}_h$, this yields, for some $c, C > 0$,

$$\langle \mathscr{N}_h \Xi_h^W \psi, \Xi_h^W \psi\rangle_{L^2(\mathbb{R}^3_+, \check{m}|\mathrm{Jac}\,\mathscr{J}^{-1}|\,\mathrm{d}u\,\mathrm{d}t)} \geqslant c Q_h^0(\Xi_h^W \psi) - Ch^{1-\eta}\|\Xi_h^W \psi\|^2 + \mathscr{O}(h^\infty)\|\psi\|^2, \tag{3-5}$$

where

$$Q_h^0(\varphi) = \|h\,\partial_t \varphi\|^2 + \|(h\check{C}_0 D_{u_1} - t\check{C}_1)\varphi\|^2 + \|(hD_{u_2} - u_1 + h\check{E}_0 D_{u_1} + t\check{E}_1)\varphi\|^2.$$

Then, using again the Young inequality, we find that

$$Q_h^0(\varphi) \geqslant \|h\,\partial_t\varphi\|^2 + \tfrac{1}{2}\|h\check{C}_0 D_{u_1}\varphi\|^2 + \tfrac{1}{2}\|(hD_{u_2} - u_1)\varphi\|^2 - 2\|h\check{E}_0 D_{u_1}\varphi\|^2 - C\|t\varphi\|^2.$$

Notice that there exists $c > 0$ such that

$$|\check{C}_0| \geqslant c, \quad |\check{E}_0| \leqslant \tfrac{1}{4}c, \tag{3-6}$$

where we recall (3-2) and Lemma 2.10. Indeed, we have $E_0(0,0) = 0$ and, for some $c_0 > 0$, we have $C_0 \geqslant c_0 > 0$. In particular, $\epsilon_0$ can be chosen small enough in the extension procedure in Section 2.3.2 that (3-6) holds. This shows that, for some $c_0 > 0$,

$$Q_h^0(\varphi) \geqslant \|h\,\partial_t\varphi\|^2 + c_0\|hD_{u_1}\varphi\|^2 + \tfrac{1}{2}\|(hD_{u_2} - u_1)\varphi\|^2 - C\|t\varphi\|^2. \tag{3-7}$$

On the support of $\Xi_h$, we have $(\upsilon_2 - u_1)^2 \geqslant ch^{1-2\eta}$ for some $c > 0$. Thus (3-4), (3-5), (3-7), and again the localization in $t$ yield

$$\langle \mathscr{N}_h \Xi_h^W \psi, \Xi_h^W \psi\rangle_{L^2(\mathbb{R}^3_+, \check{m}|\mathrm{Jac}\,\mathscr{J}^{-1}|\,\mathrm{d}u\,\mathrm{d}t)} \geqslant \tfrac{1}{2}\tilde{c}h^{1-2\eta}\|\Xi_h^W \psi\|^2 + \mathscr{O}(h^\infty)\|\psi\|^2. \tag{3-8}$$

Using classical results of composition of pseudo-differential operators, we have

$$\langle [\mathscr{N}_h, \Xi_h^W]\psi, \Xi_h^W \psi\rangle \leqslant Ch^{1+\eta}\|\underline{\Xi}_h^W \psi\|^2 + \mathscr{O}(h^\infty)\|\psi\|^2, \tag{3-9}$$

where $\underline{\Xi}$ has a support slightly larger than that of $\Xi_h$. Here we use the energy estimate

$$\|\mathscr{D}_h \Xi_h^W \psi\| = \mathscr{O}(h^{1/2})\|\underline{\Xi}_h^W \psi\| + \mathscr{O}(h^\infty)\|\psi\|,$$

which follows from rough estimates of (3-3).

Thus, by combining (3-3), (3-8), and (3-9) with the fact that $\lambda \leqslant Kh$, we obtain

$$\|\Xi_h^W \psi\|^2 \leqslant Mh^\eta\|\underline{\Xi}_h^W \psi\|^2 + \mathscr{O}(h^\infty)\|\psi\|^2.$$

Finally, by an induction argument on the size of the support of $\Xi$, we get

$$\|\Xi_h^W \psi\| = \mathscr{O}(h^\infty)\|\psi\|. \qquad \square$$

Let us consider the partial semiclassical Fourier transform $\mathscr{F}_2$ with respect to $u_2$ and the translation/dilation $T : u_1 \mapsto (u_1 - \upsilon_2)h^{-1/2} = z$. Slightly abusing notation, we identify $T$ with $\varphi \mapsto \varphi \circ T$. We mention that $\mathscr{F}_2$ is the metaplectic transform associated with the linear symplectic application $(u_2, \upsilon_2) \mapsto (\upsilon_2, -u_2)$; see, for instance, [Martinez 2002, Section 3.4]. Letting $V = \mathscr{F}_2^{-1}T$, we have

$$V^*(-ih\,\partial_{u_2} - u_1)V = -h^{1/2}z.$$

For the following it is pertinent to introduce the new semiclassical parameter

$$\hbar = h^{1/2},$$

keeping in mind that we continue to deal not only with $h$-pseudodifferential operators but also with asymptotic expansions in $\hbar$. The preceding equality then becomes

$$V^*(-ih\,\partial_{u_2} - u_1)V = -\hbar z.$$

Similarly, with the dilation $W : t \mapsto h^{-1/2}t = \hbar^{-1}t$, we get

$$W^*V^*\mathscr{D}_h VW = \hbar\mathscr{D}_\hbar^\sharp,$$

with

$$\mathscr{D}_\hbar^\sharp = \begin{pmatrix} -iC_0^\sharp\,\partial_z - tC_1^\sharp - \hbar t^2\chi(\hbar^{2\eta}t)^2C_2^\sharp - \hbar^2t^3\chi(\hbar^{2\eta}t)^3C_3^\sharp \\ -z - iE_0^\sharp\,\partial_z + tE_1^\sharp - \hbar t^2\chi(\hbar^{2\eta}t)^2E_2^\sharp - \hbar^2t^3\chi(\hbar^{2\eta}t)^3E_3^\sharp \\ -i\,\partial_t \end{pmatrix}^W,$$

where the coefficients of the conjugated operator $\mathscr{D}_\hbar^\sharp$ are now given by $P^\sharp = \breve{P}(\upsilon_2 + \hbar z, -u_2)$. Here the Weyl quantization can be considered only in the variables $(u_2, \upsilon_2)$ since $z$ is now a "space variable". We let

$$\mathscr{N}_\hbar^\sharp = [m_\hbar{}^{-1}]^\sharp\mathscr{D}_\hbar^\sharp \cdot [m_\hbar(M_\hbar)^{-1}]^\sharp\mathscr{D}_\hbar^\sharp,$$

where $m_\hbar(\,\cdot\,, t) = m(\,\cdot\,, \hbar t)$ and $M_\hbar(\,\cdot\,, t) = M(\,\cdot\,, \hbar t)$. The operator $\mathscr{N}_\hbar^\sharp$ is equipped with the domain $(VW)^{-1}\,\mathrm{Dom}\,\mathscr{N}_h$ (which is still made of functions satisfying the Neumann boundary condition). Note that $\mathscr{N}_h$ and $\hbar^2\mathscr{N}_\hbar^\sharp$ are unitarily equivalent since

$$W^*V^*\mathscr{N}_h VW = \hbar^2\mathscr{N}_\hbar^\sharp. \tag{3-10}$$

After all these elementary transforms, Proposition 3.2 can be reformulated as follows.

**Proposition 3.3.** *Let $K > 0$ and $\eta \in \left(0, \frac{1}{2}\right)$. Let $\Xi$ be a smooth function of the real variable equal to* 0 *near* 0 *and* 1 *away from a compact neighborhood of* 0. *There exists $\hbar_0 > 0$ such that, for all $\hbar \in (0, \hbar_0)$ and for all normalized eigenfunctions $\psi$ of $\mathscr{N}_\hbar^\sharp$ associated with an eigenvalue $\lambda$ such that $\lambda \leqslant K$, we have*

$$\Xi(\hbar^{2\eta}z)\psi = \mathscr{O}(\hbar^\infty).$$

**Remark 3.4.** As a consequence of the Agmon estimates and working in the coordinates $(u_1, u_2, t)$, we notice that the eigenfunctions are also roughly localized in "frequency" in the sense that, for all $(\alpha, \beta, \gamma) \in \mathbb{N}^3$ and all $\eta \in \left(0, \frac{1}{2}\right)$, there exist $C, \hbar_0 > 0$ such that, for all $\hbar \in (0, \hbar_0)$,

$$\|t^\alpha z^\beta D_z^\gamma\psi\| + \|t^\alpha z^\beta D_t^\gamma\psi\| \leqslant C\hbar^{-2\eta(\alpha+\beta+\gamma)}\|\psi\|.$$

## 4. A pseudodifferential operator with operator symbol

Proposition 3.3 invites us to insert cutoff functions in the coefficients of the operator $\mathscr{N}_\hbar^\sharp$. Working from now on with the semiclassical parameter $\hbar$, we therefore consider

$$\mathscr{N}_\hbar^\flat = ([m_\hbar{}^{-1}]^\flat)^W \mathscr{D}_\hbar^\flat \cdot ([m_\hbar(M_\hbar)^{-1}]^\flat)^W \mathscr{D}_\hbar^\flat, \tag{4-1}$$

where

$$\mathscr{D}_\hbar^\flat = \begin{pmatrix} -iC_0^\flat\,\partial_z - tC_1^\flat - \hbar t^2\chi(\hbar^{2\eta}t)^2 C_2^\flat - \hbar^2 t^3\chi(\hbar^{2\eta}t)^3 C_3^\flat \\ -z - iE_0^\flat\,\partial_z + tE_1^\flat - \hbar t^2\chi(\hbar^{2\eta}t)^2 E_2^\flat - \hbar^2 t^3\chi(\hbar^{2\eta}t)^3 E_3^\flat \\ -i\,\partial_t \end{pmatrix}^W, \tag{4-2}$$

with $P^\flat = \check{P}(\upsilon_2 + \hbar\chi_\eta(z)z, -u_2)$, where $\chi_\eta(z) = \chi_0(\hbar^{2\eta}z)$, the function $\chi_0$ being smooth, with a compact support, and equal to 1 on a neighborhood of the support of $1 - \Xi$.

### 4.1. *The symbol and its properties.*

Expanding the operator $\mathscr{N}_\hbar^\flat$ with respect to $\hbar$ (say first at a formal level) suggests that we should consider the following self-adjoint operator, depending on the parameters $(u_2, \upsilon_2)$ and acting in the variables $(z, t)$ as

$$\begin{aligned} &n_0(u_2, \upsilon_2) \\ &= (-i\check{C}_0(\upsilon_2, -u_2)\,\partial_z - t\check{C}_1(\upsilon_2, -u_2))^2 + \alpha^{-1}(\upsilon_2, -u_2)(-z - i\check{E}_0(\upsilon_2, -u_2)\,\partial_z + t\check{E}_1(\upsilon_2, -u_2))^2 - \partial_t^2, \end{aligned}$$

with the domain

$$\mathrm{Dom}(n_0) = \{\psi \in L^2(\mathbb{R}^2_+) : n_0(u_2, \upsilon_2)\psi \in L^2(\mathbb{R}^2_+),\ \partial_t\psi(z, 0) = 0\},$$

and where we recall that $C_1$ and $E_1$ are given in (2-13). The domain of $n_0(u_2, \upsilon_2)$ depends on $(u_2, \upsilon_2)$. However, we can check that it is unitarily equivalent to a self-adjoint operator with domain independent of $(u_2, \upsilon_2)$, see the proof of Proposition 4.4 below. In the following, we will use a class of operator symbols of the form

$$S(\mathbb{R}^2, \mathcal{L}(\mathscr{A}_1, \mathscr{A}_2)) = \{a \in \mathscr{C}^\infty(\mathbb{R}^2, \mathcal{L}(\mathscr{A}_1, \mathscr{A}_2)) : \forall\gamma \in \mathbb{N}^2, \exists C_\gamma > 0 : \|\partial^\gamma a\|_{\mathcal{L}(\mathscr{A}_1,\mathscr{A}_2)} \leqslant C_\gamma\},$$

where $\mathscr{A}_1$ and $\mathscr{A}_2$ are (fixed) Hilbert spaces. We also introduce

$$\mathscr{B}_k = \{\psi \in L^2(\mathbb{R}^2_+) : \forall\alpha \in \mathbb{N}^2, |\alpha| \leqslant k \Rightarrow (\langle t\rangle^k + \langle z\rangle^k)\,\partial^\alpha\psi \in L^2(\mathbb{R}^2_+)\} \tag{4-3}$$

and the class of symbols

$$S(\mathbb{R}^2, N) = \bigcap_{k \geqslant N} S(\mathbb{R}^2, \mathcal{L}(\mathscr{B}_k, \mathscr{B}_{k-N})),$$

and we notice that $n_0 \in S(\mathbb{R}^2, 2)$.

**Remark 4.1.** Note that these classes of symbols are not algebras. However, the classical Moyal product of symbols in $S(\mathbb{R}^2, N)$ and $S(\mathbb{R}^2, M)$ is well-defined and belongs to $S(\mathbb{R}^2, N + M)$; see [Keraval 2018, Theorem 2.1.12].

In fact, for $N \geqslant 2$, by using a classical trace theorem, we may also define

$$\mathscr{B}_N^{\mathrm{Neu}} = \{\psi \in \mathscr{B}_N : \partial_t \psi(z, 0) = 0\} \quad (\subset \operatorname{Dom} n_0)$$

and the associated class $S^{\mathrm{Neu}}(\mathbb{R}^2, N)$. We can also write $n_0 \in S^{\mathrm{Neu}}(\mathbb{R}^2, 2)$ to remember that the domain of $n_0$ is equipped with the Neumann condition.

By expanding $\mathscr{N}_\hbar^\flat$ in powers of $\hbar$ and by using the composition theorem for pseudodifferential operators [Keraval 2018, Theorem 2.1.12], we get the following.

**Proposition 4.2.** *The operator $\mathscr{N}_\hbar^\flat$ is an h-pseudodifferential operator with symbol in the class $S^{\mathrm{Neu}}(\mathbb{R}^2, 2)$. Moreover, we can write the expansion*

$$\mathscr{N}_\hbar^\flat = n_0^W + \hbar n_1^W + \hbar^2 n_2^W + \hbar^3 r_\hbar^W, \tag{4-4}$$

*with $n_1$, $n_2$, and $r_\hbar$ in the class $S^{\mathrm{Neu}}(\mathbb{R}^2, 8)$.*

*Proof.* Let us recall that $\mathscr{N}_\hbar^\flat$ is given in (4-1). Let us notice that the operator $\mathscr{D}_\hbar^\flat$, defined in (4-2), is indeed a pseudodifferential operator with operator-valued symbol. With respect to the variables $z$ and $t$, it is a differential operator of order 1 whose symbol is

$$\begin{pmatrix} -iC_0^\flat \partial_z - tC_1^\flat - \hbar t^2 \chi(\hbar^{2\eta} t)^2 C_2^\flat - \hbar^2 t^3 \chi(\hbar^{2\eta} t)^3 C_3^\flat \\ -z - iE_0^\flat \partial_z + tE_1^\flat - \hbar t^2 \chi(\hbar^{2\eta} t)^2 E_2^\flat - \hbar^2 t^3 \chi(\hbar^{2\eta} t)^3 E_3^\flat \\ -i\,\partial_t \end{pmatrix} \tag{4-5}$$

and belongs to $S(\mathbb{R}^2, 1)$. The functions/symbols $[m_\hbar{}^{-1}]^\flat$ and $[m_\hbar(M_\hbar)^{-1}]^\flat$ belong to $S(\mathbb{R}^2, 0)$. Combining these considerations with (4-1), it remains to apply the composition theorem for pseudodifferential operators with operator symbols, see Remark 4.1.

To get (4-4), it is sufficient to use the Taylor expansions in $\hbar$ of the symbol (4-5), $[m_\hbar{}^{-1}]^\flat$, and $[m_\hbar(M_\hbar)^{-1}]^\flat$, and to apply again the composition theorem (the worst remainders being roughly of order 8 in $(z, t)$). ☐

**Remark 4.3.** We will see that the accurate descriptions of $n_1$ and $n_2$ in (4-4) are not necessary to prove our main theorem. The use of the more restrictive class $S^{\mathrm{Neu}}(\mathbb{R}^2, 8)$ allows us to deal with the uniformity in the semiclassical expansions in $\hbar$.

Let us describe the groundstate energy of the principal symbol $n_0$. From now on, we lighten the notation by setting $(u_2, \upsilon_2) = (u, \upsilon)$.

**Proposition 4.4.** *For all $(u, \upsilon) \in \mathbb{R}^2$, the bottom of the spectrum of $n_0$ belongs to the discrete spectrum and it is a simple eigenvalue that equals $\breve{\beta}(\upsilon, -u)$. The corresponding normalized eigenfunction $\mathfrak{f}_{u,\upsilon}$ belongs to the Schwartz class and depends on $(u, \upsilon)$ smoothly.*

*Moreover, there exists $c > 0$ such that, by possibly choosing $\epsilon_0$ smaller in Lemma 2.9, we have, for all $(u, \upsilon) \in \mathbb{R}^2$,*

$$\inf \operatorname{sp}(n_0(u, \upsilon)|_{\mathfrak{f}_{u,\upsilon}^\perp}) \geqslant \beta_{\min} + c \geqslant \breve{\beta}(\upsilon, -u).$$

*Proof.* By using the Fourier transform in $z$ and then a change of gauge, we are reduced to the case when $E_0 = 0$. With a rescaling in $z$, $n_0$ is unitarily equivalent to

$$(-i\,\partial_z - t\check{C}_1)^2 + \alpha^{-1}(-\check{C}_0 z + t\check{E}_1)^2 - \partial_t^2 = (-i\,\partial_z - tb_2)^2 + (b_3 z + tb_1)^2 - \partial_t^2,$$

with

$$b_1 = \check{\mathcal{B}}_1, \quad b_2 = (\alpha^{1/2}\mathcal{B}_2)\check{}, \quad b_3 = -\check{\mathcal{B}}_3,$$

where the functions are evaluated at $(\upsilon_2, -u_2)$. Recalling (2-8), we see that the Euclidean norm of $b = (b_1, b_2, b_3)$ is

$$\|b\|_2 = \|\check{\boldsymbol{B}}\|,$$

with a slight abuse of notation. By homogeneity, we can easily scale out $\|\check{\boldsymbol{B}}\|$ and consider the operator

$$(-i\,\partial_z - tb_2)^2 - \partial_t^2 + (tb_1 + b_3 z)^2,$$

with

$$b_1 = \cos\theta\cos\varphi, \quad b_2 = \cos\theta\sin\varphi, \quad b_3 = \sin\theta.$$

Completing a square leads to the identity

$$(-i\,\partial_z - tb_2)^2 - \partial_t^2 + (tb_1 + b_3 z)^2 = -\partial_t^2 + (t\cos\theta - \sin\varphi D_z - z\sin\theta\cos\varphi)^2 + (\cos\varphi D_z - z\sin\theta\sin\varphi)^2.$$

This shows, thanks to the rescaling $z = \tilde{z}\sin\varphi$ (since $\sin\varphi$ is nonzero) and the change of gauge

$$\psi \mapsto e^{-i\frac{\tilde{z}^2}{2}\sin\theta\cos\varphi}\psi,$$

that the operator is unitarily equivalent to

$$D_t^2 + (t\cos\theta - D_{\tilde{z}})^2 + (\cot\varphi D_{\tilde{z}} - \tilde{z}\sin\theta)^2$$

and then, by the Fourier transform, to

$$D_t^2 + (t\cos\theta - \zeta)^2 + (\zeta\cot\varphi + \sin\theta D_\zeta)^2.$$

Thanks to the change of gauge

$$\psi \mapsto e^{-i\frac{\zeta^2\cot\varphi}{2\sin\theta}}\psi$$

(which is well-defined since $\sin\theta \neq 0$), this last operator is unitarily equivalent to

$$D_t^2 + D_z^2 + (t\cos\theta - z\sin\theta)^2,$$

which is nothing but the Lu–Pan operator defined in (1-2), which is unitarily equivalent to

$$\cos^2\theta D_t^2 + \sin^2\theta D_z^2 + (t - z)^2$$

(whose domain is independent of $\theta$).

The eigenfunction $\mathfrak{f}_{u,\upsilon}$ belongs to the Schwartz class by virtue of [Raymond 2009, Corollaire 5.1.2] and the stability of the Schwartz class under Fourier and gauge transforms. □

**4.2.** ***An approximate parametrix.***

**4.2.1.** *Inverting the principal symbol.*

**Lemma 4.5.** *Consider $\epsilon > 0$ and $\Lambda \leqslant \beta_{\min} + \epsilon$. We let*

$$\mathscr{P}_0(\Lambda) = \begin{pmatrix} n_0(u, \upsilon) - \Lambda & \cdot\, \mathfrak{f}_{u,\upsilon} \\ \langle\, \cdot\,, \mathfrak{f}_{u,\upsilon}\rangle & 0 \end{pmatrix}.$$

*For $\epsilon$ small enough, $\mathscr{P}_0(\Lambda) : \operatorname{Dom} n_0 \times \mathbb{C} \to L^2(\mathbb{R}^2_+) \times \mathbb{C}$ is bijective. Its inverse is denoted by $\mathscr{Q}_0$ and is given by*

$$\mathscr{Q}_0 = \mathscr{Q}_0(\Lambda) = \begin{pmatrix} (n_0(u, \upsilon) - \Lambda)^{-1}_{\perp} & \cdot\, \mathfrak{f}_{u,\upsilon} \\ \langle\, \cdot\,, \mathfrak{f}_{u,\upsilon}\rangle & \Lambda - \breve{\beta}(\upsilon, -u) \end{pmatrix},$$

*where $(n_0(u, \upsilon) - \Lambda)^{-1}_{\perp}$ is the regularized resolvent on $(\operatorname{span} \mathfrak{f}_{u,\upsilon})^{\perp}$.*

*Moreover, we have $\mathscr{Q}_0 \in S(\mathbb{R}^2, 0)$.*

*Proof.* Using the same algebraic computations as in [Keraval 2018] and the spectral gap in Proposition 4.4, we get the announced inverse. Moreover, it is also clear that $\mathscr{Q}_0$ is bounded from $L^2(\mathbb{R}^2_+)$ to $L^2(\mathbb{R}^2_+)$ uniformly in $(u, \upsilon)$. The fact that it belongs to the class $S(\mathbb{R}^2, 0)$ follows from weighted resolvent estimates similar to [Raymond 2009, pp. 100-101]; see also [Fahs et al. 2024, Appendix]. □

We let

$$\mathscr{P}_\hbar(\Lambda) = \begin{pmatrix} n_0 + \hbar n_1 + \hbar^2 n_2 + \hbar^3 r_\hbar - \Lambda & \cdot\, \mathfrak{f}_{u,\upsilon} \\ \langle\, \cdot\,, \mathfrak{f}_{u,\upsilon}\rangle & 0 \end{pmatrix} = \mathscr{P}_0(\Lambda) + \hbar \mathscr{P}_1 + \hbar^2 \mathscr{P}_2 + \hbar^3 \mathscr{R}_\hbar,$$

where $n_0$, $n_1$, $n_2$, and $r_\hbar$ are given in Proposition 4.2.

**4.2.2.** *The approximate parametrix.* Let us now construct an approximate (at order 2) inverse of $\mathscr{P}_\hbar^W$ when it acts on the Schwartz class (with Neumann condition). We consider

$$\mathscr{Q}_\hbar = \mathscr{Q}_0 + \hbar \mathscr{Q}_1 + \hbar^2 \mathscr{Q}_2 = \begin{pmatrix} Q_\hbar & Q_\hbar^+ \\ Q_\hbar^- & Q_\hbar^\pm \end{pmatrix},$$

where

$$\mathscr{Q}_1 = -\mathscr{Q}_0 \mathscr{P}_1 \mathscr{Q}_0, \quad \mathscr{Q}_2 = -\mathscr{Q}_0 \mathscr{P}_2 \mathscr{Q}_0 + \mathscr{Q}_0 \mathscr{P}_1 \mathscr{Q}_0 \mathscr{P}_1 \mathscr{Q}_0 - \frac{1}{i}\{\mathscr{Q}_0, \mathscr{P}_0\}\mathscr{Q}_0. \tag{4-6}$$

By Remark 4.1, the symbols $\mathscr{Q}_1$ and $\mathscr{Q}_2$ belong to $S(\mathbb{R}^2, M)$ for some $M \geqslant 8$. By computing products of matrices and using the exponential decay of $\mathfrak{f}_{u,\upsilon}$, we get

$$Q_\hbar^\pm(\Lambda) = \Lambda - (p_0 + \hbar p_1 + \hbar^2 p_{2,\Lambda}), \tag{4-7}$$

with $p_0 = \breve{\beta}(\upsilon, -u)$ and $p_1$, $p_{2,\Lambda} \in S_{\mathbb{R}^2}(1)$, where

$$S_{\mathbb{R}^2}(1) = \{a \in \mathscr{C}^\infty(\mathbb{R}^2, \mathbb{C}) : \forall \alpha \in \mathbb{N}^2, \exists C_\alpha > 0 : |\partial^\alpha a| \leqslant C_\alpha\}.$$

In addition, $\Lambda \mapsto p_{2,\Lambda} \in S_{\mathbb{R}^2}(1)$ is analytic in a neighborhood of $\beta_{\min}$.

**Remark 4.6.** Let us emphasize here that nothing a priori ensures that the subprincipal symbols $p_1$ and $p_{2,E}$ are real-valued since our formal operator is not self-adjoint on the canonical $L^2$-space.

The reason to consider the expressions (4-6) simply comes from the semiclassical expansion of the product $\mathscr{Q}_\hbar^W \mathscr{P}_\hbar^W$ by means of the composition theorem [Keraval 2018, Theorem 2.1.12]. These explicit choices, with the Calderón–Vaillancourt theorem [Keraval 2018, Theorem 2.1.16] to estimate the remainders, imply the following proposition.

**Proposition 4.7.** *There exists $N \geqslant 2$ such that the following holds. We have*

$$\mathscr{Q}_\hbar^W \mathscr{P}_\hbar^W = \mathrm{Id}_{\mathscr{S}^{\mathrm{Neu}}(\overline{\mathbb{R}}^3_+)\times\mathscr{S}(\mathbb{R})} + \hbar^3 \mathscr{R}_{\hbar,\ell}^W, \quad \mathscr{P}_\hbar^W \mathscr{Q}_\hbar^W = \mathrm{Id}_{\mathscr{S}(\overline{\mathbb{R}}^3_+)\times\mathscr{S}(\mathbb{R})} + \hbar^3 \mathscr{R}_{\hbar,r}^W,$$

*where $\mathscr{R}_{\hbar,\ell}$ and $\mathscr{R}_{\hbar,r}$ belong to $S(\mathbb{R}^2, N)$ and where $\mathscr{S}^{\mathrm{Neu}}(\overline{\mathbb{R}}^3_+)$ denotes the Schwartz class on $\mathbb{R}^2_+$ with Neumann condition at $t=0$.*

*In particular, we have, for all $\psi \in \mathscr{S}^{\mathrm{Neu}}(\overline{\mathbb{R}}^3_+)$,*

$$\begin{aligned} Q_\hbar^W(\mathscr{N}_\hbar^\flat - \Lambda)\psi + (Q_\hbar^+)^W \mathfrak{P}\psi &= \psi + \mathscr{O}(\hbar^3)\|\psi\|_{L^2(\mathbb{R},\mathscr{B}_N)}, \\ (Q_\hbar^-)^W(\mathscr{N}_\hbar^\flat - \Lambda)\psi + (Q_\hbar^\pm)^W \mathfrak{P}\psi &= \mathscr{O}(\hbar^3)\|\psi\|_{L^2(\mathbb{R},\mathscr{B}_N)} \end{aligned} \tag{4-8}$$

*and, for all $\varphi \in \mathscr{S}(\mathbb{R})$,*

$$\begin{aligned} (\mathscr{N}_\hbar^\flat - \Lambda)(Q_\hbar^+)^W\varphi + \mathfrak{P}^*(Q_\hbar^\pm)^W\varphi &= \mathscr{O}(\hbar^3)\|\varphi\|, \\ \mathfrak{P}(Q_\hbar^+)^W\varphi &= \varphi + \mathscr{O}(\hbar^3)\|\varphi\|. \end{aligned} \tag{4-9}$$

*Here, $\mathfrak{P} = (\langle\,\cdot\,, \mathfrak{f}_{u,\upsilon}\rangle)^W$, $\mathscr{B}_N$ is given in* (4-3), *and $\|\cdot\|_{L^2(\mathbb{R},\mathscr{B}_N)}$ is the $L^2$-norm defined thanks to the Bochner integral valued in the Banach space $\mathscr{B}_N$.*

## 5. Spectral consequences

This last section is devoted to the proof of Theorem 1.4, with the help of Proposition 4.7. The spectrum of $\mathscr{N}_\hbar^\sharp$ will be compared to the spectrum of a model operator, derived from an effective $h$-pseudodifferential operator whose symbol has the following expansion in powers of $\hbar = h^{1/2}$:

$$p_\hbar^{\mathrm{eff}} = p_0 + \hbar p_1 + \hbar^2 p_{2,\beta_{\min}}, \tag{5-1}$$

see (4-7).

**5.1.** ***A model operator.*** Let us consider

$$p_\hbar^{\mathrm{mod}}(U) = p_\hbar^{\mathrm{eff}}(0) + \tfrac{1}{2}\,\mathrm{Hess}_{(0,0)}\, p_0(U,U) + \hbar p_1^{\mathrm{lin}}(U), \quad U = (u,\upsilon),$$

where $p_1^{\mathrm{lin}}$ is the linear approximation of $p_1$ at $(0,0)$. The corresponding $h$-pseudodifferential operator $(p_\hbar^{\mathrm{mod}})^W$ is not self-adjoint due to the linear part. However, this operator still has compact resolvent, and we can compute its spectrum and estimate its resolvent. Let us explain this. Thanks to Assumption 1.3, the quadratic form $\mathrm{Hess}_{(0,0)}\, p_0(U,U)$ can be diagonalized with a rotation (which is a symplectic transformation in two dimensions). Thus, by using a metaplectic transformation (or by means of an explicit linear transformation in $u$), we may assume that the symbol is

$$p_\hbar^{\mathrm{mod}} = p_\hbar^{\mathrm{eff}}(0) + \tfrac{1}{2} d_0(u^2+\upsilon^2) + \hbar(\alpha u + \beta\upsilon)$$

for some $d_0 > 0$ and $(\alpha,\beta) \in \mathbb{C}^2$.

**Remark 5.1.** In fact, we have

$$d_0 = \sqrt{\det \operatorname{Hess}_{(0,0)} p_0} = \sqrt{\det \operatorname{Hess}_{(0,0)} \breve{\beta}(\upsilon, -u)} = \sqrt{\frac{\det \operatorname{Hess}_{x_0} \beta}{\|\boldsymbol{B}(x_0)\|^2 \sin^2 \theta(x_0)}},$$

where we use the notation introduced at the beginning of Section 3, the change of variable $\mathscr{J}$ in Lemma 2.10, and Remark 2.6.

By completing the square and recalling that we deal with $h$-quantizations and that we have let $\hbar = h^{1/2}$, we get

$$(p_\hbar^{\mathrm{mod}})^W = \tilde{p}_\hbar^{\mathrm{eff}}(0) + \frac{d_0}{2}\left(\left(u + \frac{\hbar\alpha}{d_0}\right)^2 + \left(\hbar^2 D_u + \frac{\hbar\beta}{d_0}\right)^2\right), \quad \tilde{p}_\hbar^{\mathrm{eff}}(0) = p_\hbar^{\mathrm{eff}}(0) - \frac{\alpha^2+\beta^2}{2d_0}\hbar^2.$$

For all $n \geqslant 1$, we let

$$\begin{aligned} f_n(u) &= [e^{-i\beta\cdot/d_0} H_n(\,\cdot\,)]\left(u + \frac{\alpha}{d_0}\right), \\ f_{n,\hbar}(u) &= \hbar^{-1/2} f_n(\hbar^{-1}u), \end{aligned}$$

where $H_n$ is the $n$-th normalized Hermite function.

The family $(f_{n,\hbar})_{n\geqslant 1}$ is a total family in $L^2(\mathbb{R})$ (but not necessarily orthogonal). It satisfies

$$\begin{aligned} (p_\hbar^{\mathrm{mod}})^W f_{n,\hbar} &= \lambda_n^{\mathrm{mod}}(\hbar) f_{n,\hbar}, \\ \lambda_n^{\mathrm{mod}}(\hbar) &= \tfrac{1}{2} d_0(2n-1)\hbar^2 + \tilde{p}_\hbar^{\mathrm{eff}}(0). \end{aligned} \tag{5-2}$$

By the analytic perturbation theory (see [Kato 1995, Chapter VII]), the spectrum of $(p_\hbar^{\mathrm{mod}})^W$ is made of eigenvalues of algebraic multiplicity 1, and it is given by

$$\operatorname{sp}((p_\hbar^{\mathrm{mod}})^W) = \left\{\tfrac{1}{2} d_0(2n-1)\hbar^2 + \tilde{p}_\hbar^{\mathrm{eff}}(0),\ n \geqslant 1\right\}.$$

Moreover, for all compact $K \subset \mathbb{C}$, there exists $C_K > 0$ such that, for all $\mu \in K$,

$$\|((p_\hbar^{\mathrm{mod}})^W - \tilde{p}_\hbar^{\mathrm{eff}}(0) - \hbar^2\mu)^{-1}\| \leqslant \frac{C_K}{\operatorname{dist}(\tilde{p}_\hbar^{\mathrm{eff}}(0) + \hbar^2\mu, \operatorname{sp}((p_\hbar^{\mathrm{mod}})^W))}. \tag{5-3}$$

To see this, consider the operator

$$\mathscr{A}_\hbar = (p_\hbar^{\mathrm{mod}})^W - \tilde{p}_\hbar^{\mathrm{eff}}(0) = \left(u + \frac{\hbar\alpha}{d_0}\right)^2 + \left(\hbar^2 D_u + \frac{\hbar\beta}{d_0}\right)^2.$$

When $\hbar = 1$, we have the estimate

$$\|(\mathscr{A}_1 - \mu)^{-1}\| \leqslant \frac{C_K}{\operatorname{dist}(\mu, \operatorname{sp}(\mathscr{A}_1))}, \tag{5-4}$$

which follows from the fact that the eigenvalues have algebraic multiplicity 1 (the Riesz projectors associated with the finite number of eigenvalues in $K$ have rank 1). To get (5-3), we use the rescaling $u = \hbar\tilde{u}$ and (5-4).

**5.2.** ***Refined estimates.***

**5.2.1.** *From the model operator to $\mathscr{N}_\hbar^\sharp$.* The functions $(f_{n,\hbar})$ can serve as quasimodes for $\mathscr{N}_\hbar^\sharp$ with the help of (4-9). Indeed, by taking $\Lambda = \lambda_n^{\mathrm{mod}}(\hbar)$ and $\varphi = f_{n,\hbar}$, we see that

$$(\mathscr{N}_\hbar^\flat - \lambda_n^{\mathrm{mod}}(\hbar))(Q_\hbar^+)^W f_{n,\hbar} = \mathscr{O}(\hbar^3).$$

Since $(Q_\hbar^+)^W f_{n,\hbar}$ is localized near $(z,t) = (0,0)$ (due to the exponential decay of $\mathfrak{f}_{u,\upsilon}$, which is uniform in $(u,\upsilon)$), we get

$$(\mathscr{N}_\hbar^\sharp - \lambda_n^{\mathrm{mod}}(\hbar))(Q_\hbar^+)^W f_{n,\hbar} = \mathscr{O}(\hbar^3).$$

By using the inverse Fourier transform and translation/dilation, $(Q_\hbar^+)^W f_{n,\hbar}$ becomes a quasimode for $\mathscr{N}_h$, see (3-1) and the end of Section 3. But the operator $\mathscr{N}_h$ is unitarily equivalent to a self-adjoint operator for a suitable scalar product on the usual $L^2$-space. Therefore, we can apply the spectral theorem, and we deduce that

$$\mathrm{dist}(\lambda_n^{\mathrm{mod}}(\hbar), \mathrm{sp}(\mathscr{N}_\hbar^\sharp)) \leqslant C\hbar^3.$$

In particular, this implies that, for $\hbar$ small enough, $\lambda_n^{\mathrm{mod}}(\hbar)$ is real. This shows that we necessarily have

$$p_1(0) \in \mathbb{R}, \quad p_2(0) - \frac{\alpha^2+\beta^2}{2d_0} \in \mathbb{R}.$$

This also implies that

$$\lambda_n(\mathscr{N}_\hbar^\sharp) \leqslant \lambda_n^{\mathrm{mod}}(\hbar) + C\hbar^3. \tag{5-5}$$

**5.2.2.** *From $\mathscr{N}_\hbar^\sharp$ to the model operator.* Let $n \geqslant 1$. Let us consider an eigenfunction $\psi$ of $\mathscr{N}_\hbar^\sharp$ associated with the eigenvalue $\lambda_n(\mathscr{N}_\hbar^\sharp)$.

We know that $\lambda_n(\mathscr{N}_\hbar^\sharp) = \beta_{\min} + o(1)$ and that the corresponding eigenfunctions are localized in $(z,t)$ (due to the Agmon estimates and Proposition 3.3). Thus, in (4-8), we can replace $\mathscr{N}_\hbar^\flat$ by $\mathscr{N}_\hbar^\sharp$, and we deduce that

$$((p_\hbar^{\mathrm{eff}})^W - \lambda_n(\mathscr{N}_\hbar^\sharp))\mathfrak{P}\psi = \mathscr{O}(\hbar^{3-\eta})\|\psi\|, \quad \|\psi\| \leqslant C\|\mathfrak{P}\psi\|, \tag{5-6}$$

for $\eta > 0$ as small as we want. We use Remark 3.4 to control the remainders $\|\psi\|_{L^2(\mathbb{R},\mathscr{B}_N)}$ by $\mathscr{O}(\hbar^{-\eta})\|\psi\|$. By taking the scalar product with $\mathfrak{P}\psi$, taking the real part and using the min-max principle, we get that

$$\lambda_n(\mathscr{N}_\hbar^\sharp) \geqslant \beta_{\min} + p_1(0)\hbar - C\hbar^2.$$

This establishes the two-term asymptotic estimate

$$\lambda_n(\mathscr{N}_\hbar^\sharp) = \beta_{\min} + p_1(0)\hbar + \mathscr{O}(\hbar^2).$$

Therefore, we can focus on the description of the eigenvalues of the form

$$\lambda_n(\mathscr{N}_\hbar^\sharp) = \beta_{\min} + p_1(0)\hbar + \mu_n(\hbar)\hbar^2$$

for $\mu_n(\hbar) \in D(0,R)$ with a given $R > 0$. We have

$$((p_\hbar^{\mathrm{eff}})^W - (\beta_{\min} + p_1(0)\hbar + \mu_n(\hbar)\hbar^2))\mathfrak{P}\psi_n = \mathscr{O}(\hbar^{3-\eta})\|\mathfrak{P}\psi_n\|, \tag{5-7}$$

where $\psi_n$ denotes a normalized eigenfunction associated to the $n$-th eigenvalue of $\mathscr{N}_\hbar^\sharp$. In fact, by considering (5-7) and again Proposition 4.7, the function $\mathfrak{P}\psi_n$ is microlocalized near (0, 0), the minimum of the principal symbol $p_0$. Since this minimum is nondegenerate, the quadratic approximation of the symbol shows that $\mathfrak{P}\psi_n$ is microlocalized near $(u, \upsilon) = (0, 0)$ at the scale $\hbar^{1-\eta}$ for any $\eta \in \left(0, \frac{1}{2}\right)$. In particular, we deduce that

$$((p_\hbar^{\text{mod}})^W - (\beta_{\min} + p_1(0)\hbar + \mu_n(\hbar)\hbar^2))\mathfrak{P}\psi_n = \mathscr{O}(\hbar^{3-3\eta})\|\mathfrak{P}\psi_n\|.$$

From the resolvent estimate (5-3), this implies that

$$\mu_n(\hbar) \in \bigcup_{j\geqslant 1} D\left(\frac{d_0}{2}(2j-1) + d_1, C\hbar^{1-3\eta}\right), \quad d_1 = p_2(0) - \frac{\alpha^2+\beta^2}{2d_0},$$

where $D(z, r)$ denotes the disc of center $z \in \mathbb{C}$ and radius $r > 0$. In particular, we have

$$\mu_1(\hbar) \geqslant \tfrac{1}{2}d_0 + d_1 - C\hbar^{1-3\eta}.$$

This shows that

$$\lambda_1(\mathscr{N}_\hbar^\sharp) \geqslant \beta_{\min} + p_1(0)\hbar + \left(\tfrac{1}{2}d_0 + d_1\right)\hbar^2 - C\hbar^{3-3\eta},$$

and thus, with (5-5), we get

$$\mu_1(\hbar) = \tfrac{1}{2}d_0 + d_1 + \mathscr{O}(\hbar^{1-3\eta})$$

and

$$\lambda_1(\mathscr{N}_\hbar^\sharp) = \lambda_1^{\text{mod}}(\hbar) + \mathscr{O}(\hbar^{3-3\eta}).$$

Let us now deal with $\lambda_2(\mathscr{N}_\hbar^\sharp)$ and recall (5-5). Assume by contradiction that $\mu_2(\hbar) \in D\left(\frac{1}{2}d_0+d_1, C\hbar^{1-3\eta}\right)$. Then, we have

$$|\mu_2(\hbar) - \mu_1(\hbar)| \leqslant C\hbar^{1-3\eta}.$$

We infer that

$$((p_\hbar^{\text{mod}})^W - \lambda_1^{\text{mod}}(\hbar))\mathfrak{P}\psi = \mathscr{O}(\hbar^{3-3\eta})\|\mathfrak{P}\psi\|$$

for all $\psi \in \operatorname{span}(\psi_1, \psi_2)$. Moreover, coming back to (4-8) (see also (5-7)), we also get that $\|\psi\| \leqslant C\|\mathfrak{P}\psi\|$ for all $\psi \in \operatorname{span}(\psi_1, \psi_2)$. In particular, $\mathfrak{P}(\operatorname{span}(\psi_1, \psi_2))$ is of dimension 2. Let us consider the Riesz projector (in the characteristic subspace of $(p_\hbar^{\text{mod}})^W$ associated with the smallest eigenvalue)

$$\Pi = \frac{1}{2i\pi}\int_{\mathscr{C}(\lambda_1^{\text{mod}}(\hbar), \hbar^{3-4\eta})} (\zeta - (p_\hbar^{\text{mod}})^W)^{-1}\,\mathrm{d}\zeta,$$

which is of rank 1. Then, for all $\varphi \in \mathfrak{P}(\operatorname{span}(\psi_1, \psi_2))$, we write, with the Cauchy formula,

$$\Pi\varphi = \varphi + \frac{1}{2i\pi}\int_{\mathscr{C}(\lambda_1^{\text{mod}}(\hbar), \hbar^{3-4\eta})} ((\zeta - (p_\hbar^{\text{mod}})^W)^{-1} - (\zeta - \lambda_1^{\text{mod}}(\hbar))^{-1})\varphi\,\mathrm{d}\zeta.$$

But, we have

$$(\zeta - (p_\hbar^{\text{mod}})^W)^{-1} - (\zeta - \lambda_1^{\text{mod}}(\hbar))^{-1} = (\zeta - \lambda_1^{\text{mod}}(\hbar))^{-1}(\zeta - (p_\hbar^{\text{mod}})^W)^{-1}((p_\hbar^{\text{mod}})^W - \lambda_1^{\text{mod}}(\hbar)),$$

so that, by using the resolvent estimate (5-3), we get

$$\|\Pi\varphi-\varphi\| \leqslant C\hbar^{3-4\eta}\hbar^{-3+4\eta}\hbar^{-3+4\eta}\hbar^{3-3\eta}\|\varphi\| = C\hbar^{\eta}\|\varphi\|.$$

This shows that the range of $\Pi$ is of dimension at least 2 as soon as $\hbar$ is small enough. This is a contradiction. Therefore, we must have $\mu_2(\hbar) \in D\big(3\big(\frac{1}{2}d_0\big)+d_1, C\hbar^{1-3\eta}\big)$. In particular, we have

$$\mu_2(\hbar) = 3\big(\tfrac{1}{2}d_0\big)+d_1+\mathscr{O}(\hbar^{1-3\eta}), \quad \lambda_2(\mathscr{N}_{\hbar}^{\sharp}) = \lambda_2^{\mathrm{mod}}(\hbar)+\mathscr{O}(\hbar^{3-3\eta}).$$

We proceed by induction to get that, for all $n \geqslant 1$,

$$\mu_n(\hbar) = (2n-1)\big(\tfrac{1}{2}d_0\big)+d_1+\mathscr{O}(\hbar^{1-3\eta}), \quad \lambda_n(\mathscr{N}_{\hbar}^{\sharp}) = \lambda_n^{\mathrm{mod}}(\hbar)+\mathscr{O}(\hbar^{3-3\eta}). \tag{5-8}$$

**5.2.3.** *End of the proof of Theorem 1.4.* Proposition 2.11 shows that the first eigenvalues of $\mathscr{L}_h$ coincide with those of $\mathscr{L}_h^{\mathrm{app}}$ modulo $o(h^2)$. Then, by (3-1), $\mathscr{L}_h^{\mathrm{app}}$ is unitarily equivalent to $\mathscr{N}_h$. The operator $\mathscr{N}_h$ is unitarily equivalent to $\hbar^2\mathscr{N}_{\hbar}^{\sharp}$, see (3-10). Theorem 1.4 follows from (5-8) and (5-2) (see also Remark 5.1 for the explicit formula for $d_0$).

## Acknowledgements

This work was conducted within the France 2030 framework programme, Centre Henri Lebesgue ANR-11-LABX-0020-01.

## References

[Bonnaillie-Noël et al. 2022] V. Bonnaillie-Noël, F. Hérau, and N. Raymond, "Purely magnetic tunneling effect in two dimensions", *Invent. Math.* **227**:2 (2022), 745–793. MR Zbl

[Fahs et al. 2024] R. Fahs, L. Le Treust, N. Raymond, and S. Vũ Ngọc, "Boundary states of the Robin magnetic Laplacian", *Doc. Math.* **29**:5 (2024), 1157–1200. MR Zbl

[Fournais and Helffer 2010] S. Fournais and B. Helffer, *Spectral methods in surface superconductivity*, Progr. Nonlinear Differential Equations Appl. **77**, Birkhäuser, Boston, 2010. MR Zbl

[Fournais et al. 2023] S. Fournais, B. Helffer, A. Kachmar, and N. Raymond, "Effective operators on an attractive magnetic edge", *J. Éc. polytech. Math.* **10** (2023), 917–944. MR Zbl

[Helffer and Kachmar 2023] B. Helffer and A. Kachmar, "Helical magnetic fields and semi-classical asymptotics of the lowest eigenvalue", *Doc. Math.* **28**:4 (2023), 857–901. MR Zbl

[Helffer and Morame 2002] B. Helffer and A. Morame, "Magnetic bottles for the Neumann problem: the case of dimension 3", pp. 71–84 in *Spectral and inverse spectral theory* (Goa, 2000), vol. 112, edited by P. D. Hislop and M. Krishna, 2002. MR Zbl

[Helffer and Morame 2004] B. Helffer and A. Morame, "Magnetic bottles for the Neumann problem: curvature effects in the case of dimension 3 (general case)", *Ann. Sci. École Norm. Sup.* (4) **37**:1 (2004), 105–170. MR Zbl

[Helffer et al. 2016] B. Helffer, Y. Kordyukov, N. Raymond, and S. Vũ Ngọc, "Magnetic wells in dimension three", *Anal. PDE* **9**:7 (2016), 1575–1608. MR Zbl

[Hérau and Raymond 2024] F. Hérau and N. Raymond, "Semiclassical spectral gaps of the 3D Neumann Laplacian with constant magnetic field", *Ann. Inst. Fourier* (*Grenoble*) **74**:3 (2024), 915–972. MR Zbl

[Kato 1995] T. Kato, *Perturbation theory for linear operators*, Springer, 1995. MR Zbl

[Keraval 2018] P. Keraval, *Formules de Weyl par réduction de dimension*: *applications à des Laplaciens électro-magnétiques*, Ph.D. thesis, Université Rennes 1, 2018, available at https://tinyurl.com/PKeraval-Thesis.

[Lu and Pan 2000] K. Lu and X.-B. Pan, "Surface nucleation of superconductivity in 3-dimensions", *J. Differential Equations* **168**:2 (2000), 386–452. MR Zbl

[Martinez 2002] A. Martinez, *An introduction to semiclassical and microlocal analysis*, Springer, 2002. MR Zbl

[Morin et al. 2023] L. Morin, N. Raymond, and S. Vũ Ngọc, "Eigenvalue asymptotics for confining magnetic Schrödinger operators with complex potentials", *Int. Math. Res. Not.* **2023**:17 (2023), 14547–14593. MR Zbl

[Raymond 2009] N. Raymond, *Méthodes spectrales et théorie des cristaux liquides*, Ph.D. thesis, Université Paris-Sud 11, 2009, available at https://theses.hal.science/tel-00424859v1.

[Raymond 2010a] N. Raymond, "On the semiclassical 3D Neumann Laplacian with variable magnetic field", *Asymptot. Anal.* **68**:1-2 (2010), 1–40. MR Zbl

[Raymond 2010b] N. Raymond, "Uniform spectral estimates for families of Schrödinger operators with magnetic field of constant intensity and applications", *Cubo* **12**:1 (2010), 67–81. MR Zbl

[Raymond 2012] N. Raymond, "Semiclassical 3D Neumann Laplacian with variable magnetic field: a toy model", *Comm. Partial Differential Equations* **37**:9 (2012), 1528–1552. MR Zbl

[Raymond 2017] N. Raymond, *Bound states of the magnetic Schrödinger operator*, EMS Tracts in Mathematics **27**, European Mathematical Society (EMS), Zürich, 2017. MR Zbl

[Sjöstrand and Zworski 2007] J. Sjöstrand and M. Zworski, "Elementary linear algebra for advanced spectral problems", *Ann. Inst. Fourier* (*Grenoble*) **57**:7 (2007), 2095–2141. MR

[Zworski 2012] M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics **138**, Amer. Math. Soc., Providence, RI, 2012. MR Zbl

Received 11 Jul 2023. Revised 22 May 2024. Accepted 28 Aug 2024.

MAHA AAFARANI: maha.aafarani94@gmail.com
*Laboratoire Angevin de Recherche Mathématique, Faculté des Sciences, LAREMA, Université d'Angers, Angers, France*

KHALED ABOU ALFA: khaled.abou-alfa@etu.univ-nantes.fr
*Laboratoire de Mathématiques Jean Leray, Université de Nantes, Nantes, France*

FRÉDÉRIC HÉRAU: herau@univ-nantes.fr
*Laboratoire de Mathématiques Jean Leray, Université de Nantes, Nantes, France*

NICOLAS RAYMOND: nicolas.raymond@univ-angers.fr
*Laboratoire Angevin de Recherche Mathématique, Faculté des Sciences, LAREMA, Université d'Angers, Angers, France*

ANALYSIS AND PDE
Vol. 18 (2025), No. 8, pp. 1921–1954
DOI: 10.2140/apde.2025.18.1921

# CHARACTERIZATION OF WEIGHTED HARDY SPACES ON WHICH ALL COMPOSITION OPERATORS ARE BOUNDED

PASCAL LEFÈVRE, DANIEL LI, HERVÉ QUEFFÉLEC AND LUIS RODRÍGUEZ-PIAZZA

We give a complete characterization of the sequences $\beta = (\beta_n)$ of positive numbers for which all composition operators on $H^2(\beta)$ are bounded, where $H^2(\beta)$ is the space of analytic functions $f$ on the unit disk $\mathbb{D}$ such that $\sum_{n=0}^{\infty} |a_n|^2 \beta_n < +\infty$ if $f(z) = \sum_{n=0}^{\infty} a_n z^n$. We prove that all composition operators are bounded on $H^2(\beta)$ if and only if $\beta$ is essentially decreasing and slowly oscillating. We also prove that every automorphism of the unit disk induces a bounded composition operator on $H^2(\beta)$ if and only if $\beta$ is slowly oscillating. We give applications of our results.

## 1. Introduction

Let $\beta = (\beta_n)_{n\geq 0}$ be a sequence of positive numbers such that

$$\liminf_{n\to\infty} \beta_n^{1/n} \geq 1. \tag{1-1}$$

The associated weighted Hardy space $H^2(\beta)$ is defined to be the Hilbertian space of analytic functions $f(z) = \sum_{n=0}^{\infty} a_n z^n$ such that

$$\|f\|^2 := \sum_{n=0}^{\infty} |a_n|^2 \beta_n < \infty. \tag{1-2}$$

Condition (1-1) is equivalent to the inclusion $H^2(\beta) \subseteq \mathcal{H}ol(\mathbb{D})$. Indeed, if (1-1) holds, we have $H^2(\beta) \subseteq \mathcal{H}ol(\mathbb{D})$ since $|a_n|^2\beta_n$ is bounded and thanks to the Hadamard formula. Conversely, testing the inclusion $H^2(\beta) \subseteq \mathcal{H}ol(\mathbb{D})$ on the function $f(z) = \sum_{n=1}^{\infty} (n\sqrt{\beta(n)})^{-1} z^n \in H^2(\beta)$, we get (1-1) from the Hadamard formula.

Condition (1-1) will therefore be assumed throughout this paper, without repeating it.

When $\beta_n \equiv 1$, we recover the usual Hardy space $H^2$; the Bergman space corresponds to $\beta_n = 1/(n+1)$ and the Dirichlet space to $\beta_n = n+1$.

Recall that a symbol is a (nonconstant) analytic self-map $\varphi\colon \mathbb{D} \to \mathbb{D}$, and the associated composition operator $C_\varphi\colon H^2(\beta) \to \mathcal{H}ol(\mathbb{D})$ is defined as

$$C_\varphi(f) = f \circ \varphi. \tag{1-3}$$

An important question in the theory is to decide when $C_\varphi$ is bounded on $H^2(\beta)$, i.e., when

$$C_\varphi\colon H^2(\beta) \to H^2(\beta).$$

*MSC2020:* primary 47B33; secondary 30H10.
*Keywords:* composition operator, weighted Hardy space, slowly oscillating sequence, automorphism of the unit disk.

© 2025 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

This question appears in the literature in several places. For instance, it is Problem 1 in the thesis of Nina Zorboska [1988, p. 49]. This thesis contains many interesting results, in particular Propositions 3.1 and 4.2 of the present paper (we discovered the content of Zorboska's thesis once the present paper was almost finished). See also Question 36 raised by Deddens in [Shields 1974, p. 122.c].

When $H^2(\beta)$ is the usual Hardy space $H^2$ (i.e., when $\beta_n \equiv 1$), it is well known, as a consequence of the Littlewood subordination principle [1925], that all symbols generate bounded composition operators [Shapiro 1993, pp. 13–17]. On the other hand, for the Dirichlet space, corresponding to $\beta_n = n+1$, not all composition operators are bounded since there exist symbols $\varphi$ not belonging to the Dirichlet space (e.g., any infinite Blaschke product).

Note that, by definition of the norm of $H^2(\beta)$, all rotations $R_\theta$, defined by $R_\theta(z) = e^{i\theta} z$, with $\theta \in \mathbb{R}$, induce bounded and surjective composition operators on $H^2(\beta)$ and send isometrically $H^2(\beta)$ into itself.

Our goal in this paper is to characterize the sequences $\beta$ for which all composition operators act boundedly on the space $H^2(\beta)$, i.e., send $H^2(\beta)$ into itself.

In Shapiro's presentation for the Hardy space $H^2$, the main point is the case $\varphi(0) = 0$ and a subordination principle for subharmonic functions (Littlewood's subordination principle). The case of automorphisms is claimed to be simple, using an integral representation for the norm and some change of variable. For general weights $\beta$, the situation is different, as we will see in this paper, and it turns out that the conditions on $\beta$ for the boundedness of the composition operators $C_\varphi$ on $H^2(\beta)$ are not the same depending on whether we consider the class of all symbols such that $\varphi(0) = 0$, or the class of symbols $\varphi = T_a$, where

$$T_a(z) = \frac{a+z}{1+\bar{a}z} \quad \text{for } a \in \mathbb{D}. \tag{1-4}$$

It is clear that when these two classes of composition operators are bounded, then all composition operators are bounded. Recall that every symbol $\varphi$ can be written as the composition $\varphi = T_a \circ \psi$, where $\psi(0) = 0$ and $a = \varphi(0)$, and then $C_\varphi = C_\psi \circ C_{T_a}$.

In many occurrences, the weight $\beta$ is defined as

$$\beta_n = \int_0^1 t^n \, d\sigma(t), \tag{1-5}$$

where $\sigma$ is a positive measure on (0, 1); more specifically the following definition is often used: let $G\colon (0,1) \to \mathbb{R}_+$ be an integrable function, and let $H^2_G$ be the space of analytic functions $f\colon \mathbb{D} \to \mathbb{C}$ such that

$$\|f\|^2_{H^2_G} := \int_{\mathbb{D}} |f(z)|^2 G(1-|z|^2)\, dA(z) < \infty. \tag{1-6}$$

Such weighted Bergman-type spaces are used, for instance, in [Kellay and Lefèvre 2012; Kriete and MacCluer 1995; Li et al. 2014]. We have $H^2_G = H^2(\beta)$ with

$$\beta_n = 2\int_0^1 r^{2n+1} G(1-r^2)\, dr = \int_0^1 t^n G(1-t)\, dt, \tag{1-7}$$

and the sequence $\beta = (\beta_n)_n$ is *nonincreasing* (actually, the representation (1-5) is equivalent, by the Hausdorff moment theorem, to a high regularity of the sequence $\beta$, namely its *complete monotony*).

When the weight $\beta$ is nonincreasing (or more generally, essentially decreasing), all the symbols vanishing at the origin induce a bounded composition operator. This was proved by C. Cowen [1990, Corollary, p. 31], using Hadamard multiplication. We can also use Kacnelson's theorem (see [Chalendar and Partington 2014] or [Lefèvre et al. 2021, Theorem 3.12]). Actually that follows from an older theorem of Goluzin [1951] (see [Duren 1983, Theorem 6.3]), which itself uses a self-refinement observed by Rogosinski of Littlewood's principle [Duren 1983, Theorem 6.2].

For weights defined as in (1-5), we have at our disposal integral representations for the norm in $H^2(\beta)$, and, as in the Hardy space case, this integral representation rather easily allows us to decide when the boundedness of $C_{T_a}$ on $H^2(\beta)$ occurs. This is not always the case, as shown by T. Kriete and B. MacCluer in [Kriete and MacCluer 1995]. They consider spaces of Bergman-type $A^2_{\widetilde{G}} := H^2_G$, where $\widetilde{G}(r) = G(1-r^2)$, defined as the spaces of analytic functions in $\mathbb{D}$ such that

$$\int_{\mathbb{D}} |f(z)|^2 \widetilde{G}(|z|)\, dA < \infty$$

for a positive nonincreasing continuous function $\widetilde{G}$ on $[0, 1)$. They prove [Kriete and MacCluer 1995, Theorem 3] that, for

$$\widetilde{G}(r) = \exp\Big(-B\frac{1}{(1-r)^\alpha}\Big), \qquad B > 0, \;\; 0 < \alpha \le 2,$$

and

$$\varphi(z) = z + t(1-z)^\gamma, \qquad 1 < \gamma \le 3, \;\; 0 < t < 2^{1-\gamma},$$

$\varphi$ is a symbol and $C_\varphi$ is bounded on $A^2_{\widetilde{G}}$ if and only if $\gamma \ge \alpha + 1$.

Here

$$\beta_n = \int_0^1 t^n \mathrm{e}^{-B/(1-\sqrt{t})^\alpha}\, dt \lesssim \exp(-cn^{\alpha/(\alpha+1)}).$$

We point out that $\beta$ is nonincreasing, so, for every symbol $\varphi$ fixing the origin, the composition operator $C_\varphi$ is bounded. Nevertheless, choosing $\gamma < \alpha + 1$, there exist symbols inducing an unbounded composition operator, hence not all the $C_{T_a}$ are bounded. Actually, for every $\alpha \in (0, 2]$, no $C_{T_a}$ is bounded because $\beta$ has no polynomial lower estimate (see Proposition 4.5 below).

*Contents of the paper.* In Section 2, we introduce several notions of growth or regularity for a sequence $\beta$ — essentially decreasing, polynomial lower and upper bounds, slow oscillation — and give some connections between them. In Section 3, we consider the composition operators whose symbol vanishes at the origin. We show that, in order for all these operators to be bounded, it is necessary that $\beta$ be bounded above. We show that $\beta$ is essentially decreasing if and only if all these operators are bounded and

$$\sup_{\varphi(0)=0} \|C_\varphi\| < +\infty.$$

In Theorem 3.3, we give a sufficient condition for having all the composition operators $C_\varphi$ with $\varphi(0) = 0$ bounded, allowing us to give an example of a sequence $\beta$ for which this happens even though $\sup_{\varphi(0)=0} \|C_\varphi\| = +\infty$ (Theorem 3.7). In Section 4, we prove that all $C_{T_a}$ are bounded on $H^2(\beta)$ if and only if $\beta$ is slowly oscillating (Theorems 4.6 and 4.9). We state our main result.

**Theorem 1.1.** *Let $\beta$ be a sequence of positive numbers, and let*

$$T_a(z) = \frac{a+z}{1+\bar{a}z}$$

*for $a \in \mathbb{D}$. The following assertions are equivalent*:

(1) *For some $a \in \mathbb{D} \setminus \{0\}$, the map $T_a$ induces a bounded composition operator $C_{T_a}$ on $H^2(\beta)$.*

(2) *For all $a \in \mathbb{D}$, the maps $T_a$ induce bounded composition operators $C_{T_a}$ on $H^2(\beta)$.*

(3) *$\beta$ is slowly oscillating.*

The deep implication is (2) $\Rightarrow$ (3). Its proof requires some sharp estimates on the mean of Taylor coefficients of $T_a$ for $a$ belonging to a subinterval of (0, 1). Once we found the equivalence of (1) and (2), we realized that it already appeared in the thesis of Zorboska [1988].

In Section 5, we show (Theorem 5.1) that if $\beta$ is slowly oscillating, and moreover all composition operators are bounded on $H^2(\beta)$, then $\beta$ is essentially decreasing. We thus obtain the following theorem.

**Theorem 1.2.** *Let $\beta$ be a sequence of positive numbers. Then all composition operators on $H^2(\beta)$ are bounded if and only if $\beta$ is essentially decreasing and slowly oscillating.*

For the notion of essentially decreasing and slowly oscillating sequences, see Definitions 2.1 and 2.2.

We end the paper with some results about multipliers.

A first version of this paper, not including the complete characterization given here, was put on arXiv on 30 November 2020 (and a second version on 21 March 2022) under the title “Boundedness of composition operators on general weighted Hardy spaces of analytic functions”.

## 2. Definitions, notation, and preliminary results

The open unit disk of $\mathbb{C}$ is denoted by $\mathbb{D}$ and we write $\mathbb{T}$ for its boundary $\partial\mathbb{D}$. We set

$$e_n(z) = z^n, \quad n \geq 0.$$

The weighted Hardy space $H^2(\beta)$ defined in the introduction is a Hilbert space with the canonical orthonormal basis

$$e_n^\beta(z) = \frac{1}{\sqrt{\beta_n}} z^n, \quad n \geq 0, \tag{2-1}$$

and the reproducing kernel $K_w$ given, for all $w \in \mathbb{D}$, by

$$K_w(z) = \sum_{n=0}^{\infty} e_n^\beta(z)\overline{e_n^\beta(w)} = \sum_{n=0}^{\infty} \frac{1}{\beta_n} \overline{w}^n z^n. \tag{2-2}$$

Note that $H^2$ is continuously embedded in $H^2(\beta)$ if and only if $\beta$ is bounded above. In particular, this is the case when $\beta$ is nonincreasing. In this paper, we need a slightly more general notion.

**Definition 2.1.** A sequence of positive numbers $\beta = (\beta_n)_{n\geq 0}$ is said to be *essentially decreasing* if, for some constant $C \geq 1$, we have, for all $m \geq n \geq 0$,

$$\beta_m \leq C\beta_n. \tag{2-3}$$

Note: saying that $\beta$ is essentially decreasing means that the shift operator on $H^2(\beta)$ is power bounded. If $\beta$ is essentially decreasing and if we set

$$\tilde{\beta}_n = \sup_{m\geq n} \beta_m,$$

the sequence $\tilde{\beta} = (\tilde{\beta}_n)$ is nonincreasing and we have $\beta_n \leq \tilde{\beta}_n \leq C\beta_n$. In particular, $H^2(\beta) = H^2(\tilde{\beta})$ (with equivalent norms) and $H^2$ is continuously embedded in $H^2(\beta)$.

**Definition 2.2.** A sequence $\beta$ is *slowly oscillating* if there are positive constants $c < 1 < C$ such that

$$c \leq \frac{\beta_m}{\beta_n} \leq C \quad \text{when } n/2 \leq m \leq 2n. \tag{2-4}$$

We may remark that this is equivalent to the existence of some function $\rho\colon (0, \infty) \to (0, \infty)$ which is bounded above on each compact subset of $(0, \infty)$ and for which $\beta_m/\beta_n \leq \rho(m/n)$, equivalently

$$\frac{1}{\rho(n/m)} \leq \frac{\beta_m}{\beta_n} \leq \rho(m/n).$$

**Definition 2.3.** The sequence of positive numbers $\beta = (\beta_n)$ is said to have a *polynomial lower bound* if there are positive constants $c$ and $\alpha$ such that, for all integers $n \geq 1$,

$$\beta_n \geq cn^{-\alpha}. \tag{2-5}$$

This means that $H^2(\beta)$ is continuously embedded in the weighted Bergman space $\mathfrak{B}^2_{\alpha-1}$ of the analytic functions $f\colon \mathbb{D} \to \mathbb{C}$ such that

$$\|f\|^2_{\mathfrak{B}^2_{\alpha-1}} := \alpha \int_{\mathbb{D}} |f(z)|^2 (1 - |z|^2)^{\alpha-1}\, dA(z) < \infty$$

since $\mathfrak{B}^2_{\alpha-1} = H^2(\gamma)$ with $\gamma_n \approx n^{-\alpha}$.

**Definition 2.4.** The sequence of positive numbers $\beta = (\beta_n)$ is said to have a *polynomial upper bound* if there are positive constants $C$ and $\gamma$ such that, for all integers $n \geq 1$,

$$\beta_n \leq Cn^{\gamma}. \tag{2-6}$$

The following simple proposition links these notions.

**Proposition 2.5.** (1) *Every slowly oscillating sequence $\beta$ has polynomial lower and upper bounds.*

(2) *There are sequences that are essentially decreasing and with polynomial lower bound but are not slowly oscillating.*

(3) *There are bounded sequences that are slowly oscillating but not essentially decreasing.*

*Proof.* (1) This is clear because, for some $c \in (0, 1)$, if $2^j \leq n < 2^{j+1}$, then

$$\beta_n \geq c\beta_{2^j} \geq c^{j+1}\beta_1 \geq c\beta_1 n^{-\alpha},$$

with $\alpha = \log(1/c)/\log 2$, and, for some $C > 1$,

$$\beta_n \leq C\beta_{2^j} \leq C^{j+1}\beta_1 \leq C\beta_1 n^{\gamma},$$

with $\gamma = \log C/\log 2$.

(2) Let $\delta > 0$. We set $\beta_0 = \beta_1 = 1$ and, for $n \geq 2$,

$$\beta_n = \frac{1}{(k!)^\delta} \quad \text{when } k! < n \leq (k+1)!.$$

The sequence $\beta$ is nonincreasing.

For $n$ and $k$ as above, we have

$$\beta_n = \frac{1}{(k!)^\delta} \geq \frac{1}{n^\delta};$$

hence $\beta$ has arbitrarily slow polynomial lower bound. However we have, for $k \geq 2$,

$$\frac{\beta_{2(k!)}}{\beta_{k!}} = \frac{(k!)^{-\delta}}{[(k-1)!]^{-\delta}} = \frac{1}{k^\delta} \xrightarrow[k\to\infty]{} 0,$$

so $\beta$ is not slowly oscillating.

(3) We define $\beta_n$ as follows. Let $(a_k)$ be an increasing sequence of positive square integers such that $\lim_{k\to\infty} a_{k+1}/a_k = \infty$, for example $a_k = 4^{k^2}$, and let $b_k = \sqrt{a_k a_{k+1}}$; with our choice, this is an integer and we clearly have $a_k < b_k < a_{k+1}$. We set

$$\beta_n = \begin{cases} a_k/n & \text{for } a_k \leq n < b_k, \\ (a_k/b_k^2)n = (1/a_{k+1})n & \text{for } b_k \leq n < a_{k+1}. \end{cases}$$

The sequence $(\beta_n)$ is slowly oscillating by construction. Indeed, since the other cases are obvious, it suffices to check that, for $a_k \leq n/2 < b_k \leq n < a_{k+1}$, the quotient $\beta_m/\beta_n$ remains lower and upper bounded when $n/2 \leq m \leq n$ (it will then be automatically also satisfied when $n \leq m \leq 2n$). But, for $n/2 \leq m < b_k$, we have

$$\frac{\beta_m}{\beta_n} = \frac{a_k/m}{n/a_{k+1}} = \frac{a_k a_{k+1}}{mn} = \frac{b_k^2}{mn},$$

which is $\leq 2b_k^2/n^2 \leq 2$ and $\geq b_k^2/n^2 \geq (n/2)^2/n^2 = \frac{1}{4}$; and, for $b_k \leq m$, we have

$$\frac{\beta_m}{\beta_n} = \frac{m/a_{k+1}}{n/a_{k+1}} = \frac{m}{n} \in \left[\tfrac{1}{2}, 1\right].$$

However, even though $(\beta_n)$ is bounded, since $\beta_n \leq 1$ for $a_k \leq n < b_k$ and, for $b_k \leq n < a_{k+1}$,

$$\beta_n \leq \beta_{a_{k+1}-1} = \frac{1}{a_{k+1}}(a_{k+1}-1) \leq 1,$$

it is not essentially decreasing, since

$$\frac{\beta_{a_{k+1}-1}}{\beta_{b_k}} = \frac{1}{\sqrt{a_k a_{k+1}}}(a_{k+1}-1) \sim \sqrt{\frac{a_{k+1}}{a_k}} \xrightarrow[k\to\infty]{} \infty. \qquad \square$$

Now we are going to recall some well-known facts about matrix representation of an operator $T$ defined on a Hilbert space with an orthonormal basis $(\mathrm{e}_n)_{n\geq 0}$ and explain how they translate into our framework.

The entry $a_{m,n}$ (where $m, n \geq 0$) is defined by the $m$-th coordinate of $T(\mathrm{e}_n)$:

$$a_{m,n} = \mathrm{e}_m^*(T(\mathrm{e}_n)),$$

where $\mathrm{e}_k^*(x)$ stands for the $k$-th coordinate of the vector $x$.

We shall use the notation $\hat{f}(k)$ for the $k$-th Fourier coefficient of a function $f \in L^1(-\pi,\pi)$:

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}\,dt.$$

Let us point out that when the operator is the composition operator $C_\varphi$ associated to the symbol $\varphi$, viewed on $H^2(\beta)$, its matrix representation in the basis $(e_n^\beta)_{n\geq 0}$ has an entry $(m,n)$, which we write as

$$(e_m^\beta)^*(C_\varphi(e_n^\beta)) = \frac{\sqrt{\beta_m}}{\sqrt{\beta_n}} e_m^*(\varphi^n) = \frac{\sqrt{\beta_m}}{\sqrt{\beta_n}}\widehat{\varphi^n}(m)$$

since the $m$-th Taylor coefficient of $\varphi^n$ coincides with its $m$-th Fourier coefficient.

We say that the reproducing kernels $K_w$ have a *slow growth* if

$$\|K_w\| \leq \frac{C}{(1-|w|)^s} \tag{2-7}$$

for positive constants $C$ and $s$. We have the following equivalence.

**Proposition 2.6.** *The sequence $\beta$ has polynomial lower bound if and only if the reproducing kernels $K_w$ of $H^2(\beta)$ have a slow growth.*

*Proof.* Assume that the reproducing kernels have a slow growth. Since

$$\|K_w\|^2 = \sum_{k=0}^{\infty} \frac{|w|^{2k}}{\beta_k},$$

we get, for any $k \geq 2$,

$$\frac{|w|^{2k}}{\beta_k} \leq \frac{C^2}{(1-|w|)^{2s}}.$$

Taking $w = 1 - 1/k$, we obtain $\beta_k \geq C' k^{-2s}$.

For the necessity, we only have to see that

$$\|K_w\|^2 = \frac{1}{\beta_0} + \sum_{n=1}^{\infty}\frac{|w|^{2n}}{\beta_n} \leq \frac{1}{\beta_0} + \delta^{-1}\sum_{n=1}^{\infty} n^\alpha |w|^{2n} \leq \frac{C}{(1-|w|^2)^{\alpha+1}}. \qquad \square$$

## 3. Boundedness of composition operators whose symbol vanishes at the origin

**3.1.** ***Necessary conditions.*** We begin with this simple observation, see [Zorboska 1988, Proposition 3.1].

**Proposition 3.1.** *If all composition operators with symbol vanishing at* 0 *are bounded on $H^2(\beta)$, then the sequence $\beta$ is bounded above.*

*Proof.* Let $f \in H^\infty$. Write $f = A\varphi + f(0)$, where $A$ is a constant and $\varphi$ a symbol vanishing at 0. We have $\varphi = C_\varphi(z) \in H^2(\beta)$, by hypothesis, and so $f \in H^2(\beta)$ and $H^\infty \subseteq H^2(\beta)$. It follows (by the closed graph theorem, since the convergence in norm implies pointwise convergence) that there exists a constant $M$ such that $\|f\|_{H^2(\beta)} \leq M\|f\|_\infty$ for all $f \in H^\infty$. Testing this with $f(z) = z^n$, we get $\beta_n \leq M^2$. $\square$

Let us point out that boundedness of $\beta_n$ does not suffice. For example, let $(\beta_n)$ be a sequence such that $\beta_{4k+2}/\beta_{2k+1} \xrightarrow[k\to\infty]{} \infty$ (for instance $\beta_{2k} = 1$ and $\beta_{2k+1} = 1/(k+1)$); if $\varphi(z) = z^2$, then

$$\|C_\varphi(z^{2n+1})\|^2 = \|z^{2(2n+1)}\|^2 = \beta_{2(2n+1)};$$

since $\|z^{2n+1}\|^2 = \beta_{2n+1}$, the operator $C_\varphi$ is not bounded on $H^2(\beta)$.

A partial characterization is given in the next proposition.

**Proposition 3.2.** *The following assertions are equivalent*:

(1) *All symbols $\varphi$ such that $\varphi(0) = 0$ induce bounded composition operators $C_\varphi$ on $H^2(\beta)$* **and**

$$\sup_{\varphi(0)=0} \|C_\varphi\| < \infty. \tag{3-1}$$

(2) *$\beta$ is an essentially decreasing sequence.*

Of course, by the uniform boundedness principle, (3-1) is equivalent to

$$\sup_{\varphi(0)=0} \|f\circ\varphi\| < \infty \quad \text{for all } f \in H^2(\beta).$$

Let us point out an important fact: we shall see in Theorem 3.7 that there are weights $\beta$ for which all composition operators $C_\varphi$ with $\varphi(0) = 0$ are bounded but $\sup_{\varphi(0)=0} \|C_\varphi\| = +\infty$.

*Proof.* (2) $\Rightarrow$ (1) We may assume that $\beta$ is nonincreasing. Then the Goluzin–Rogosinski theorem [Duren 1983, Theorem 6.3] gives the result; in fact, writing

$$f(z) = \sum_{n=0}^{\infty} c_n z^n \quad \text{and} \quad (C_\varphi f)(z) = \sum_{n=0}^{\infty} d_n z^n,$$

it says that

$$\sum_{0\le k\le n} |d_k|^2 \le \sum_{0\le k\le n} |c_k|^2 \quad \text{for all } n \ge 0,$$

and hence, by Abel summation,

$$\|C_\varphi f\|^2 = \sum_{n=0}^{\infty} |d_n|^2\beta_n \le \sum_{n=0}^{\infty} |c_n|^2\beta_n = \|f\|^2,$$

leading to $C_\varphi$ bounded and $\|C_\varphi\| \le 1$. This same result was also proved by Cowen [1990, Corollary of Theorem 7]. Alternatively, we can use a result of Kacnelson [1972]; see also [Chalendar and Partington 2014; 2017, Corollary 2.2; Lefèvre et al. 2021, Theorem 3.12].

(1) $\Rightarrow$ (2) Set $M = \sup_{\varphi(0)=0} \|C_\varphi\|$. Let $m > n$, and take

$$\varphi(z) = \varphi_{m,n}(z) = z\big(\tfrac{1}{2}(1 + z^{m-n})\big)^{1/n}.$$

Then $\varphi(0) = 0$ and $[\varphi(z)]^n = \frac{1}{2}(z^n + z^m)$; hence

$$\tfrac{1}{4}(\beta_n + \beta_m) = \|\varphi^n\|^2 = \|C_\varphi(e_n)\|^2 \le \|C_\varphi\|^2\|e_n\|^2 \le M^2\beta_n,$$

so $\beta$ is essentially decreasing. □

**Remark.** Let us mention the following example. For $0 < r < 1$, let $\beta_n = \pi n r^{2n}$ for $n \geq 1$ and $\beta_0 = 1$. This sequence is eventually decreasing, so it is essentially decreasing. The quantity $\|f\|^2_{H^2(\beta)} - |f(0)|^2$ is the area of the part of the Riemann surface on which $r\mathbb{D}$ is mapped by $f$. E. Reich [1954], generalizing Goluzin's result [1951] (see [Duren 1983, Theorem 6.3]), proved that, for all symbols $\varphi$ such that $\varphi(0) = 0$, the composition operator $C_\varphi$ is bounded on $H^2(\beta)$ and

$$\|C_\varphi\| \leq \sup_{n\geq 1} \sqrt{n} r^{n-1} \leq \frac{1}{\sqrt{2\mathrm{e}}} \frac{1}{r\sqrt{\log(1/r)}}.$$

For $0 < r < 1/\sqrt{2}$, Goluzin's theorem asserts that $\|C_\varphi\| \leq 1$.

Note that this sequence $\beta$ is not slowly oscillating, since $\beta_{2n}/\beta_n = 2r^{2n}$. Hence, from Theorem 4.9 below, we get that no composition operator $C_{T_a}$ is bounded on $H^2(\beta)$.

However, that the weight $\beta$ is essentially decreasing is not necessary for the boundedness of all composition operators $C_\varphi$, with symbol $\varphi$ vanishing at 0, as we will see later (Theorem 3.7).

### 3.2. *Sufficient condition.*

**Theorem 3.3.** *Let $\beta = (\beta_n)_{n=0}^\infty$ be a sequence of positive numbers that is **weakly decreasing**, i.e.,*

$$\begin{aligned} &\text{for every } \delta > 0, \text{ there exists a positive constant } C = C(\delta) \\ &\text{such that } \beta_m \leq C\beta_n \text{ whenever } m > (1+\delta)n. \end{aligned} \tag{3-2}$$

*Then, for all symbols $\varphi\colon \mathbb{D} \to \mathbb{D}$ vanishing at 0, the composition operator $C_\varphi$ is bounded on $H^2(\beta)$.*

Let us point out that (3-2) implies that $\beta$ is bounded.

Note that Zorboska showed [1988, Example 1, pp. 14-15] that, for $\beta_n = \exp(n^a)$, with $0 < a < 1$, which is unbounded, the symbol $\varphi(z) = z^k$, $k \geq 2$, induces an unbounded composition operator on $H^2(\beta)$.

To prove Theorem 3.3, we need several lemmas.

**Lemma 3.4.** *Let $\varphi\colon \mathbb{D} \to \mathbb{D}$ be an analytic self-map such that $\varphi(0) = 0$ and $|\varphi'(0)| < 1$. Then there exists $\rho > 0$ such that, for all integers $n$ and $m$,*

$$|\widehat{\varphi^n}(m)| \leq \exp(-[(1+\rho)n - m]).$$

*Proof.* Since $\varphi(0) = 0$, we can write $\varphi(z) = z\varphi_1(z)$. Since $|\varphi'(0)| < 1$, we have $\varphi_1\colon \mathbb{D} \to \mathbb{D}$. Now let $M(r) = \sup_{|z|=r} |\varphi_1(z)|$. Cauchy's inequalities say that $|\widehat{\varphi_1^n}(m)| \leq [M(r)]^n / r^m$. We have $M(r) < 1$, so there exists a positive number $\rho = \rho(r)$ such that $M(r) = r^\rho$. We get

$$|\widehat{\varphi^n}(m)| = |\widehat{\varphi_1^n}(m-n)| \leq \frac{r^{\rho n}}{r^{m-n}} = r^{(1+\rho)n - m},$$

and the result follows by taking $r = \mathrm{e}^{-1}$. ☐

The following result of V. È. Kacnelson [1972] was used in [Chalendar and Partington 2014; 2017, Corollary 2.2]; see also [Lefèvre et al. 2021, Theorem 3.12].

**Theorem 3.5** [Kacnelson 1972]. *Let $H$ be a separable complex Hilbert space, and let $(e_i)_{i\geq 0}$ be a fixed orthonormal basis of $H$. Let $M\colon H \to H$ be a bounded linear operator. We assume that the matrix of $M$ with respect to this basis is lower-triangular: $\langle Me_j \,|\, e_i\rangle = 0$ for $i < j$.*

*Let $(\gamma_j)_{j\geq 0}$ be a nondecreasing sequence of positive real numbers, and let $\Gamma$ be the (possibly unbounded) diagonal operator such that $\Gamma(e_j) = \gamma_j e_j$, $j \geq 0$. Then the operator $\Gamma^{-1}M\Gamma\colon H \to H$ is bounded and, moreover,*

$$\|\Gamma^{-1}M\Gamma\| \leq \|M\|.$$

We need the following generalization of Kacnelson's theorem, which is implicitly used in [Lefèvre et al. 2021, p. 13]. The matrix $A$ only needs to be lower-triangular with respect to the order induced by the sequence $(d_n)_n$.

**Lemma 3.6.** *Let $A\colon \ell_2 \to \ell_2$ be a bounded operator represented by the matrix $(a_{m,n})_{m,n}$, i.e., $a_{m,n} = \langle Ae_n, e_m\rangle$, where $(e_n)_{n\geq 0}$ is the canonical basis of $\ell_2$.*

*Let $(d_n)$ be a sequence of positive numbers such that, for every $m$ and $n$,*

$$d_m < d_n \quad\Longrightarrow\quad a_{m,n} = 0. \tag{3-3}$$

*Then, $D$ being the (possibly unbounded) diagonal operator with entries $d_n$, we have*

$$\|D^{-1}AD\| \leq \|A\|.$$

We will propose two different proofs. The first one, using complex variables, is an adaptation of that of Kacnelson, and we reproduce it for the convenience of the reader; the second one is new and uses real variables.

*Proof 1.* Let $\mathbb{C}_0$ be the right-half-plane $\mathbb{C}_0 = \{z \in \mathbb{C} : \Re\mathrm{e}\, z > 0\}$. We set $H_N = \operatorname{span}\{e_n : n \leq N\}$ and

$$A_N = P_N A J_N,$$

where $P_N$ is the orthogonal projection from $\ell_2$ onto $H_N$ and $J_N$ is the canonical injection from $H_N$ into $\ell_2$. We consider, for $z \in \overline{C}_0$,

$$A_N(z) = D^{-z} A_N D^z\colon H_N \to H_N,$$

where $D^z(e_n) = d_n^z e_n$.

If $(a_{m,n}(z))_{m,n}$ is the matrix of $A_N(z)$ on the basis $\{e_n : n \leq N\}$ of $H_N$, we clearly have

$$a_{m,n}(z) = a_{m,n}(d_n/d_m)^z.$$

In particular, we have, thanks to (3-3),

$$a_{m,n}(z) = 0 \quad \text{if } d_m < d_n$$

and

$$|a_{m,n}(z)| \leq \sup_{k,l} |a_{k,l}| := M \quad \text{for all } z \in \overline{C}_0.$$

Since

$$\|A_N(z)\|^2 \leq \|A_N(z)\|_{HS}^2 = \sum_{m,n\leq N} |a_{m,n}(z)|^2 \leq (N+1)^2 M^2,$$

we get

$$\|A_N(z)\| \leq (N+1)M \quad \text{for all } z \in \overline{C}_0.$$

Let us consider the function $u_N : \overline{C}_0 \to \overline{C}_0$ defined by

$$u_N(z) = \|A_N(z)\|. \tag{3-4}$$

This function $u_N$ is continuous on $\overline{C}_0$, bounded above by $(N+1)M$, and subharmonic in $\mathbb{C}_0$. Moreover, thanks to (3-3), the maximum principle gives

$$\sup_{\overline{C}_0} u_N(z) = \sup_{\partial\mathbb{C}_0} u_N(z).$$

Since $\|D^z\| = \|D^{-z}\| = 1$ for $z \in \partial\mathbb{C}_0$, we have

$$\|A_N(z)\| \le \|A_N\| \quad \text{for } z \in \partial\mathbb{C}_0,$$

and we get

$$\sup_{\overline{C}_0} u_N(z) \le \|A_N\| \le \|A\|.$$

In particular, $u_N(1) \le \|A\|$ and, letting $N$ go to infinity, we obtain $\|D^{-1}AD\| \le \|A\|$.

*Proof 2.* Since $d_n$ is positive, we can write $d_n = \mathrm{e}^{-\rho_n}$, where $\rho_n \in \mathbb{R}$. Let $x = (x_n)_{n\ge 0}$ and $y = (y_n)_{n\ge 0}$ be in $\ell^2$ with finite support. We are interested in controlling the sum

$$S = \sum_{m,n} a_{m,n} \frac{d_n}{d_m} x_n \overline{y_m},$$

which can also be written

$$S = \sum_{m,n} a_{m,n} \mathrm{e}^{-|\rho_n - \rho_m|} x_n \overline{y_m}$$

since the nontrivial part of the sum runs over the pairs $(m, n)$ such that $d_m \ge d_n$, i.e., $\rho_n \ge \rho_m$.

Now we introduce the function

$$f(t) = \frac{1}{\pi(1+t^2)} \quad \text{for } t \in \mathbb{R},$$

which is positive and belongs to the unit ball of $L^1(\mathbb{R})$. Moreover, its Fourier transform satisfies, for every $x \in \mathbb{R}$,

$$\mathcal{F}(f)(-x) = \int_{\mathbb{R}} f(t)\mathrm{e}^{ixt}\, dt = \mathrm{e}^{-|x|}.$$

We get

$$S = \int_{\mathbb{R}} f(t) \left( \sum_{m,n} a_{m,n} x_n \mathrm{e}^{i\rho_n t} \overline{y_m \mathrm{e}^{i\rho_m t}} \right) dt = \int_{\mathbb{R}} f(t) \langle A(x(t)), y(t) \rangle_{\ell^2}\, dt,$$

where

$$x(t) = (x_n \mathrm{e}^{i\rho_n t})_{n\ge 0} \quad \text{and} \quad y(t) = (y_n \mathrm{e}^{i\rho_n t})_{n\ge 0}.$$

We obtain

$$|S| \le \int_{\mathbb{R}} f(t) \|A\| \|x(t)\| \|y(t)\|\, dt = \int_{\mathbb{R}} f(t) \|A\| \|x\| \|y\|\, dt = \|A\| \|x\| \|y\|$$

since $\|f\|_{L^1(\mathbb{R})} = 1$.

Since $x$ and $y$ are arbitrary, this proves $\|D^{-1}AD\| \le \|A\|$. □

*Proof of Theorem 3.3.* First, if $|\varphi'(0)| = 1$, we have $\varphi(z) = \alpha z$ for some $\alpha$ with $|\alpha| = 1$, and the result is trivial.

So, we assume that $|\varphi'(0)| < 1$. Then, by Lemma 3.4, there exists $\rho > 0$ such that, for all $m$, $n$,

$$|\widehat{\varphi^n}(m)| \leq \exp(-[(1+\rho)n - m]).$$

It follows that, with $\delta = \frac{1}{2}\rho$, we have

$$|\widehat{\varphi^n}(m)| \leq \exp(-\delta n) \quad \text{when } m \leq (1+\delta)n.$$

Since $\varphi(0) = 0$, we also know that $\widehat{\varphi^n}(m) = 0$ if $m < n$.

Now, using property (3-2), there exists $M \geq 1$ such that

$$\beta_m \leq M\beta_n \quad \text{when } m \geq (1+\delta)n.$$

Define now a new sequence $\gamma = (\gamma_n)$ as

$$\gamma_n = \max\Big\{\beta_n, \sup_{m>(1+\delta)n} \beta_m\Big\}.$$

We have

(1) $\beta_n \leq \gamma_n \leq M\beta_n$,

(2) $\gamma_m \leq \gamma_n$ if $m \geq (1+\delta)n$.

Item (1) implies that $H^2(\gamma) = H^2(\beta)$, and we are reduced to proving that $C_\varphi \colon H^2(\gamma) \to H^2(\gamma)$ is bounded.

Let $A = (a_{m,n})_{m,n} = (\widehat{\varphi^n}(m))_{m,n}$. We have to prove that

$$B = (\gamma_m^{1/2}\gamma_n^{-1/2} a_{m,n})_{m,n}$$

represents a bounded operator on $\ell_2$.

Define the matrix

$$A_1 = (a_{m,n}\mathbb{1}_{\{(m,n):m\leq(1+\delta)n\}})_{m,n},$$

and set $A_2 = A - A_1$. Define analogously $B_1$ and $B_2 = B - B_1$.

Then $A_1$ is a Hilbert–Schmidt operator because (recall that $a_{m,n} = 0$ if $m < n$) we have

$$\sum_{n=1}^{\infty} \sum_{m=1}^{(1+\delta)n} |a_{m,n}|^2 \leq \sum_{n=1}^{\infty} \sum_{m=n}^{(1+\delta)n} \exp(-2\delta n) \leq \sum_{n=1}^{\infty} (\delta n + 1)\exp(-2\delta n) < +\infty.$$

Since $A$ is bounded, it follows that $A_2 = A - A_1$ is bounded.

We now remark that, writing $A_2 = (\alpha_{m,n})_{m,n}$, we have, with $d_n = 1/\sqrt{\gamma_n}$,

$$d_m < d_n \quad \Longrightarrow \quad \gamma_m > \gamma_n \quad \Longrightarrow \quad m < (1+\delta)n \quad \Longrightarrow \quad \alpha_{m,n} = 0.$$

Hence we can apply Lemma 3.6 to the matrix $A_2$, which implies that $B_2$ is bounded.

Now, we have $\liminf \beta_n^{1/n} \geq 1$, so $\beta_n \geq \mathrm{e}^{-\delta n}$ for $n$ large enough; hence we have $\beta_n \geq c\mathrm{e}^{-\delta n}$ for every $n \geq 1$.

Since $\gamma$ is bounded (like $\beta$ is), we have, for some positive constant $C$,

$$\sum_{n=1}^{\infty}\sum_{m=n}^{(1+\delta)n}\frac{\gamma_m}{\gamma_n}|a_{m,n}|^2 \le \sum_{n=1}^{\infty}\sum_{m=n}^{(1+\delta)n}\frac{C}{\beta_n}\exp(-2\delta n) \le \sum_{n=1}^{\infty}\frac{C}{c}(\delta n+1)\exp(-\delta n) < +\infty,$$

meaning that $B_1$ is a Hilbert–Schmidt operator.

Therefore $B = B_1 + B_2$ is bounded, as desired. □

As a corollary of Theorem 3.3, we can provide the following example.

**Theorem 3.7.** *There exists a bounded sequence $\beta$, with polynomial lower bound, which is **not essentially decreasing**, and for which every composition operator with symbol vanishing at 0 is bounded on $H^2(\beta)$.*

*We hence have* $\sup_{\varphi(0)=0}\|C_\varphi\| = +\infty$.

It should be noted that, for this weight, the composition operators are not all bounded, as we will see in Proposition 4.10.

*Proof.* Define $\beta_n = 1$ for $n \le 3!$, and, for $k \ge 3$,

$$\begin{cases}\beta_n = 1/k! & \text{for } k! < n \le (k+1)!-2 \text{ and for } n=(k+1)!,\\ \beta_n = 1/(k+1)! & \text{for } n=(k+1)!-1.\end{cases}$$

Note that, for $m > n$, we have $\beta_m > \beta_n$ only if $n = (k+1)! - 1$ and $m = (k+1)! = n+1$ for some $k \ge 3$.

However $\beta$ is not essentially decreasing since, for every $k \ge 3$, we have $\beta_{n+1}/\beta_n = k+1$ if $n = (k+1)!-1$.

The sequence $\beta$ has a polynomial lower bound because $\beta_n \ge 1/(2n)$ for all $n \ge 1$. In fact, for $k \ge 3$, we have $\beta_n \ge (k+1)/n \ge 1/n$ if $k! < n \le (k+1)!-2$ or if $n = (k+1)!$, and, for $n = (k+1)!-1$, we have $n\beta_n = [(k+1)!-1]/(k+1)! \ge \frac{1}{2}$. It has a polynomial upper bound since it is bounded above by 1.

Now, it remains to check (3-2) in order to apply Theorem 3.3 and finish the proof of Theorem 3.7. Note first that we have $\beta_m/\beta_n \le 1$ if $m \ge n+2$. Next, for given $\delta > 0$, there exists an integer $N$ such that $(1+\delta)n \ge n+2$ for every $n \ge N$, so $\beta_m/\beta_n \le 1$ if $m \ge (1+\delta)n$ and $n \ge N$. It suffices to take $C = \max_{1\le n\le N}\beta_{n+1}/\beta_n$ to obtain (3-2). The last assertion follows from Proposition 3.2. □

## 4. Boundedness of composition operators of the symbol $T_a$

Recall that, for $a \in \mathbb{D}$, we defined

$$T_a(z) = \frac{a+z}{1+\bar{a}z}, \quad z \in \mathbb{D}. \tag{4-1}$$

It is well known that $T_a$ is an automorphism of $\mathbb{D}$ and that $T_a(0) = a$ and $T_a(-a) = 0$.

Though we do not really need this, we remark that $(T_a)_{a\in(-1,1)}$ is a group and $(T_a)_{a\in(0,1)}$ is a semigroup. It suffices to see that $T_a \circ T_b = T_{a*b}$, with

$$a*b = \frac{a+b}{1+ab}. \tag{4-2}$$

In this section, we are going to prove a necessary and sufficient condition for the statement that all composition operators $C_{T_a}$ for $a \in \mathbb{D}$ are bounded on $H^2(\beta)$. Namely, we have the following theorem, the proof of which will occupy Sections 4.2 and 4.3.

**Theorem 4.1.** *All composition operators $C_{T_a}$, with $a \in \mathbb{D}$, are bounded on $H^2(\beta)$ if and only if $\beta$ is slowly oscillating.*

Before that, let us note the following fact; see also [Zorboska 1988, Proposition 3.6]. Recall that if $\varphi$ and $\psi$ are two symbols, then $C_\varphi \circ C_\psi = C_{\psi \circ \varphi}$.

**Proposition 4.2.** *If $C_{T_a}$ is bounded on $H^2(\beta)$ for some $a \in \mathbb{D} \setminus \{0\}$, then $C_{T_b}$ is bounded on $H^2(\beta)$ for all $b \in \mathbb{D}$.*

*Moreover, the maps $C_{T_b}$ are uniformly bounded on the compact subsets of $\mathbb{D}$.*

We decompose the proof into lemmas. The first one was first proved in [Zorboska 1988] (see also [Gallardo-Gutiérrez and Partington 2013, Proposition 2.1]) and follows from the fact that if $b = \rho e^{i\theta}$ and $R_\theta$ is the rotation $R_\theta(z) = e^{i\theta} z$, which induces a unitary operator $C_{R_\theta}$ on $H^2(\beta)$, then $T_b = R_\theta \circ T_\rho \circ R_{-\theta}$ and $C_{T_b} = C_{R_{-\theta}} \circ C_{T_\rho} \circ C_{R_\theta}$.

**Lemma 4.3.** *The composition operator $C_{T_b}$ is bounded if and only if $C_{T_{|b|}}$ is bounded, with equal norms.*

**Lemma 4.4.** *Let $r \in (0, 1)$ such that $C_{T_r}$ is bounded. For any $b \in \mathbb{D}$ satisfying $|b| \le 2r/(1+r^2)$, $C_{T_b}$ is bounded and we have $\|C_{T_b}\| \le \|C_{T_r}\|^2$.*

*Proof.* Let $S$ be the circle $C(0, r)$ and $u \colon S \to \mathbb{R}_+$ be the continuous function defined by

$$u(s) = \left| \frac{s+r}{1+\bar{s}r} \right|. \tag{4-3}$$

By connectedness, $u(S)$ contains the segment $[0, 2r/(1+r^2)] = [u(-r), u(r)]$. Let now

$$b \in D\Big(0, \frac{2r}{1+r^2}\Big).$$

By the above, there exists $s \in S$ such that $|b| = u(s)$. This means that

$$|T_b(0)| = |b| = |u(s)| = |T_s(r)| = |(T_s \circ T_r)(0)|.$$

Therefore, $T_b(0) = e^{i\alpha}(T_s \circ T_r)(0)$ for some $\alpha \in \mathbb{R}$, and hence, by Schwarz's lemma, there is some $\theta \in \mathbb{R}$ such that $T_b = R_\alpha \circ T_s \circ T_r \circ R_\theta$. We then have $C_{T_b} = C_{R_\theta} \circ C_{T_r} \circ C_{T_s} \circ C_{R_\alpha}$. Since $C_{R_\theta}$ and $C_{R_\alpha}$ are unitary, we get, using Lemma 4.3 for $C_{T_s}$,

$$\|C_{T_b}\| = \|C_{T_r} \circ C_{T_s}\| \le \|C_{T_r}\| \|C_{T_s}\| = \|C_{T_r}\|^2. \qquad \square$$

*Proof of Proposition 4.2.* It suffices to use Lemmas 4.3 and 4.4 and do an iteration, noting that if $r_0 = |a| > 0$ and $r_{n+1} = 2r_n/(1+r_n^2) = r_n * r_n$, then $(r_n)_{n\ge 0}$ increases to 1. $\square$

**4.1.** ***An elementary necessary condition.*** We begin with an elementary necessary condition. It is implied by Theorem 4.9, but its statement deserves to be pointed out. Moreover, its proof is simple and highlights the role of the reproducing kernel.

**Proposition 4.5.** *Let $a \in (0, 1)$, and assume that $T_a$ induces a bounded composition operator on $H^2(\beta)$. Then $\beta$ has polynomial lower bound.*

*Proof.* Since

$$\|K_x\|^2 = \sum_{n=0}^{\infty} \frac{x^{2n}}{\beta_n},$$

we have $\|K_x\| \le \|K_y\|$ for $0 \le x \le y < 1$.

We define by induction a sequence $(u_n)_{n\ge 0}$ with

$$u_0 = 0 \quad \text{and} \quad u_{n+1} = T_a(u_n).$$

Since $T_a(1) = 1$ (recall that $a \in (0, 1)$), we have

$$1 - u_{n+1} = \int_{u_n}^{1} T_a'(t)\,dt = \int_{u_n}^{1} \frac{1-a^2}{(1+at)^2}\,dt;$$

hence

$$\frac{1-a}{1+a}(1-u_n) \le 1 - u_{n+1} \le (1-a^2)(1-u_n).$$

Let $0 < x < 1$. We can find $N \ge 0$ such that $u_N \le x < u_{N+1}$. Then

$$1 - x \le 1 - u_N \le (1-a^2)^N.$$

On the other hand, since $C_{T_a}^* K_z = K_{T_a(z)}$ for all $z \in \mathbb{D}$, we have

$$\|K_x\| \le \|K_{u_{N+1}}\| \le \|C_{T_a}\|\|K_{u_N}\| \le \|C_{T_a}\|^{N+1}\|K_{u_0}\| = \frac{1}{\sqrt{\beta_0}}\|C_{T_a}\|^{N+1}.$$

Let $s \ge 0$ such that $(1-a^2)^{-s} = \|C_{T_a}\|$. We obtain

$$\|K_x\| \le \frac{1}{\sqrt{\beta_0}(1-x)^s}\|C_{T_a}\|. \tag{4-4}$$

We get the result by using Proposition 2.6. □

**Remarks.** (1) For example, when $\beta_n = \exp[-c(\log(n+1))^2]$, with $c > 0$, no $T_a$ induces a bounded composition operator on $H^2(\beta)$, even though $C_\varphi$ is bounded for all symbols $\varphi$ with $\varphi(0) = 0$, since $\beta$ is decreasing, as we saw in Proposition 3.2.

(2) For the Dirichlet space $\mathcal{D}^2$, we have $\beta_n = n + 1$, but all the maps $T_a$ induce bounded composition operators on $\mathcal{D}^2$; see [Lefèvre et al. 2021, Remark before Theorem 3.12]. In this case $\beta$ has polynomial upper bound even though it is not bounded above.

(3) However, even for decreasing sequences, a polynomial lower bound for $\beta$ is not enough for some $T_a$ to induce a bounded composition operator. Indeed, we saw in Proposition 2.5 an example of a decreasing sequence $\beta$ with polynomial lower bound but not slowly oscillating, and we will see in Theorem 4.9 that this condition is needed to have some $T_a$ induce a bounded composition operator.

(4) Gallardo-Gutiérrez and Partington [2013] give estimates for the norm of $C_{T_a}$, with $a \in (0, 1)$, when $C_{T_a}$ is bounded on $H^2(\beta)$. More precisely, they proved that if $\beta$ is bounded above and $C_{T_a}$ is bounded, then

$$\|C_{T_a}\| \ge \left(\frac{1+a}{1-a}\right)^{\sigma},$$

where $\sigma = \inf\{s \geq 0 : (1-z)^{-s} \notin H^2(\beta)\}$, and

$$\|C_{T_a}\| \leq \Big(\frac{1+a}{1-a}\Big)^{\tau},$$

where $\tau = \frac{1}{2} \sup \Re\mathrm{e}\, W(A)$, with $A$ the infinitesimal generator of the continuous semigroup $(S_t)$ defined as $S_t = C_{T_{\tanh t}}$, namely $(Af)(z) = f'(z)(1-z^2)$, and $W(A)$ its numerical range.

For $\beta_n = 1/(n+1)^{\nu}$ with $0 \leq \nu \leq 1$, the two bounds coincide, so they get

$$\|C_{T_a}\| = \Big(\frac{1+a}{1-a}\Big)^{(\nu+1)/2}.$$

**4.2. *Sufficient condition.*** The following sufficient condition explains in particular why all composition operators $C_{T_a}$ are bounded on the Dirichlet space.

**Theorem 4.6.** *If $\beta$ is slowly oscillating, then all symbols that extend analytically in a neighborhood of $\overline{\mathbb{D}}$ induce a bounded composition operator on $H^2(\beta)$.*

*In particular, all $C_{T_a}$, for $a \in \mathbb{D}$, are bounded on $H^2(\beta)$.*

To prove Theorem 4.6, we begin with a very elementary fact.

**Lemma 4.7.** *Let $\varphi : \mathbb{D} \to \mathbb{D}$ have an analytic extension to an open neighborhood $\Omega$ of $\overline{\mathbb{D}}$. Then there are a constant $b > 0$ and an integer $\lambda > 1$ such that*

$$|\widehat{\varphi^n}(m)| \leq \begin{cases} \mathrm{e}^{-bn} & \text{if } n \geq \lambda m, \\ \mathrm{e}^{-bm} & \text{if } m \geq \lambda n. \end{cases}$$

*Proof.* Let $R > 1$ such that $\overline{D(0, R)} \subseteq \Omega$. For $0 < r \leq R$, we set

$$M(r) = \sup_{|z|=r} |\varphi(z)|.$$

Take any $r \in (0, 1)$, for instance $r = \mathrm{e}^{-1}$. We have $M(r) < 1$, so we can write $M(r) = \mathrm{e}^{-\rho}$ for some positive $\rho$.

Cauchy's inequalities give

$$|\widehat{\varphi^n}(m)| \leq \frac{[M(r)]^n}{r^m} = \mathrm{e}^{m-\rho n}.$$

Choose $\lambda_1 = \max(2, 2/\rho)$ and $b_1 = \rho - \lambda_1^{-1}$. Then $|\widehat{\varphi^n}(m)| \leq \mathrm{e}^{-b_1 n}$ if $n \geq \lambda_1 m$.

For the second inequality, write $R =: \mathrm{e}^{\beta}$, with $\beta > 0$. Let $\alpha > 0$ with $M(R) \leq \mathrm{e}^{\alpha}$. Cauchy's inequalities again give

$$|\widehat{\varphi^n}(m)| \leq \frac{[M(R)]^n}{R^m} \leq \mathrm{e}^{\alpha n - \beta m}.$$

Choose $\lambda_2 = \max(2, 2\alpha/\beta)$ and $b_2 = \beta - \alpha\lambda_2^{-1}$. Then $|\widehat{\varphi^n}(m)| \leq \mathrm{e}^{-b_2 m}$ if $m \geq \lambda_2 n$. We get the conclusion taking $b = \min(b_1, b_2)$ and choosing an integer $\lambda \geq \max(\lambda_1, \lambda_2)$. □

**Lemma 4.8.** *Let $(\beta_n)$ be a slowly oscillating sequence of positive numbers. Let $A = (a_{m,n})_{m,n}$ be the matrix of a bounded operator on $\ell_2$. Assume that, for some integer $\lambda > 1$ and some constants $c$, $b$, we have:*

(1) $|a_{m,n}| \le c\mathrm{e}^{-bn}$ *when* $n \ge \lambda m$,

(2) $|a_{m,n}| \le c\mathrm{e}^{-bm}$ *when* $m \ge \lambda n$.

*Then the matrix* $\tilde{A} = (a_{m,n}\sqrt{\beta_m/\beta_n})_{m,n}$ *also defines a bounded operator on* $\ell_2$.

*Proof.* In the sequel $\|\cdot\|$ stands for the $\ell^2$-norm.

Since $\beta$ is slowly oscillating, it has polynomial lower and upper bounds: for some $\alpha, \gamma > 0$ and $\delta \in (0,1)$, we have $\delta(n+1)^{-\alpha} \le \beta_n \le \delta^{-1}(n+1)^{\gamma}$.

The matrix $\tilde{A}$ is Hilbert–Schmidt far from the diagonal since

$$\sum_{n=1}^{\infty}\sum_{\lambda m<n}|a_{m,n}|^2\frac{\beta_m}{\beta_n} \lesssim \sum_{n=1}^{\infty}\sum_{\lambda m<n}(n+1)^{\alpha+\gamma}|a_{m,n}|^2 \lesssim \sum_{n=1}^{\infty}(n+1)^{\alpha+\gamma+1}\mathrm{e}^{-2bn} < +\infty$$

and

$$\sum_{n=0}^{\infty}\sum_{m>\lambda n}|a_{m,n}|^2\frac{\beta_m}{\beta_n} \lesssim \sum_{n=0}^{\infty}\sum_{m>\lambda n}(n+1)^{\alpha+\gamma}|a_{m,n}|^2 \lesssim \sum_{n=0}^{\infty}(n+1)^{\alpha+\gamma}\Bigg(\sum_{m>\lambda n}\mathrm{e}^{-2bm}\Bigg) < +\infty.$$

Since $\beta_m/\beta_n$ remains bounded from above and below around the diagonal, the matrix $\tilde{A}$ behaves like $A$ near the diagonal. More precisely, if $I, J$ are blocks of integers such that $(m,n) \in I \times J$ implies that $n/\lambda^2 \le m \le \lambda^2 n$, then, with obvious notation (e.g., $P_I$ is the orthogonal projection on $\mathrm{span}(e_n, n \in I)$), the slow oscillation of $\beta$ gives, for some $C > 0$,

$$\Bigg|\sum_{(m,n)\in I\times J}a_{m,n}x_n\overline{y_m}\sqrt{\frac{\beta_m}{\beta_n}}\Bigg| \le \|A\|\Bigg(\sum_{(m,n)\in I\times J}|x_n|^2|y_m|^2\frac{\beta_m}{\beta_n}\Bigg)^{1/2} \le C^{1/2}\|A\|\|P_J x\|\|P_I y\|.$$

For $k = 0, 1, 2, \ldots$, let $J_k = [\lambda^k, \lambda^{k+1}[$ and, for $k = 1, 2, \ldots$, we define $I_k = [\lambda^{k-1}, \lambda^{k+2}[$. We also define $I_0 = [0, \lambda^2[$.

We define the matrix $R$ to have entries

$$r_{m,n} = \begin{cases}\sqrt{\beta_m/\beta_n}\,a_{m,n} & \text{if } (m,n) \in \bigcup_{k=0}^{\infty}(I_k \times J_k),\\ 0 & \text{elsewhere.}\end{cases}$$

Let $H_k$ be the subspace of the sequences $(x_n)_{n\ge 0}$ in $\ell_2$ such that $x_n = 0$ for $n \notin I_k$, i.e.,

$$H_k = \mathrm{span}\{e_n : n \in I_k\} \quad \text{and} \quad \widetilde{H}_k = \mathrm{span}\{e_n : n \in J_k\}.$$

Let $P_k$ be (the matrix of) the orthogonal projection of $\ell_2$ with range $H_k$ and $Q_k$ that with range $\widetilde{H}_k$. Then $R_k = P_k A Q_k$ is the matrix with entries $a_{m,n}$ when $(m,n) \in I_k \times J_k$ and 0 elsewhere. By the above discussion, we have

$$|(R_k x \mid y)| \le C^{1/2}\|A\|\|Q_k x\|\|P_k y\|.$$

We point out that, for every $y \in \ell^2$, we have $\sum \|P_k y\|^2 \le 3\|y\|^2$ since each integer belongs to at most three intervals $I_k$.

In the same way, for every $x \in \ell^2$, we have $\sum \|Q_k x\|^2 \le \|x\|^2$ since the subspaces $\widetilde{H}_k$ are orthogonal. Summing up over $k$, we get the boundedness of $R = \sum_{k=0}^{\infty} R_k$.

Now let us check when the entries of $R$ do not coincide with the entries of $\tilde{A}$. Actually, it happens when $(m, n)$ does not belong to the union of the $I_k \times J_k$. When $n \geq 1$, it means that $n$ belongs to some $J_p$ but $m \notin I_p$: either $m < \lambda^{p-1}$ or $m \geq \lambda^{p+2}$, and hence either $m/n < \lambda^{-1}$ or $m/n > \lambda$. Therefore the nonzero entries $(m, n)$ of $\tilde{A} - R$ satisfy either $n > \lambda m$ or $m > \lambda n$.

That ends the proof since we have seen at the beginning that $\tilde{A} - R$ is Hilbert–Schmidt. □

**Remark.** The proof shows that, instead of (1) and (2), it is enough to have

$$\sum_{m<C_1 n} n^{\alpha+1}|a_{m,n}|^2 < \infty \quad \text{and} \quad \sum_{m>C_2 n} m^{\alpha}|a_{m,n}|^2 < \infty.$$

Moreover, the proof also shows that, when $\beta$ is slowly oscillating, if we set

$$E = \{(m, n) : C_1 n \leq m \leq C_2 n\} \quad \text{for some } C_1, C_2 > 0,$$

then the matrix $(\sqrt{\beta_m/\beta_n}\,\mathbb{1}_E(m, n))$ is a Schur multiplier over *all* the bounded matrices, while Kacnelson's theorem (Theorem 3.5) says that, if $\gamma = (\gamma_n)$ is nonincreasing, the matrix $(\gamma_m/\gamma_n)$ is a Schur multiplier of all bounded *lower-triangular* matrices.

*Proof of Theorem 4.6.* Thanks to Lemma 4.7, the hypotheses of Lemma 4.8 are fulfilled by the matrix whose entries are $a_{m,n} = \widehat{\varphi^n}(m)$. It follows (with the notation of Lemma 4.8) that $\tilde{A}$ is bounded on $\ell^2$, which means exactly that $T_a$ is bounded on $H^2(\beta)$. □

**4.3.** ***Necessary condition.*** The main theorem of this section is the following.

**Theorem 4.9.** *If the composition operator $C_{T_a}$ is bounded on $H^2(\beta)$ for some $a \in \mathbb{D} \setminus \{0\}$, then $\beta$ is slowly oscillating.*

Let us give a corollary of this result.

**Proposition 4.10.** *For the weight $\beta$ constructed in the proof of Theorem 3.7, no automorphism $T_a$ with $0 < a < 1$ can be bounded.*

*Proof.* Indeed, it is clear that $\beta$ is not slowly oscillating, since

$$\frac{\beta_{(k+1)!-1}}{\beta_{(k+1)!}} = \frac{1}{k+1} \xrightarrow[k\to\infty]{} 0. \qquad \square$$

To prove Theorem 4.9, we need estimates on the Taylor coefficients of $T_a^n$. Actually, the Taylor coefficients of $T_a^n$ are the Fourier coefficients of $x \in \mathbb{R} \mapsto T_a^n(e^{ix})$, and we shall denote them with the same notation $\widehat{T_a^n}$. Sharp such estimates are given in [Szehr and Zarouf 2020; 2021], and we thank R. Zarouf for interesting information on this subject (see also [Borichev et al. 2024]). Our method, using stationary phase and the van der Corput lemma, is a variant of that used in [Szehr and Zarouf 2020; 2021] and goes back at least to [Girard 1973]. However, we need minorizations of $|\widehat{T_a^n}(m)|$ when $m$ is close to $n$, and Szehr and Zarouf's estimates show that this quantity oscillates and, for individual $a$, can be too small for our purpose, so we cannot use them and have to prove an estimate *in mean* for $a$ in some subinterval of (0, 1).

We begin with a standard fact, which we give with its proof for the convenience of the reader.

**Lemma 4.11.** *Let $a \in (0, 1)$, and let*

$$P_{-a}(x) = \frac{1-a^2}{1+2a\cos x + a^2}$$

*be the Poisson kernel at the point $-a$. Then, for all $x \in [-\pi, \pi]$,*

$$T_a(e^{ix}) = \exp[i V_a(x)], \tag{4-5}$$

*where*

$$V_a(x) = \int_0^x P_{-a}(t)\, dt. \tag{4-6}$$

*Proof.* For $t \in [-\pi, \pi]$, write

$$\psi(t) := \frac{e^{it}+a}{1+ae^{it}} = \exp(i v(t)),$$

with $v$ a real-valued, $\mathcal{C}^1$-function on $[-\pi, \pi]$ such that $v(0) = 0$. This is possible since $|\psi(e^{it})| = 1$ and $\psi(0) = 1$. Differentiating both sides with respect to $t$, we get

$$i e^{it} \frac{1-a^2}{(1+ae^{it})^2} = i v'(t) \frac{e^{it}+a}{1+ae^{it}}.$$

This implies

$$v'(t) = \frac{1-a^2}{|1+ae^{it}|^2} = P_{-a}(t),$$

and the result follows since $v(0) = 0 = V_a(0)$. □

Let us note that, with $V_a$ the function of Lemma 4.11, the Fourier formulas give, since $\widehat{T_a^n}(m)$ is real or since $nV_a(x) - mx$ is odd,

$$2\pi \widehat{T_a^n}(m) = \int_{-\pi}^{\pi} \exp(i[nV_a(x) - mx])\, dx = 2\,\Re\mathrm{e}\, I_{m,n}, \tag{4-7}$$

where

$$I_{m,n} = \int_0^{\pi} \exp i[nV_a(x) - mx]\, dx. \tag{4-8}$$

Now the main ingredient for proving Theorem 4.9 is the following.

**Proposition 4.12.** *Let $I := \left[\frac{1}{2}, \frac{2}{3}\right]$. There exist constants $\alpha > 1$, e.g., $\alpha = \frac{5}{4}$, and $\delta \in \left(0, \frac{1}{2}\right)$ such that, for $n$ large enough ($n \geq n_0$), we have*

$$\int_I |\widehat{T_a^n}(m)|^2\, da \geq \frac{\delta}{n} \quad \textit{for all } m \in [\alpha^{-1}n, \alpha n]. \tag{4-9}$$

*Proof.* We will set once and for all

$$q = \frac{m}{n}, \tag{4-10}$$

so that $\alpha^{-1} \leq q \leq \alpha$ where $\alpha = \frac{5}{4}$ (say). We will only consider pairs $(a, q)$ satisfying

$$a \in I = \left[\tfrac{1}{2}, \tfrac{2}{3}\right], \quad q \in J := \left[\tfrac{4}{5}, \tfrac{5}{4}\right]. \tag{4-11}$$

Such pairs will be called *admissible*.

With this notation, we set, for $0 \le x \le \pi$,

$$F_q(x) = V_a(x) - \frac{m}{n}x = \int_0^x P_{-a}(t)\,dt - qx, \tag{4-12}$$

where $P_{-a}$ is the Poisson kernel at $-a$. We have

$$F_q'(x) = \frac{(1-a^2)}{1+2a\cos x + a^2} - q,$$

and the unique (if it exists) critical point $x_q = x_q(a)$ of $F_q$ in $[0, \pi]$ is given by $P_{-a}(x_q) = q$, that is,

$$\cos x_q = \frac{1}{q}\frac{1-a^2}{2a} - \frac{1+a^2}{2a} =: h_q(a). \tag{4-13}$$

We now proceed through a series of simple lemmas and begin by estimates on $h_q$ and $x_q$.

**Lemma 4.13.** *There are positive constants $C > 1$ and $\delta \in \left(0, \frac{1}{2}\right)$ such that, for every admissible pair $(a, q)$, we have*

$$|h_q(a)| \le 1 - \delta \quad \textit{and} \quad |h_q'(a)| \le C, \tag{4-14}$$

*so there is one critical point $x_q(a)$ satisfying*

$$\delta \le x_q(a) \le \pi - \delta \quad \textit{and} \quad \sin x_q(a) \ge \delta; \tag{4-15}$$

*moreover,*

$$|x_q'(a)| \le C \quad \textit{and} \quad \delta \le |P_{-a}'(x_q)| \le C. \tag{4-16}$$

*Proof.* We have

$$h_q(a) = \left(\frac{1}{q}\frac{1-a^2}{2a}\right) + \left(-\frac{1+a^2}{2a}\right) =: u(a) + v(a),$$

with $u$ and $v$ respectively decreasing and increasing on $[0, 1]$ and with $v \le 0$, so that we have, for $q \in J$,

$$h_q(a) \le u\left(\frac{1}{2}\right) = \frac{3}{4q} \le \frac{15}{16}.$$

Similarly:

$$h_q(a) \ge u\left(\frac{2}{3}\right) + v\left(\frac{1}{2}\right) = \frac{5}{12q} - \frac{5}{4} \ge \frac{1}{3} - \frac{5}{4} = -\frac{11}{12}.$$

Next, $2h_q'(a) = (1 - 1/q)1/a^2 - (1 + 1/q)$; hence $|h_q'(a)| \le C$. So, writing $x_q = x_q(a) = \arccos h_q(a)$, we get, with another constant $C > 0$,

$$|x_q'(a)| = \frac{|h_q'(a)|}{\sqrt{1 - h_q(a)^2}} \le C$$

since $h_q(a)^2 \le 1 - \delta$. Finally, $\frac{1}{9} \le (1-a)^2 \le 1 + 2a\cos x_q + a^2 \le 4$, and since

$$P_{-a}'(x_q) = \frac{2a(1-a^2)\sin x_q}{(1+2a\cos x_q + a^2)^2},$$

we get the final estimates, ending the proof. □

Back to Proposition 4.12.

We saw in (4-7) that the value of $a_{m,n} := \widehat{T_a^n}(m)$ is given by the formula

$$a_{m,n} = \frac{1}{\pi}\,\mathfrak{Re}\, I_{m,n}. \tag{4-17}$$

We have the following estimate, whose proof is postponed (recall that $q = m/n$ and $x_q = x_q(a)$).

**Proposition 4.14.** *We have*

$$I_{m,n} = \sqrt{2\pi}\, n^{-1/2} \frac{\mathrm{e}^{i[nF_q(x_q)+\pi/4]}}{\sqrt{|P'_{-a}(x_q)|}} + O(n^{-3/5}), \tag{4-18}$$

*where the $O$ only depends on $a$ and so is absolute as long as $(a, q)$ is admissible.*

Note that $\frac{3}{5} > \frac{1}{2}$. We hence have

$$a_{m,n} = \sqrt{\frac{2}{\pi}}\, n^{-1/2} \frac{\cos[\pi/4 + nF_q(x_q)]}{\sqrt{|P'_{-a}(x_q)|}} + O(n^{-3/5}). \tag{4-19}$$

It will be convenient to introduce $\varphi_q(a)$, which we do by setting

$$F_q(x_q(a)) = \varphi_q(a). \tag{4-20}$$

Then, since $\frac{1}{2} + \frac{3}{5} = \frac{11}{10}$, $\cos^2(\pi/4 + x) = \frac{1}{2}(1 - \sin 2x)$ and $|P'_{-a}(x_q)| \geq \delta$ by Lemma 4.13, we have

$$a_{m,n}^2 = \frac{1}{\pi} n^{-1} \frac{1 - \sin[2n\varphi_q(a)]}{|P'_{-a}(x_q)|} + O(n^{-11/10})$$

implying, since $|P'_{-a}(x_q)| \leq C$ by Lemma 4.13 (again for $(a, q)$ admissible) and changing $\delta$,

$$a_{m,n}^2 \geq \delta n^{-1}(1 - \sin[2n\varphi_q(a)]) + O(n^{-11/10}). \tag{4-21}$$

We will also need estimates on the derivatives of $\varphi_q(a)$.

**Lemma 4.15.** *If $(a, q)$ is admissible, then $\varphi_q$ decreases on $I$ and, moreover,*

(1) $|\varphi'_q(a)| \geq \delta$,

(2) $|\varphi''_q(a)| \leq C$.

*Proof.* Note, in passing, that, with $x = x_q(a) \in [0, \pi]$ (thanks to (4-12)),

$$\varphi_q(a) = \int_0^x [P_{-a}(t) - P_{-a}(x)]\, dt \leq 0$$

since the integrand is negative. Next, if $f$ and $g$ are real $\mathcal{C}^1$-functions and

$$\Phi(a) = \int_0^{f(a)} g(a, t)\, dt,$$

the chain rule gives

$$\Phi'(a) = f'(a) g(a, f(a)) + \int_0^{f(a)} \frac{\partial g}{\partial a}(a, t)\, dt.$$

With $g(a, t) = P_{-a}(t)$ and $f(a) = x_q(a)$, we get, remembering that $x_q(a)$ is critical for $F_q$,

$$\varphi'_q(a) = [P_{-a}(x_q(a)) - q] x'_q(a) + \int_0^{x_q(a)} \frac{\partial P_{-a}}{\partial a}(a, t)\, dt = \int_0^{x_q(a)} \frac{\partial P_{-a}}{\partial a}(a, t)\, dt.$$

But $P_{-a}(t) = 1 + 2\sum_{k=1}^{\infty}(-a)^k \cos kt$, so we have

$$\varphi_q'(a) = \int_0^{x_q(a)} \Big(-2\sum_{k=1}^{\infty} k(-a)^{k-1}\cos kt\Big)\,dt = \frac{2}{a}\sum_{k=0}^{\infty}(-a)^k \sin[kx_q(a)],$$

that is,

$$\varphi_q'(a) = \frac{2}{a}\,\Im\mathrm{m}\,\frac{1}{1+a\mathrm{e}^{ix_q(a)}} = \frac{-2\sin x_q(a)}{1+2a\cos x_q(a)+a^2} < 0. \tag{4-22}$$

Now, (4-15) gives (1).

Since $|x_q'(a)| \le C$ by Lemma 4.13, the chain rule and (4-22) clearly give the uniform boundedness of $|\varphi_q''(a)|$ when $(a, q)$ is admissible, and this ends the proof. □

Lemmas 4.13 and 4.15 will now be exploited through a simple variant of the van der Corput inequalities.

**Lemma 4.16.** *Let $f : [A, B] \to \mathbb{R}$, with $A < B$, be a $\mathcal{C}^2$-function satisfying $|f'| \ge \delta$ and $|f''| \le C$, and let us put $M = \int_A^B \mathrm{e}^{inf(x)}\,dx$. Then*

$$|M| \le \frac{2}{n\delta} + \frac{C(B-A)}{n\delta^2}.$$

*Proof.* Write

$$\mathrm{e}^{inf} = \frac{(\mathrm{e}^{inf})'}{inf'}$$

and integrate by parts to get

$$M = \left[\frac{\mathrm{e}^{inf}}{inf'}\right]_A^B - \frac{i}{n}\int_A^B \mathrm{e}^{inf(x)}\frac{f''(x)}{[f'(x)]^2}\,dx =: M_1 + M_2,$$

with $|M_1| \le 2/(n\delta)$ and $|M_2| \le ((B-A)/n)\cdot C/\delta^2$. □

*End of proof of Proposition 4.12.* The preceding lemma can be applied with $A = \frac{1}{2}$, $B = \frac{2}{3}$, $f = \varphi_q$ and $n$ changed into $2n$, since Lemma 4.15 shows that this $f$ meets the assumptions of Lemma 4.16. This gives us, uniformly with respect to $(a, q)$ admissible,

$$\left|\int_I \sin[2n\varphi_q(a)]\,da\right| \le \left|\int_I \mathrm{e}^{2in\varphi_q(a)}\,da\right| \le \frac{C}{n}. \tag{4-23}$$

Now, integrating (4-21) on $I$ and using (4-23) gives, for some numerical $\delta \in \big(0, \frac{1}{2}\big)$,

$$\int_I |\widehat{T_a^n}(m)|^2\,da \ge \delta n^{-1} + O(n^{-2}) + O(n^{-11/10}) \ge \tfrac{1}{2}\delta n^{-1}$$

for $n \ge n_0$ and $\alpha^{-1} \le m/n \le \alpha$ (recall that $a_{m,n} = \widehat{T_a^n}(m)$). This ends the proof of Proposition 4.12. □

*Proof of Theorem 4.9.* By Proposition 4.2, $C_{T_a}$ is bounded for all $a \in \mathbb{D}$, and, thanks to Lemma 4.4,

$$K := \sup_{1/2 \le a \le 2/3} \|C_{T_a}\| < +\infty.$$

Matricially, this can be written, for all $a \in \big(\frac{1}{2}, \frac{2}{3}\big)$,

$$\left\|\left(\widehat{T_a^n}(m)\sqrt{\frac{\beta_m}{\beta_n}}\right)_{m,n}\right\| \le K.$$

In particular, for every $n \geq 1$, we have, considering the columns and rows of the previous matrix,

$$\sum_{m=1}^{\infty} |\widehat{T_a^n}(m)|^2 \frac{\beta_m}{\beta_n} \leq K^2, \quad \text{i.e.,} \sum_{m=1}^{\infty} |\widehat{T_a^n}(m)|^2 \beta_m \leq K^2 \beta_n,$$

and, for every $m \geq 1$,

$$\sum_{n=1}^{\infty} |\widehat{T_a^n}(m)|^2 \frac{\beta_m}{\beta_n} \leq K^2, \quad \text{i.e.,} \sum_{n=1}^{\infty} |\widehat{T_a^n}(m)|^2 \frac{1}{\beta_n} \leq \frac{K^2}{\beta_m}.$$

In particular, for every $n \geq 1$,

$$\sum_{(4/5)n \leq j \leq (5/4)n} |\widehat{T_a^n}(j)|^2 \beta_j \leq K^2 \beta_n \tag{4-24}$$

and, for every $m \geq 1$,

$$\sum_{(4/5)m \leq k \leq (5/4)m} |\widehat{T_a^k}(m)|^2 \frac{1}{\beta_k} \leq \frac{K^2}{\beta_m}. \tag{4-25}$$

Integrating on $a \in \left(\frac{1}{2}, \frac{2}{3}\right)$ and using Proposition 4.12, we get, from (4-24), for $n$ large enough,

$$\frac{\delta}{n} \sum_{(4/5)n \leq j \leq (5/4)n} \beta_j \leq \frac{K^2}{6} \beta_n \tag{4-26}$$

and, from (4-25), for $m$ large enough, we have both

$$\frac{\delta}{m} \sum_{(4/5)m \leq k \leq m} \frac{1}{\beta_k} \leq \frac{5K^2}{24} \frac{1}{\beta_m} \tag{4-27}$$

and

$$\frac{\delta}{m} \sum_{m \leq k \leq (5/4)m} \frac{1}{\beta_k} \leq \frac{5K^2}{24} \frac{1}{\beta_m}. \tag{4-28}$$

Since the harmonic mean (over the sets of integers $\left[\frac{4}{5}m, m\right]$ and $\left[m, \frac{5}{4}m\right]$, which have cardinality $\approx n \approx m$) is less than the arithmetical mean, we obtain, from (4-27) and (4-28), both

$$\beta_m \leq \frac{125}{24\delta} \frac{K^2}{m} \sum_{(4/5)m \leq k \leq m} \beta_k \tag{4-29}$$

and

$$\beta_m \leq \frac{10}{3\delta} \frac{K^2}{m} \sum_{m \leq k \leq (5/4)m} \beta_k. \tag{4-30}$$

Now assume that $n \leq m \leq \frac{5}{4}n$. From (4-29), we have

$$\beta_m \lesssim \frac{1}{m} \sum_{(4/5)m \leq k \leq m} \beta_k \lesssim \frac{1}{n} \sum_{(4/5)n \leq k \leq (5/4)n} \beta_k \lesssim \beta_n$$

thanks to (4-26). From (4-30) and (4-26), we treat the case $\frac{4}{5}n \leq m \leq n$ in the same way. We conclude that, for some constant $c > 0$, we have, for $n$ and $m$ large enough satisfying $\frac{4}{5}n \leq m \leq \frac{5}{4}n$,

$$\beta_m \leq c\beta_n, \tag{4-31}$$

which means that $\beta$ is slowly oscillating. □

*Proof of Proposition 4.14.* We will use a variant of [Titchmarsh 1986, Lemma 4.6, p. 72] on the van der Corput's version of the stationary phase method. A careful reading of the proof in [Titchmarsh 1986, p. 72] gives the version below, which only needs local estimates on the second derivative $F''$, as occurs in our situation. For the sake of completeness, we will give a proof, postponed to the Appendix.

**Proposition 4.17** (stationary phase). *Let $F$ be a real function with continuous derivatives up to the third order on the interval $[A, B]$ and $F'' > 0$ throughout $]A, B[$. Assume that there is a (unique) point $c$ in $]A, B[$ such that $F'(c) = 0$ and that, for some positive numbers $\lambda_2$, $\lambda_3$, and $\eta$, the following assertions hold*:

(1) $[c-\eta, c+\eta] \subseteq [A, B]$,

(2) $F''(x) \geq \lambda_2$ *for all* $x \in [c-\eta, c+\eta]$,

(3) $|F'''(x)| \leq \lambda_3$ *for all* $x \in [A, B]$.

*Then*

$$\int_A^B e^{iF(x)}\,dx = \sqrt{2\pi}\frac{e^{i[F(c)+\pi/4]}}{F''(c)^{1/2}} + O\Big(\frac{1}{\eta\lambda_2} + \eta^4\lambda_3\Big), \tag{4-32}$$

*where the $O$ involves an absolute constant.*

We will show that Proposition 4.17 is applicable with $F = nF_q$ and

$$[A, B] = [0, \pi], \quad c = x_q, \quad \lambda_2 = \kappa_0 n, \quad \lambda_3 = C_0 n, \quad \eta = (\lambda_2\lambda_3)^{-1/5}.$$

The parameter $\eta$ is chosen to make both error terms in Proposition 4.17 equal: $(\eta\lambda_2)^{-1} = \eta^4\lambda_3$; so

$$\eta = \kappa n^{-2/5}$$

and

$$\frac{1}{\eta\lambda_2} + \eta^4\lambda_3 = \tilde{\kappa} n^{-3/5} = O(n^{-3/5}) \tag{4-33}$$

(with $\kappa = (\kappa_0 C_0)^{-1/5}$ and $\tilde{\kappa} = 2/\kappa_0\kappa$).

The slight technical difficulty encountered here is that $F_q''(x)$ vanishes at 0 and $\pi$. But Proposition 4.17 covers this case. We have

$$F''(x) = nF_q''(x) = nP_{-a}'(x) = 2a(1-a^2)\frac{\sin x}{(1+2a\cos x + a^2)^2}n,$$

and there are some positive (and absolute) constants $\kappa_0$ and $\sigma$ such that

$$F''(x) \geq \kappa_0 n = \lambda_2 \quad \text{for } x \in [\sigma, \pi-\sigma]. \tag{4-34}$$

Now (for $n$ large enough), we have $[x_q - \eta, x_q + \eta] \subseteq [\sigma, \pi - \sigma]$. Hence assumptions (1) and (2) of Proposition 4.17 are satisfied.

Finally, since $F(x) = nF_q(x) = n[V_a(x) - qx]$ and $F''' = nF_q''' = nV_a''' = nP_{-a}''$, we have, for all $x \in [0, \pi]$ and $(a, q)$ admissible,

$$|F'''(x)| \leq C_0 n = \lambda_3,$$

where $C_0$ is absolute and assertion (3) of Proposition 4.17 holds.

With (4-33) this ends the proof of (4-18), once we note that $nV_a''(x_q) = F''(x_q)$. □

## 5. Boundedness of all composition operators

In this section, we characterize all the sequences $\beta$ for which all composition operators are bounded on $H^2(\beta)$. The main remaining step is the following theorem.

**Theorem 5.1.** *Assume that all composition operators $C_\varphi$ are bounded on $H^2(\beta)$. Then $\beta$ is essentially decreasing.*

As an immediate consequence, we obtain Theorem 1.2.

*Proof of Theorem 1.2.* Assume that $\beta$ is essentially decreasing and slowly oscillating. All composition operators $C_\psi$ with $\psi(0) = 0$ are bounded on $H^2(\beta)$ (see the introduction or Proposition 3.2). Since $\beta$ is slowly oscillating, all the composition operators $C_{T_a}$, with $a \in \mathbb{D}$, are bounded thanks to Theorem 4.6. Now it is very classical that we can get the boundedness of every composition operators. Indeed given a symbol $\varphi$, the symbol $\psi = T_a \circ \varphi$ fixes the origin for $a = -\varphi(0)$. Since $C_\varphi = C_\psi \circ C_{T_{-a}}$, the conclusion follows.

Assume that all composition operators are bounded on $H^2(\beta)$; in particular, the $C_{T_a}$ ones are bounded on $H^2(\beta)$, and $\beta$ is slowly oscillating, thanks to Theorem 4.9. It also follows from Theorem 5.1 that $\beta$ is essentially decreasing. □

We will use the following elementary, but crucial, lemma.

**Lemma 5.2.** *Let $u$ be a function analytic in an open neighborhood $\Omega$ of $\overline{\mathbb{D}}$. Then, for every $\varepsilon > 0$, there exists an integer $N \geq 1$ such that*

$$\sum_{j=Np}^{\infty} |\widehat{u^p}(j)|^2 \leq \varepsilon \quad \textit{for all } p \geq 1. \tag{5-1}$$

*Proof.* From Lemma 4.7, we know that there exist some integer $\lambda > 1$ and a constant $b > 0$ such that $|\widehat{u^p}(j)| \leq \mathrm{e}^{-bj}$ when $j \geq \lambda p$. Therefore, for any $N \geq \lambda$, we have

$$\sum_{j=Np}^{\infty} |\widehat{u^p}(j)|^2 \leq (1-\mathrm{e}^{-2b})^{-1}\mathrm{e}^{-2bNp} \leq (1-\mathrm{e}^{-2b})^{-1}\mathrm{e}^{-2bN} \leq \varepsilon$$

as soon as $N$ is chosen large enough. □

*Proof of Theorem 5.1.* Thanks to Theorem 4.9, we know that $\beta$ is slowly oscillating.

Now, assume that the sequence $\beta$ is not essentially decreasing.

We are going to construct an analytic function $\varphi\colon \mathbb{D} \to \mathbb{D}$ such that the composition operator $C_\varphi$ is not bounded on $H^2(\beta)$. This function $\varphi$ will be a Blaschke product of the form

$$\varphi(z) = \prod_{k=1}^{\infty} T_{a_k}(z^{n_k}) = \prod_{k=1}^{\infty} \frac{z^{n_k} + a_k}{1 + a_k z^{n_k}}$$

for a sequence of numbers $a_k \in (0, 1)$ such that $\sum_{k\geq 1}(1-a_k) < +\infty$ and a sequence of positive integers $n_k$ increasing to infinity.

Observe that $\varphi$ will be indeed a convergent Blaschke product, with $n_k$ zeroes of modulus $a_k^{1/n_k}$, $k = 1, 2, \dots$, because, for $T_a(z) = (z + a)/(1 + az)$, with $0 < a < 1$, we have

$$|T_{a_k}(z^{n_k}) - 1| \le \frac{2(1 - a_k)}{1 - |z|}$$

and, setting $a_k = \mathrm{e}^{-\varepsilon_k}$, we get

$$\sum_k n_k(1 - a_k^{1/n_k}) \le \sum_k n_k(\varepsilon_k/n_k) = \sum_k \varepsilon_k < +\infty.$$

These sequences will be constructed by induction, together with another sequence of integers $(m_k)_{k\ge1}$.

Since $\beta$ is not essentially decreasing, there exist integers $n_1 > m_1 \ge 4$ such that $\beta_{n_1} \ge 2\beta_{m_1}$. We start with

$$a_1 = 1 - \frac{1}{m_1} \ge \frac{3}{4}.$$

Using Lemma 5.2 with $u = T_{a_1}$, we get $N_0 \ge 1$ such that

$$\sum_{j=N_0 m}^{\infty} |\widehat{T_{a_1}^m}(j)|^2 \le 2^{-15} \quad \text{for all } m \ge 1.$$

Assume now that we have constructed increasing sequences of integers

$$m_1, m_2, \dots, m_k, \quad n_1, n_2, \dots, n_k, \quad N_0, N_1, \dots, N_{k-1}$$

such that, for $1 \le l \le k - 1$, we have

$$m_{l+1} \ge 4m_l \quad \text{and} \quad n_{l+1} \ge 4n_l$$

and, for $1 \le l \le k$,

$$n_l \ge N_{l-1} m_l \quad \text{and} \quad \beta_{n_l} \ge 2^l \beta_{m_l}$$

and

$$\sum_{j=N_{l-1}m_l}^{\infty} |\widehat{\varphi_l^m}(j)|^2 \le 2^{-15},$$

where

$$a_l = 1 - \frac{1}{m_l} \quad \text{and} \quad \varphi_l(z) = T_{a_l}(z^{n_l}).$$

We then apply Lemma 5.2 again to the function $u = u_k = \varphi_1 \cdots \varphi_k$. We get $N_k > N_{k-1}$ such that

$$\sum_{j=N_k m}^{\infty} |\widehat{u_k^m}(j)|^2 \le 2^{-15} \quad \text{for all } m \ge 1. \tag{5-2}$$

Since $\beta$ is not essentially decreasing but is slowly oscillating, there exist $m_{k+1} \ge 4m_k$ and $n_{k+1} \ge 4n_k$ such that

$$n_{k+1} \ge N_k m_{k+1} \quad \text{and} \quad \beta_{n_{k+1}} \ge 2^{k+1} \beta_{m_{k+1}}.$$

We set

$$a_{k+1} = 1 - \frac{1}{m_{k+1}} \quad \text{and} \quad \varphi_{k+1}(z) = T_{a_{k+1}}(z^{n_{k+1}}).$$

This ends the induction.

It remains to check that

$$\sum_{k=1}^{\infty}(1-a_k)=\sum_{k=1}^{\infty}\frac{1}{m_k}\le\sum_{k=1}^{\infty}4^{-k}=\frac{1}{3}<+\infty$$

to get that the infinite product $\varphi=\prod_{k\ge 1}\varphi_k$ converges uniformly on compact subsets of $\mathbb{D}$.

To show that the composition operator $C_\varphi$ is not bounded on $H^2(\beta)$, it suffices to show that, for some constant $c_1>0$, we have, for all $k\ge 2$,

$$\sum_{j=n_k}^{2n_k}|\widehat{\varphi^{m_k}}(j)|^2\ge c_1. \tag{5-3}$$

Indeed, since $\beta$ is slowly oscillating, there is a positive constant $\delta<1$ such that

$$\beta_j\ge\delta\beta_{n_k}\quad\text{for } j=n_k,n_k+1,\dots,2n_k.$$

Then, if we set $e_k(z)=z^{m_k}$, we have, since $C_\varphi(e_k)=\varphi^{m_k}$,

$$\frac{\|C_\varphi(e_k)\|^2_{H^2(\beta)}}{\|e_k\|^2_{H^2(\beta)}}\ge\frac{\sum_{j=n_k}^{2n_k}|\widehat{\varphi^{m_k}}(j)|^2\beta_j}{\beta_{m_k}}\ge\frac{c_1\delta\beta_{n_k}}{\beta_{m_k}}\ge 2^k c_1\delta\xrightarrow[k\to\infty]{}+\infty,$$

and so $C_\varphi$ is not bounded on $H^2(\beta)$.

We now have to show (5-3). Let us agree to write formally, for an analytic function $f(z)=\sum_{k=0}^{\infty}f_k z^k$ and an arbitrary positive integer $p$,

$$f(z)=\sum_{k=0}^{p}f_k z^k+O(z^{p+1}).$$

For that, we set

$$G_k(z)=\prod_{l=k+1}^{\infty}\varphi_l(z)=\prod_{l=k+1}^{\infty}a_l+O(z^{n_{k+1}}).$$

We have, for $k\ge 2$,

$$\varphi(z)=v_k(z)\varphi_k(z)G_k(z),$$

where $v_k=\varphi_1\cdots\varphi_{k-1}$.

Remark now that, for $0<a<1$, we have

$$T_a(z)=a+(1-a^2)z+O(z^2),$$

so

$$\varphi_k(z)=T_{a_k}(z^{n_k})=a_k+(1-a_k^2)z^{n_k}+O(z^{2n_k}).$$

Then

$$[G_k(z)]^{m_k}=\Bigl(\prod_{l=k+1}^{\infty}a_l\Bigr)^{m_k}+O(z^{n_{k+1}}) \tag{5-4}$$

and

$$[\varphi_k(z)]^{m_k}=a_k^{m_k}+(1-a_k^2)m_k a_k^{m_k-1}z^{n_k}+O(z^{2n_k}). \tag{5-5}$$

But

$$a_k^{m_k-1} = \Big(1 - \frac{1}{m_k}\Big)^{m_k-1} \geq \mathrm{e}^{-1} := c_2 \tag{5-6}$$

and

$$(1-a_k^2)m_k a_k^{m_k-1} \geq (1-a_k)m_k a_k^{m_k-1} \geq c_2. \tag{5-7}$$

Moreover, since $1-x \geq \mathrm{e}^{-2x}$ for $0 \leq x \leq \frac{1}{2}$, we have

$$\Big(\prod_{l=k+1}^{\infty} a_l\Big)^{m_k} \geq \exp\Big(-2\Big(\sum_{l=k+1}^{\infty} \frac{1}{m_l}\Big)m_k\Big) \geq \exp\Big(-2\sum_{l=1}^{\infty} 4^{-l}\Big) = \exp\big(-\tfrac{2}{3}\big) := c_3. \tag{5-8}$$

Afterwards, by (5-2), we have

$$\sum_{j=N_{k-1}m_k}^{\infty} |\widehat{v_k^{m_k}}(j)|^2 \leq 2^{-15}. \tag{5-9}$$

Set $v_k^{m_k} = g_1 + g_2$, with

$$\begin{cases} g_1(z) = \sum_{j=0}^{N_{k-1}m_k} \widehat{v_k^{m_k}}(j) z^j, \\ g_2(z) = \sum_{j>N_{k-1}m_k} \widehat{v_k^{m_k}}(j) z^j. \end{cases}$$

By (5-9), we have, with $\|\cdot\|_2 = \|\cdot\|_{L^2(\mathbb{T})}$,

$$\|g_2\|_2^2 = \sum_{j>N_{k-1}m_k} |\widehat{v_k^{m_k}}(j)|^2 \leq 2^{-15}.$$

Besides, since $\varphi_k$ is inner as a product of inner functions, we have $|v_k(z)| = 1$ for all $z \in \mathbb{T}$, so

$$\|g_1\|_2^2 = \|v_k\|_2^2 - \|g_2\|_2^2 \geq 1 - 2^{-15}.$$

Now, $\varphi^{m_k} = v_k^{m_k} \varphi_k^{m_k} G_k^{m_k} = F_1 + F_2$, with

$$F_1 = g_1 \varphi_k^{m_k} G_k^{m_k} \quad \text{and} \quad F_2 = g_2 \varphi_k^{m_k} G_k^{m_k}.$$

Using (5-4), (5-5), (5-7) and (5-8), we get

$$\sum_{j=n_k}^{n_k+N_{k-1}m_k} |\widehat{F_1}(j)|^2 = \Big(\prod_{l=k+1}^{\infty} a_l\Big)^{2m_k} [(1-a_k^2)m_k a_k^{m_k-1}]^2 \sum_{j=0}^{N_{k-1}m_k} |\widehat{g_1}(j)|^2 \geq (1-2^{-15})c_2^2 c_3^2.$$

As

$$\|F_2\|_2^2 \leq \|g_2\|_2^2 \|\varphi_k^{m_k}\|_\infty^2 \|G_k^{m_k}\|_\infty^2 \leq 2^{-15},$$

we get, using the inequality $|a+b|^2 \geq \frac{1}{2}|a|^2 - |b|^2$,

$$\begin{aligned} \sum_{j=n_k}^{2n_k} |\widehat{\varphi^{m_k}}(j)|^2 &\geq \sum_{j=n_k}^{n_k+N_{k-1}m_k} |\widehat{F_1}(j) + \widehat{F_2}(j)|^2 \geq \frac{1}{2}\sum_{j=n_k}^{n_k+N_{k-1}m_k} |\widehat{F_1}(j)|^2 - \sum_{j=n_k}^{n_k+N_{k-1}m_k} |\widehat{F_2}(j)|^2 \\ &\geq \frac{1}{2}(1-2^{-15})c_2^2c_3^2 - 2^{-15} = \frac{1}{2}(1-2^{-15})\mathrm{e}^{-10/3} - 2^{-15} \geq 2^{-9} - 2^{-15} > 0. \qquad \square \end{aligned}$$

## 6. Some results on multipliers

In this section, we give some results on the multipliers on $H^2(\beta)$, which show how the different notions of regularity for $\beta$ come into play.

The set $\mathcal{M}(H^2(\beta))$ of multipliers of $H^2(\beta)$ is by definition the vector space of functions $h$ analytic on $\mathbb{D}$ such that $hf \in H^2(\beta)$ for all $f \in H^2(\beta)$. When $h \in \mathcal{M}(H^2(\beta))$, the operator $M_h$ of multiplication by $h$ is bounded on $H^2(\beta)$ by the closed graph theorem. The space $\mathcal{M}(H^2(\beta))$ equipped with the operator norm is a Banach space. We note the obvious property

$$\mathcal{M}(H^2(\beta)) \hookrightarrow H^\infty \text{ contractively.} \tag{6-1}$$

Indeed, if $h \in \mathcal{M}(H^2(\beta))$, we easily get, for all $w \in \mathbb{D}$,

$$M_h^*(K_w) = \overline{h(w)} K_w,$$

and so by taking norms and simplifying, we are left with $|h(w)| \leq \|M_h\|$, showing that $h \in H^\infty$ with $\|h\|_\infty \leq \|M_h\|$.

**Proposition 6.1.** *We have $\mathcal{M}(H^2(\beta)) = H^\infty$ isomorphically if and only if $\beta$ is essentially decreasing.*

*Proof.* The sufficient condition is proved in [Lefèvre et al. 2021, beginning of the proof of Proposition 3.16]. For the necessity, we then have $\|M_h\| \approx \|h\|_\infty$ for every $h \in H^\infty$ by the Banach isomorphism theorem. Now, for $m > n$ (recall that $e_n(z) = z^n$),

$$e_m(z) = z^{m-n} z^n = (M_{e_{m-n}} e_n)(z);$$

so, since $\|M_{e_{m-n}}\| \leq C \|e_{m-n}\|_\infty = C$ for some positive constant $C$,

$$\beta_m = \|e_m\|^2 \leq C^2 \|e_n\|^2 = C^2 \beta_n. \qquad \square$$

In [Lefèvre et al. 2021, Section 3.6], we gave the following notion of an *admissible* Hilbert space of analytic functions.

**Definition 6.2.** A Hilbert space $H$ of analytic functions on $\mathbb{D}$, containing the constants, and with reproducing kernels $K_a$, $a \in \mathbb{D}$, is said to be *admissible* if

(i) $H^2$ is continuously embedded in $H$,

(ii) $\mathcal{M}(H) = H^\infty$,

(iii) the automorphisms of $\mathbb{D}$ induce bounded composition operators on $H$,

(iv) $\dfrac{\|K_a\|_H}{\|K_b\|_H} \leq h\left(\dfrac{1-|b|}{1-|a|}\right)$ for $a, b \in \mathbb{D}$, where $h \colon \mathbb{R}^+ \to \mathbb{R}^+$ is a nondecreasing function.

We proved in that paper that every weighted Hilbert space $H^2(\beta)$ with $\beta$ nonincreasing is admissible under the additional hypothesis that the automorphisms of $\mathbb{D}$ induce bounded composition operators. In view of Theorem 4.6, we get the following result.

**Proposition 6.3.** *Let $\beta$ be a weight.*

(1) *If $\beta$ is essentially decreasing, then we have* (i), (ii), (iii) *in Definition 6.2.*

(2) *If $\beta$ is slowly oscillating, then we have* (iv) *in Definition 6.2.*

Let us give a different proof from the one in [Lefèvre et al. 2021].

*Proof.* (1) Let us assume that $\beta$ is essentially decreasing. Then item (i) holds, as well as item (ii), by Proposition 6.1. Item (iii) is Theorem 4.6.

(2) Now we assume that $\beta$ is slowly oscillating.

Let $0 < s < r < 1$.

Without loss of generality, we may assume that $r, s \geq \frac{1}{2}$. It is enough to prove

$$\|K_r\|^2 \leq C\|K_{r^2}\|^2 \tag{6-2}$$

for some constant $C > 1$. Indeed, iteration of (6-2) gives

$$\|K_r\|^2 \leq C^k\|K_{r^{2^k}}\|^2,$$

and if $k$ is the smallest integer such that $r^{2^k} \leq s$, we have

$$2^{k-1}\log r > \log s$$

and

$$2^k \leq D\frac{1-s}{1-r},$$

where $D$ is a numerical constant. Writing $C = 2^\alpha$ with $\alpha > 1$, we obtain

$$\left(\frac{\|K_r\|}{\|K_s\|}\right)^2 \leq C^k = (2^k)^\alpha \leq D^\alpha\left(\frac{1-s}{1-r}\right)^\alpha.$$

To prove (6-2), we pick some $M > 1$ such that

$$\beta_{2n} \geq M^{-1}\beta_n$$

and

$$\beta_{2n-1} \geq M^{-1}\beta_n$$

for all $n \geq 1$, since $\beta$ is slowly oscillating. Write $t = r^2$. We have

$$\|K_r\|^2 = \frac{1}{\beta_0} + \sum_{n=1}^{\infty}\frac{t^{2n}}{\beta_{2n}} + \sum_{n=1}^{\infty}\frac{t^{2n-1}}{\beta_{2n-1}},$$

implying, since $t^{2n-1} \leq 4t^{2n}$,

$$\|K_r\|^2 \leq \frac{1}{\beta_0} + M\sum_{n=1}^{\infty}\frac{t^{2n}}{\beta_n} + 4M\sum_{n=1}^{\infty}\frac{t^{2n}}{\beta_n} \leq 5M\|K_t\|^2.$$ □

The notion of an admissible Hilbert space $H$ is useful for the set of conditional multipliers:

$$\mathcal{M}(H, \varphi) = \{w \in H : w(f \circ \varphi) \in H \text{ for all } f \in H\}.$$

As a corollary of [Lefèvre et al. 2021, Theorem 3.18], we get the following.

**Corollary 6.4.** *Let $\beta$ be essentially decreasing and slowly oscillating. Then*

(1) $\mathcal{M}(H^2, \varphi) \subseteq \mathcal{M}(H^2(\beta), \varphi)$,

(2) $\mathcal{M}(H^2(\beta), \varphi) = H^2(\beta)$ *if and only if* $\|\varphi\|_\infty < 1$,

(3) $\mathcal{M}(H^2(\beta), \varphi) = H^\infty$ *if and only if* $\varphi$ *is a finite Blaschke product.*

We add here as another application of our results an answer to a question appearing in Problem 5 in the thesis of Zorboska [1988].

**Theorem 6.5.** *Let $\beta$ be a weight such that $H^2(\beta)$ is disc-automorphism-invariant, and let $\varphi$ be a symbol inducing a compact composition operator on $H^2(\beta)$. Then the Denjoy–Wolff point of $\varphi$ must be in $\mathbb{D}$.*

In other words, $\varphi$ has a fixed point in $\mathbb{D}$.

In the statement, "$H^2(\beta)$ is disc-automorphism-invariant" means that, for all the automorphisms $T_a$, where $a \in \mathbb{D}$, we have that $C_{T_a}$ is bounded on $H^2(\beta)$ (equivalently it is bounded for at least one $a \in \mathbb{D} \setminus \{0\}$).

For the definition of the Denjoy–Wolff point, we refer to [Shapiro 1993].

*Proof.* From Theorem 4.9, we know that $\beta$ is slowly oscillating, and from Proposition 6.3, we know that

$$\frac{\|K_a\|_{H^2(\beta)}}{\|K_b\|_{H^2(\beta)}} \leq h\left(\frac{1-|b|}{1-|a|}\right) \quad \text{for every } a, b \in \mathbb{D}, \tag{6-3}$$

where $h \colon \mathbb{R}^+ \to \mathbb{R}^+$ is a nondecreasing function.

Now we split the proof into two cases:

• If $\sum 1/\beta_n < \infty$, then $H^2(\beta) \subset A(\mathbb{D})$ (continuously) thanks to the Cauchy–Schwarz inequality. It follows from [Shapiro 1987, Theorem 2.1] that $\|\varphi\|_\infty < 1$, and the conclusion follows obviously.

• If $\sum 1/\beta_n = \infty$, then the normalized reproducing kernel $K_z/\|K_z\|$ is weakly converging to 0 when $|z| \to 1^-$ since $\|K_z\| \to +\infty$.

Since $C_\varphi$ is compact, $C_\varphi^*$ is compact as well, and we get

$$\frac{K_{\varphi(z)}}{\|K_z\|} \to 0 \quad \text{when } |z| \to 1^-$$

and equivalently

$$\frac{\|K_z\|}{\|K_{\varphi(z)}\|} \to +\infty \quad \text{when } |z| \to 1^-.$$

But, from (6-3), we get

$$h\left(\frac{1-|\varphi(z)|}{1-|z|}\right) \to +\infty \quad \text{when } |z| \to 1^-;$$

hence, since $h$ is nondecreasing,

$$\frac{1-|\varphi(z)|}{1-|z|} \to +\infty \quad \text{when } |z| \to 1^-.$$

By the Denjoy–Wolff theorem [Shapiro 1993], the conclusion follows in this case too. □

## Appendix

In this appendix, we give the proof of Proposition 4.17.

The following lemma can be found in [Montgomery 1994, Lemma 1, p. 47].

**Lemma A.1.** *Let $F\colon [u, v] \to \mathbb{R}$, with $u < v$, be a $\mathcal{C}^2$-function with $F'' > 0$ and $F'$ not vanishing on $[u, v]$. Let*

$$J = \int_u^v \mathrm{e}^{iF(x)}\, dx.$$

*Then*:

(a) *if $F' > 0$ on $[u, v]$, we have $|J| \le 2/F'(u)$,*

(b) *if $F' < 0$ on $[u, v]$, we have $|J| \le 2/|F'(v)|$.*

*Proof of Proposition 4.17.* Write now the integral $I$ of Proposition 4.17 on $[A, B]$ as $I = I_1 + I_2 + I_3$, with

$$I_1 = \int_A^{c-\eta} \mathrm{e}^{iF(x)}\, dx, \quad I_2 = \int_{c-\eta}^{c+\eta} \mathrm{e}^{iF(x)}\, dx, \quad I_3 = \int_{c+\eta}^{B} \mathrm{e}^{iF(x)}\, dx.$$

Lemma A.1 with $u = A$ and $v = c - \eta$ implies, since $F' < 0$ on $[A, c - \eta]$,

$$|I_1| \le \frac{2}{|F'(c-\eta)|} \le \frac{2}{\eta\lambda_2}, \tag{A-1}$$

where, for the last inequality, we just have to write

$$|F'(c-\eta)| = F'(c) - F'(c-\eta) = \eta F''(\xi)$$

for some $\xi \in [c - \eta, c]$ and to note that $F''(\xi) \ge \lambda_2$, by hypothesis.

Similarly, Lemma A.1 with $u = c + \eta$ and $v = B$ implies

$$|I_3| \le \frac{2}{F'(c+\eta)} \le \frac{2}{\eta\lambda_2}. \tag{A-2}$$

We can now estimate $I_2$. The Taylor formula shows that

$$F(x) = F(c) + \tfrac{1}{2}(x-c)^2 F''(c) + R,$$

with

$$|R| \le \tfrac{1}{6}|x-c|^3 \lambda_3.$$

Hence

$$I_2 = \mathrm{e}^{iF(c)} \int_0^\eta 2\exp\bigl(\tfrac{1}{2} i x^2 F''(c)\bigr)\, dx + S,$$

with

$$|S| \le \lambda_3 \int_0^\eta \tfrac{1}{3} x^3\, dx = \tfrac{1}{12}\eta^4 \lambda_3.$$

Finally, set

$$K = \int_0^\eta 2\exp\bigl(\tfrac{1}{2} i x^2 F''(c)\bigr)\, dx.$$

We make the change of variable $x = \sqrt{2/F''(c)}\sqrt{t}$. Recall that $\int_0^\infty e^{it}/\sqrt{t}\,dt = \sqrt{\pi}e^{i\pi/4}$ is the classical Fresnel integral and that an integration by parts gives, for $m > 0$,

$$\left|\int_m^\infty \frac{e^{it}}{\sqrt{t}}\,dt\right| \le \frac{2}{\sqrt{m}}.$$

Therefore, with $m = \frac{1}{2}\eta^2 F''(c)$,

$$K = \sqrt{\frac{2}{F''(c)}}\int_0^m \frac{e^{it}}{\sqrt{t}}\,dt = \sqrt{\frac{2\pi}{F''(c)}}e^{i\pi/4} + R_m,$$

with

$$|R_m| \le C\sqrt{\frac{1}{F''(c)}}\frac{1}{\sqrt{m}} \le \frac{C}{\eta\lambda_2}.$$

All in all, we proved that

$$I_2 = \sqrt{\frac{2\pi}{F''(c)}}\exp[i(F(c)+\pi/4)] + O\Big(\frac{1}{\eta\lambda_2} + \eta^4\lambda_3\Big), \tag{A-3}$$

and the same estimate holds for $I$, thanks to (A-1) and (A-2).

We have hence proved Proposition 4.17. □


## Acknowledgements

L. Rodríguez-Piazza is partially supported by the projects PGC2018-094215-B-I00 and PID2022-136320NB-I00 (Spanish Ministerio de Ciencia, Innovación y Universidades, and FEDER funds). Parts of this paper were written when he visited the Université d'Artois in Lens and the Université de Lille in January 2020 and September 2022. It is his pleasure to thank all his colleagues at these universities for their warm welcome.

Parts of this paper were written during an invitation of Li and Queffélec by the IMUS at the Universidad de Sevilla; it is their pleasure to thank all the people in that university who made this stay possible and very pleasant and especially Manuel Contreras.

Queffélec was partly supported by the Labex CEMPI (ANR-LABX-0007-01).

This work is also partially supported by the grant ANR-17-CE40-0021 of the French National Research Agency ANR (project Front).

We warmly thank R. Zarouf for useful discussions and information.


## References


[Borichev et al. 2024] A. Borichev, K. Fouchet, and R. Zarouf, "On the Fourier coefficients of powers of a Blaschke factor and strongly annular functions", *Constr. Approx.* **60**:1 (2024), 33–86. MR Zbl

[Chalendar and Partington 2014] I. Chalendar and J. R. Partington, "Norm estimates for weighted composition operators on spaces of holomorphic functions", *Complex Anal. Oper. Theory* **8**:5 (2014), 1087–1095. MR Zbl

[Chalendar and Partington 2017] I. Chalendar and J. R. Partington, "Compactness and norm estimates for weighted composition operators on spaces of holomorphic functions", pp. 81–89 in *Harmonic analysis, function theory, operator theory, and their applications*, edited by P. Jaming et al., Theta Ser. Adv. Math. **19**, Theta, Bucharest, 2017. MR Zbl

[Cowen 1990] C. C. Cowen, "An application of Hadamard multiplication to operators on weighted Hardy spaces", *Linear Algebra Appl.* **133** (1990), 21–32. MR Zbl

[Duren 1983] P. L. Duren, *Univalent functions*, Grundl. Math. Wissen. **259**, Springer, 1983. MR Zbl

[Gallardo-Gutiérrez and Partington 2013] E. A. Gallardo-Gutiérrez and J. R. Partington, "Norms of composition operators on weighted Hardy spaces", *Israel J. Math.* **196**:1 (2013), 273–283. MR Zbl

[Girard 1973] D. M. Girard, "The behavior of the norm of an automorphism of the unit disk", *Pacific J. Math.* **47** (1973), 443–456. MR Zbl

[Goluzin 1951] G. M. Goluzin, "On majorants of subordinate analytic functions, I", *Mat. Sb.*, *Nov. Ser.* **29** (1951), 209–224. In Russian. MR Zbl

[Kacnelson 1972] V. E. Kacnelson, "A remark on canonical factorization in certain spaces of analytic functions", *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.* (*LOMI*) **30** (1972), 163–164. In Russian; translated in *J. Soviet Math.* **4**, 1975, no. 2 1976, 444–445. MR Zbl

[Kellay and Lefèvre 2012] K. Kellay and P. Lefèvre, "Compact composition operators on weighted Hilbert spaces of analytic functions", *J. Math. Anal. Appl.* **386**:2 (2012), 718–727. MR Zbl

[Kriete and MacCluer 1995] T. L. Kriete and B. D. MacCluer, "A rigidity theorem for composition operators on certain Bergman spaces", *Michigan Math. J.* **42**:2 (1995), 379–386. MR Zbl

[Lefèvre et al. 2021] P. Lefèvre, D. Li, H. Queffélec, and L. Rodríguez-Piazza, "Comparison of singular numbers of composition operators on different Hilbert spaces of analytic functions", *J. Funct. Anal.* **280**:3 (2021), art. id. 108834. MR Zbl

[Li et al. 2014] D. Li, H. Queffélec, and L. Rodríguez-Piazza, "A spectral radius type formula for approximation numbers of composition operators", *J. Funct. Anal.* **267**:12 (2014), 4753–4774. MR Zbl

[Littlewood 1925] J. E. Littlewood, "On inequalities in the theory of functions", *Proc. London Math. Soc.* (2) **23**:7 (1925), 481–519. MR Zbl

[Montgomery 1994] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, CBMS Regional Conference Series in Mathematics **84**, Amer. Math. Soc., Providence, RI, 1994. MR Zbl

[Reich 1954] E. Reich, "An inequality for subordinate analytic functions", *Pacific J. Math.* **4** (1954), 259–274. MR Zbl

[Shapiro 1987] J. H. Shapiro, "Compact composition operators on spaces of boundary-regular holomorphic functions", *Proc. Amer. Math. Soc.* **100**:1 (1987), 49–57. MR Zbl

[Shapiro 1993] J. H. Shapiro, *Composition operators and classical function theory*, Springer, 1993. MR Zbl

[Shields 1974] A. L. Shields, "Weighted shift operators and analytic function theory", pp. 49–128 in *Topics in operator theory*, edited by C. Pearcy, Math. Surveys **13**, Amer. Math. Soc., Providence, RI, 1974. MR Zbl

[Szehr and Zarouf 2020] O. Szehr and R. Zarouf, "$l_p$-norms of Fourier coefficients of powers of a Blaschke factor", *J. Anal. Math.* **140**:1 (2020), 1–30. MR Zbl

[Szehr and Zarouf 2021] O. Szehr and R. Zarouf, "Explicit counterexamples to Schäffer's conjecture", *J. Math. Pures Appl.* (9) **146** (2021), 1–30. MR Zbl

[Titchmarsh 1986] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, 2nd ed., Oxford University Press, 1986. MR Zbl

[Zorboska 1988] N. Zorboska, *Composition operators on weighted Hardy spaces*, Ph.D. thesis, University of Toronto, 1988, available at https://www.proquest.com/docview/303632738. MR

Received 23 Oct 2023. Revised 30 Jul 2024. Accepted 20 Sep 2024.

PASCAL LEFÈVRE: pascal.lefevre@univ-artois.fr
*Université d'Artois, UR 2462, Laboratoire de Mathématiques de Lens (LML), F-62300 Lens, France*

DANIEL LI: daniel.li@univ-artois.fr
*Université d'Artois, UR 2462, Laboratoire de Mathématiques de Lens (LML), F-62300 Lens, France*

HERVÉ QUEFFÉLEC: herve.queffelec@univ-lille.fr
*Université de Lille, CNRS, UMR 8524 – Laboratoire Paul Painlevé, F-59000 Lille, France*

LUIS RODRÍGUEZ-PIAZZA: piazza@us.es
*Dpto. de Análisis Matemático & IMUS, Facultad de Matemáticas, Universidad de Sevilla, Sevilla, Spain*

ANALYSIS AND PDE
Vol. 18 (2025), No. 8, pp. 1955–2032
DOI: 10.2140/apde.2025.18.1955

# LONG-TIME BEHAVIOR OF THE STOKES-TRANSPORT SYSTEM IN A CHANNEL

ANNE-LAURE DALIBARD, JULIEN GUILLOD AND ANTOINE LEBLOND

We consider here a two-dimensional incompressible fluid in a periodic channel, whose density is advected by pure transport, and whose velocity is given by the Stokes equation with gravity source term. Dirichlet boundary conditions are taken for the velocity field on the bottom and top of the channel and periodic conditions in the horizontal variable. We prove that the affine stratified density profile is stable under small perturbations in Sobolev spaces and prove convergence of the density to another limiting stratified density profile for large time with an explicit algebraic decay rate. Moreover, we are able to precisely identify the limiting profile as the decreasing vertical rearrangement of the initial density. Finally, we show that boundary layers are formed for large times in the vicinity of the upper and lower boundaries. These boundary layers, which had not been identified in previous works, are given by a self-similar ansatz and driven by a linear mechanism. This allows us to precisely characterize the long-time behavior beyond the constant limiting profile and reach more optimal decay rates.



## 1. Introduction

The Stokes-transport system

$$\begin{cases} \partial_t \rho + \boldsymbol{u} \cdot \nabla \rho = 0, \\ -\Delta \boldsymbol{u} + \nabla p = -\rho \boldsymbol{e}_z, \\ \operatorname{div} \boldsymbol{u} = 0, \\ \rho|_{t=0} = \rho_0 \end{cases} \tag{1-1a}$$

models the evolution of an incompressible inhomogeneous fluid with density $\rho$ and velocity and pressure fields $(\boldsymbol{u}, p)$. For physical reasons and without loss of generality, we assume that the initial density $\rho_0$ is

*MSC2020:* primary 35B35, 35B40, 35Q49, 76D07, 76D10; secondary 35D35, 35M13.
*Keywords:* Stokes-transport, channel, stability, affine profile, decreasing vertical rearrangement, long-time behavior, boundary layers, Dirichlet boundary conditions, optimal decay rate.

© 2025 MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY).
Open Access made possible by subscribing institutions via Subscribe to Open.

nonnegative. This equation will be studied in a two-dimensional periodic strip, namely $\Omega = \mathbb{T} \times (0, 1)$ with variables $(x, z) \in \Omega$ and with Dirichlet boundary condition of the velocity field:

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \partial\Omega. \tag{1-1b}$$

It consists of a coupling of the transport equation for the density of the fluid with a velocity field satisfying for all times the Stokes equation with gravity forcing $-\rho \boldsymbol{e}_z$, where $\boldsymbol{e}_z$ is the unitary vertical vector. This equation has been studied in particular in [Höfer 2018; Mecherbet 2021] showing that (1-1a) is a model obtained as the homogenization limit of inertialess particles in a fluid-satisfying Stokes equation. The more recent paper [Grayer 2023] shows that this system is obtained as a formal limit where the Prandtl number is infinite. In this paper, the domain is chosen as $\Omega = \mathbb{T} \times (0, 1)$, which describes a physically meaningful situation including Dirichlet boundary conditions.

***Well-posedness.*** The well-posedness of this system has been shown in [Antontsev et al. 2000] for piecewise constant initial data in bounded domains of $\mathbb{R}^n$ and in [Leblond 2022] for arbitrary $L^\infty$ data in bounded domains of $\mathbb{R}^2$ and $\mathbb{R}^3$ or in the infinite strip $\mathbb{R} \times (0, 1)$, the well-posedness in $\Omega = \mathbb{T} \times (0, 1)$ being a direct consequence.

Well-posedness in Sobolev spaces is required for our results. Since this result does not seem to appear in the literature, we provide a concise proof of the global well-posedness of this problem in Appendix A for the sake of completeness. More precisely, for any $\rho_0 \in H^m$ with $m \geq 3$, there exists a unique strong solution $(\rho, \boldsymbol{u})$ of (1-1a) with $\rho \in C(\mathbb{R}_+; H^m(\Omega))$ and $\boldsymbol{u} \in C(\mathbb{R}_+; H^{m+2}(\Omega))$. Well-posedness in other domains and spaces has also been proven; see for example the recent results [Mecherbet and Sueur 2024; Inversi 2023].

***Steady states.*** Before going further let us observe that the stationary states, i.e., states such that $\partial_t \rho = 0$, of this system are precisely the *stratified* density profiles, which means in this paper density profiles depending only on the vertical variable $z$. Indeed, for such a map $\rho_s = \rho_s(z)$,

$$(\rho, \boldsymbol{u}, p) = \left(\rho_s, \boldsymbol{0}, -\int^z \rho_s(z')\,\mathrm{d}z'\right)$$

is a solution of (1-1a). To show the converse, let us introduce the potential energy associated to a density profile $\rho$,

$$\mathrm{E}(\rho) := \int_\Omega z\rho \,\mathrm{d}x\,\mathrm{d}z.$$

The energy balance is

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}(\rho) = \int_\Omega z\partial_t\rho = -\int_\Omega z\boldsymbol{u}\cdot\nabla\rho = \int_\Omega \boldsymbol{u}\cdot\boldsymbol{e}_z\rho = -\|\nabla\boldsymbol{u}\|_{L^2}^2, \tag{1-2}$$

where the divergence-free and the Dirichlet boundary conditions on $\boldsymbol{u}$ are used in the integration by parts. The last equality is simply the basic estimate of the Stokes equation. The potential energy dissipates exactly through the viscosity effects. From this observation we see that the whole evolution is nonreversible; the fluid only rearranges in states of lower potential energy. Moreover, a stationary state is exactly a state for which $\boldsymbol{u} = \boldsymbol{0}$; therefore it means that the density $\rho$ and the pressure $p$ must satisfy

$$\nabla p = -\rho\boldsymbol{e}_z,$$

so that the pressure is independent of the $x$-variable, implying $\rho$ depends only on the $z$-variable.

The aim of this paper is to study the long-time behavior of perturbations of stratified initial data in the stable regime, with lighter fluid on top and heavier fluid on the bottom. We will prove three different results: The first one, Theorem 1.1, provides the stability of such stratified profiles, together with some decay estimates. The second one, Theorem 1.2, gives an explicit asymptotic decomposition of solutions of a linear version of (1-1a) as $t \to \infty$. In particular, we identify boundary layer profiles in the vicinity of the top and bottom boundaries. Eventually, in Theorem 1.3, we go back to the nonlinear system (1-1a) and provide a more precise description of the solutions as $t \to \infty$, building on the analysis from Theorem 1.2. A striking consequence of our results lies in the fact that the boundaries slow down the relaxation towards the asymptotic state. This new observation could probably be adapted to other systems; see Remark 1.4.

***Main stability result.*** For simplicity and in the rest of this paper, we consider perturbations of the affine profile $\rho_s(z) = 1 - z$, although more general profiles such that $\partial_z \rho_s < 0$ could be considered; see Remark 2.8.

Our main stability result for perturbations vanishing on the boundary is the following:

**Theorem 1.1.** *There exists a small universal constant $\varepsilon_0 > 0$ such that, for any $\rho_0 \in H^6(\Omega)$ satisfying $\|\rho_0 - \rho_s\|_{H^6} \leq \varepsilon_0$ and $\rho_0 - \rho_s \in H_0^2(\Omega)$, the solution $\rho$ of* (1-1a) *satisfies*

$$\|\rho - \rho_\infty\|_{L^2(\Omega)} \lesssim \frac{\varepsilon_0}{1+t}, \quad \|\rho - \rho_\infty\|_{H^4(\Omega)} \lesssim \varepsilon_0, \tag{1-3}$$

*where $\rho_\infty$ is given by the decreasing vertical rearrangement of $\rho_0$*:

$$\rho_\infty(z) := \int_0^\infty \mathbf{1}_{0 \leq z \leq |\{\rho_0 > \lambda\}|} \, \mathrm{d}\lambda.$$

Note that the condition on $H_0^2(\Omega)$ is equivalent to the following requirements, discussed in the following subsections:

$$\rho_0|_{\partial\Omega} = \rho_s|_{\partial\Omega}, \quad \partial_{\boldsymbol{n}} \rho_0|_{\partial\Omega} = \partial_{\boldsymbol{n}} \rho_s|_{\partial\Omega}.$$

This theorem will be proven in Section 2, and we provide a scheme of proof at the end of this section.

***Remarks on the main stability result.***

- Since the set of steady states is not discrete, it is expected that $\rho_s$ is not asymptotically stable, and that the long-time behavior is given by a slightly modified density profile. In general, this asymptotic profile depends on the entire nonlinear dynamics in a very nonexplicit way. However the transport equation is remarkable since it preserves the measure of the level sets. This property combined with the fact that the asymptotic profile is strictly decreasing (as a smooth perturbation of $\rho_s$) allows us to identify the asymptotic profile as the decreasing vertical rearrangement of $\rho_0$, which can be computed directly from $\rho_0$ without dependence on the full nonlinear dynamics. See Section 2.4 for details.

- In fact the identification of the limit as the decreasing vertical rearrangement is quite general and only requires two properties. First the density needs to converge to a stratified (i.e., independent of $x$) and decreasing limiting profile. Second, the density should satisfy a transport equation with well-defined characteristics. This is in particular the case for the incompressible porous medium equation; see Section 2.4 for details.

• This result proves the stability of the particular state $\rho_s(z) = 1 - z$. We believe that the result generalizes and our proofs adapt to the case of stratified $\rho_s \in H^6$ satisfying

$$\sup_{(0,1)} \partial_z \rho_s < 0.$$

This remark is detailed at the end of Section 2.2. We note that without the monotony assumption, lighter fluid might be below heavy one, and physical instabilities — similar to the Rayleigh–Bénard or Rayleigh–Taylor instabilities — are expected to develop (see [Drazin and Reid 2004]). Some weak convergence up to extraction toward a stationary state could be proven, but the limit might be a nontrivial $\omega$-limit set in general. In any case, it is not clear whether convergence to the rearranging steady state holds.

• One can of course wonder about the strong regularity requirement in Theorem 1.1. It turns out that one can adapt a strategy developed in [Kiselev and Yao 2023] about the instability of the incompressible porous media equation. Indeed, the arguments are essentially geometric, and the result is the same: there exist smooth perturbations small in $H^{2-}(\Omega)$-norm such that $\limsup_{t\to\infty} \|\rho(t) - \rho_s\|_{H^s(\Omega)} = \infty$ for any $s > 1$. Therefore, this shows the existence of a regularity threshold between $H^2(\Omega)$ and $H^6(\Omega)$ between stability and instability. The details are provided in the thesis of Antoine Leblond [2023].

• Finally, an interesting question is the optimality of the $(1+t)^{-1}$ decay in (1-3). The dynamics of the equation preserve the fact that the perturbation and its normal derivative are vanishing on $\partial\Omega$ i.e., $\rho - \rho_s \in H_0^2(\Omega)$. For higher normal derivatives this property is not preserved, and this is the main reason why the time-decay is limited. This is one of the main motivations to study the formation of boundary layers in this system, together with the possibility to allow nonvanishing perturbations on $\partial\Omega$. Theorem 1.3 below indicates that the optimal decay rate under the assumptions of Theorem 1.1 is very likely $(1+t)^{-9/8}$. We refer to the discussion at the top of page 1961 for more details.

***Related results and comparison with the incompressible porous medium equation.*** In [Gancedo et al. 2025] the interface problem for (1-1a) is considered also in the domain $\Omega = \mathbb{T} \times (0, 1)$. The interface problem treats the case where the density is equal to two different constants below and above an interface $\Gamma(t) \subset \Omega$. The question lies in the regularity of the interface, the well-posedness for $L^\infty$ densities being established in [Leblond 2022; Antontsev et al. 2000]. More precisely, the authors prove local well-posedness for the interface in $C^{1,\gamma}$ for $0 < \gamma < 1$ as well as the global well-posedness and decay of small perturbation in $H^3(\mathbb{T})$ of the flat interface with lighter fluid on top. The proof is very different from ours as it uses a contour dynamics equation, but the spirit of the stability result is pretty similar.

Let us also compare the results and properties of the Stokes-transport equation and of the incompressible porous medium equation, namely (1-1a) where the Stokes equation is replaced by Darcy's law,

$$\begin{cases} \partial_t \rho + \boldsymbol{u} \cdot \nabla \rho = 0, \\ \boldsymbol{u} + \nabla p = -\rho \boldsymbol{e}_z, \\ \operatorname{div} \boldsymbol{u} = 0, \\ \rho|_{t=0} = \rho_0. \end{cases} \tag{1-4}$$

This equation has been intensively studied and we only cite comparable results. The question of well-posedness is much more difficult than for the Stokes-transport equation. In particular global well-posedness like Theorem A.1 seems to remain an open question. Local in time well-posedness has been shown in [Córdoba et al. 2007; Xue 2009; Yu and He 2014; Constantin et al. 2015], whereas ill-posedness through nonuniqueness in some spaces has been shown in [Córdoba et al. 2011; Shvydkoy 2011; Isett and Vicol 2015].

Concerning classical global solutions, the only known results have been proven for initial data close enough in Sobolev space to the stratified initial data $\rho_s(z) = 1 - z$ by [Elgindi 2017] in $\mathbb{R}^2$ and $\mathbb{T}^2$ and later generalized in [Castro et al. 2019a] to the domain $\mathbb{T} \times (0, 1)$. More precisely, these results prove that the profile $\rho_s(z) = 1 - z$ is asymptotically stable under small perturbations in $H^m$ for some $m$. Let us also mention the recent work [Park 2025], which revisits these results, relying mainly on energy estimates.

In $\mathbb{T} \times (0, 1)$, the boundary conditions identified and used by Castro, Córdoba and Lear [Castro et al. 2019a] ensure that the main linearized structure remains stable by differentiation, which makes the analysis of that work similar to the one of the periodic or whole space case. This analysis has then been extended by the same authors to the Boussinesq system with a damping velocity term in [Castro et al. 2019b]. In particular integrations by parts of high-order derivatives are possible to obtain uniform bounds in Sobolev spaces of high enough regularity. By using similar boundary conditions for Stokes-transport in $\mathbb{T} \times (0, 1)$, the results of [Castro et al. 2019a] could be adapted in a straightforward way. In our situation, *the presence of the Dirichlet boundary condition is the major obstacle*. In particular, uniform bounds in high-regularity Sobolev spaces are no longer valid, as Theorem 1.3 below will highlight. This is due to the presence of boundary layers, as explained above. More details are provided below in the scheme of the proof.

Eventually, let us mention that the existence of the limiting profile was obtained in [Elgindi 2017; Castro et al. 2019a] through a fixed-point argument. One contribution of the present paper is to precisely identify the long-time asymptotic profile as the decreasing vertical rearrangement of $\rho_0$. As explained previously, our method to identify the limiting profile is robust and in particular also applies to show that the long-time asymptotic profile for the incompressible porous media equation is also given by the decreasing vertical rearrangement.

***Linear asymptotic expansion for nonvanishing perturbation on $\partial\Omega$.*** Theorem 1.1 is only valid under the assumption that the perturbation and its normal derivative are vanishing on $\partial\Omega$, i.e., when $\rho_0 - \rho_s \in H_0^2(\Omega)$. If the perturbation does not vanish on the boundary, this question is nontrivial even for the linearized equations around $\rho_s = 1 - z$: denoting by $\theta$ the perturbed density, we consider

$$\begin{cases} \partial_t \theta - \boldsymbol{u} \cdot \boldsymbol{e}_z = 0, \\ -\Delta \boldsymbol{u} + \nabla p = -\theta \boldsymbol{e}_z, \\ \operatorname{div} \boldsymbol{u} = 0, \\ \theta|_{t=0} = \theta_0. \end{cases} \tag{1-5}$$

It can be easily checked that Theorem 1.1 is also valid for (1-5). In other words, if $\theta_0 \in H^6 \cap H_0^2$, then $\|\theta(t)\|_{L^2} \lesssim (1+t)^{-1}$. Note that there is no smallness assumption in this case because the system is linear.

If $\theta_0$ or $\partial_{\boldsymbol{n}}\theta_0$ do not vanish on the boundary, however, it turns out that $\theta$ vanishes as $t\to\infty$ but with a much slower rate. This is due to the formation of boundary layers of typical size $t^{-1/4}$ as $t\to\infty$, in the vicinity of $z=0$ and $z=1$. More precisely, we will prove the following result in Section 3:

**Theorem 1.2.** *Let $\theta_0\in H^s(\Omega)$ for some $s$ sufficiently large. Then the solution of* (1-5) *satisfies*

$$\theta=\bar\theta_0+\theta^{\mathrm{BL}}+O(t^{-1})\quad \text{in } L^2(\Omega) \text{ as } t\to\infty,$$

*where $\bar\theta_0(z)=\frac{1}{2\pi}\int_0^{2\pi}\theta_0(x,z)\,\mathrm{d}x$ is the horizontal average of the initial data and $\theta^{\mathrm{BL}}$ is the boundary layer part whose leading terms are*

$$\theta^{\mathrm{BL}}=\Theta^0_{\mathrm{top}}(x,t^{\frac14}(1-z))+\Theta^0_{\mathrm{bot}}(x,t^{\frac14}z)+\text{l.o.t.},\tag{1-6}$$

*with $\Theta^0_{\mathrm{top}}$ (resp. $\Theta^1_{\mathrm{bot}}$) decaying exponentially as $Z_{\mathrm{top}}=t^{1/4}(1-z)\to\infty$ (resp. as $Z_{\mathrm{bot}}=t^{1/4}z\to\infty$). Furthermore, for $t\ge 1$,*

$$C^{-1}\|\theta_0|_{\partial\Omega}\|_{L^2}t^{-\frac18}+O(t^{-\frac38})\le\|\theta^{\mathrm{BL}}\|_{L^2(\Omega)}\le C\|\theta_0|_{\partial\Omega}\|_{L^2}t^{-\frac18}+O(t^{-\frac38}).$$

***Remarks on the linear asymptotic expansion.***

- In fact, the definition of $\theta^{\mathrm{BL}}$ is more involved and is given as a sum in powers of $t^{-1/4}$ of different boundary layer profiles. For instance, in the vicinity of $z=0$ and for $t>1$,

$$\theta^{\mathrm{BL}}=\sum_{j=0}^{4}t^{-\frac{j}{4}}\Theta^j_{\mathrm{bot}}(x,t^{\frac14}z).$$

Furthermore, the construction can be iterated. Up to a stronger regularity requirement on the initial data, we could probably push the expansion of $\theta^{\mathrm{BL}}$ further and prove that $\theta=\theta^{\mathrm{BL}}+O(t^{-k})$ for $k$ arbitrarily large. In this case, the definition of $\theta^{\mathrm{BL}}$ has to be modified in order to include profiles up to $j=4k$. We shall give more details on this matter in Remark 3.8.

- Note that the scaling of boundary layers is consistent with the estimates of Theorem 1.1: heuristically, one power of $t^{1/4}$ is lost with each differentiation (with respect to $z$.)

***Nonlinear asymptotic expansion.*** Let us now go back to the nonlinear problem in the case where $\rho_0-\rho_s\in H^2_0(\Omega)$. In this case, $\rho(t)-\rho_s$ and $\partial_{\boldsymbol{n}}(\rho(t)-\rho_s)$ vanish on the boundary for all $t\ge 0$ (see Lemma 2.1). As a consequence, the advection term is negligible in the vicinity of the boundary, and we expect the dynamics to be driven by a linear mechanism in this zone at main order. Building on the analysis of Theorem 1.2, we then derive uniform bounds in $H^8(\Omega)$, modulo some boundary layer terms:

**Theorem 1.3.** *There exists $\varepsilon_0>0$ small such that, for any $\rho_0\in H^{14}(\Omega)$ satisfying $\|\rho_0-\rho_s\|_{H^{14}}\le\varepsilon_0$ and $\rho_0-\rho_s\in H^3_0(\Omega)$, the solution $\rho$ of* (1-1a) *satisfies*

$$\rho=\rho_\infty+\theta^{\mathrm{BL}}+O(t^{-2})\quad \text{in } L^2(\Omega) \text{ as } t\to\infty,$$

*where $\theta^{\mathrm{BL}}$ is the boundary layer part given by*

$$\theta^{\mathrm{BL}}=\frac1t\Theta_{\mathrm{top}}(x,t^{\frac14}(1-z))+\frac1t\Theta_{\mathrm{bot}}(x,t^{\frac14}z)+\text{l.o.t.}$$

*with $\Theta_{\mathrm{top}}$ (resp. $\Theta_{\mathrm{bot}}$) decaying exponentially as $Z_{\mathrm{top}}=t^{1/4}(1-z)\to\infty$ (resp. as $Z_{\mathrm{bot}}=t^{1/4}z\to\infty$).*

A more precise version of the theorem, including $H^s$ estimates on the remainder, will be provided in Proposition 4.1. We note that $\|\theta^{\mathrm{BL}}\|_{L^2(\Omega)} \lesssim (1+t)^{-9/8}$, so this result strongly suggests that the optimal decay of $\rho-\rho_\infty$ is like $t^{-9/8}$ in $L^2(\Omega)$, which is close to the rate $t^{-1}$ obtained in Theorem 1.1. Theorem 1.3 also shows that the decay rate is dictated by boundary layers. Nevertheless, it is not excluded that the nonlinear dynamics drive the system to the case where these boundary layer terms always vanish, although we expect this behavior to be rather unlikely. Let us emphasize that the formation of boundary layers, let alone the construction of boundary layer profiles, had not been identified in previous works, even in the linear setting of Theorem 1.2. We believe that our analysis could be extended to the incompressible porous media (IPM) system (1-4), for which similar boundary layers are expected to develop; see Remark 1.4 below. Let us also recall that in the cases without boundaries (see [Elgindi 2017; Castro et al. 2019a; Park 2025]), the rate of decay of $\rho-\rho_\infty$ can be arbitrarily large, provided the initial data is sufficiently smooth.

Let us now say a few words about the case when the initial data $\rho_0$ of (1-1a) is such that $\rho_0-\rho_s$ or $\partial_{\boldsymbol{n}}(\rho_0-\rho_s)$ do not vanish on the boundary. We expect the scaling of the boundary layers to be different. Indeed, if the ansatz of the linear case (1-6) is plugged into (1-1a), we find that the quadratic term becomes dominant close to the boundary, and cannot be balanced by other terms in the equation. As a consequence, studying (1-1a) when $\rho-\rho_s \notin H_0^2(\Omega)$ goes beyond the scope of this paper. We expect that the boundary layer equations become nonlinear in this setting.

Note that Theorem 1.3 requires more stringent assumptions on the initial data than Theorem 1.1, since the initial perturbation is assumed to be small in $H^{14}$ (rather than $H^6$), and its second normal derivative is also assumed to vanish on the boundary (i.e., $\partial_z^2(\rho-\rho_s)|_{\partial\Omega}=0$). Actually, the latter condition can be slightly weakened; see Proposition 4.1 for a more precise statement.

**Remark 1.4** (extension to the incompressible porous medium equation). We believe that Theorem 1.3 could be extended to the IPM equation (1-4) when the initial datum $\rho_0$ is sufficiently smooth and such that $\rho_0-\rho_s \in H_0^1$ (i.e., the trace of $\rho_0-\rho_s$ vanishes on the boundary). In this case, $(\rho(t)-\rho_s)|_{\partial\Omega}=0$ for all $t \geq 0$ (see Lemma 2.1).

In this setting, the boundary layer ansatz from Theorem 1.3 should be replaced with

$$\theta^{\mathrm{BL}} = \frac{1}{t}\Theta^{\mathrm{IPM}}_{\mathrm{top}}(x, t^{\frac{1}{2}}(1-z)) + \frac{1}{t}\Theta^{\mathrm{IPM}}_{\mathrm{bot}}(x, t^{\frac{1}{2}}z) + \mathrm{l.o.t.},$$

where the profiles $\Theta^{\mathrm{IPM}}_a$ for $a \in \{\mathrm{top}, \mathrm{bot}\}$ satisfy

$$-\Theta^{\mathrm{IPM}}_a + \tfrac{1}{2} Z\partial_Z \Theta^{\mathrm{IPM}}_a = \partial_x \Psi^{\mathrm{IPM}}_a \partial_Z^2 \Psi^{\mathrm{IPM}}_a = \partial_x \Theta^{\mathrm{IPM}}_a.$$

This system should be compared with (4-18), and is endowed with the boundary conditions

$$\Psi^{\mathrm{IPM}}_a|_{Z=0} = \Theta^{\mathrm{IPM}}_a|_{Z=0} = 0, \quad \partial_Z^2 \Theta^{\mathrm{IPM}}_a|_{Z=0} = \gamma_a,$$

where $\gamma_{\mathrm{top}}(x) = \lim_{t\to\infty} \partial_z^2\theta(t,x,1)$, $\gamma_{\mathrm{bot}}(x) = \lim_{t\to\infty} \partial_z^2\theta(t,x,0)$.

Therefore the situation is very similar to the one of Theorem 1.3: the main difference lies in the thickness of the boundary layer ($t^{-1/2}$ for IPM vs. $t^{-1/4}$ for Stokes-transport), which is consistent with the order of the damping term ($\partial_x^2\Delta^{-1}$ for IPM vs. $\partial_x^2\Delta^{-2}$ for Stokes-transport). Furthermore, if $\partial_z^{2\ell}\theta_0|_{\partial\Omega}=0$

for $0 \le \ell \le k$ and for some $k \ge 1$, then the above ansatz should be replaced by

$$\theta^{\mathrm{BL}} = \frac{1}{t^{k+1}} \Theta^{\mathrm{IPM}}_{\mathrm{top}}(x, t^{\frac{1}{2}}(1-z)) + \frac{1}{t^{k+1}} \Theta^{\mathrm{IPM}}_{\mathrm{bot}}(x, t^{\frac{1}{2}} z) + \text{l.o.t.}$$

Note that this is consistent with the results of [Park 2025] (see also [Castro et al. 2019b]), in which the author proves that $\|\rho(t) - \rho_s\|_{L^2} \lesssim t^{-k-1/2}$ under a slightly more stringent version of the previous assumption.

However, if the trace of $\rho_0 - \rho_s$ on the boundary does not vanish, the situation is different. In this scenario, the nonlinear terms are expected to be of leading order close to the boundary, and we expect that nonlinear boundary layers are created.

Interpolating between the IPM system and the Stokes-transport system, it is also natural to wonder what happens for fractional equations such as

$$\partial_t \theta + \boldsymbol{u} \cdot \nabla \theta = \partial_x^2 (-\Delta)^{-\alpha} \theta,$$

with $\alpha \in (1, 2)$. One should however define carefully the fractional operator $(-\Delta)^{-\alpha}$ in this setting, since the domain $\mathbb{T} \times (0, 1)$ is bounded in the vertical direction (the boundedness in the horizontal direction is not really an issue since we can rely on a Fourier definition of the fractional laplacian in the horizontal variable.) One canonical choice is to use a spectral definition of the fractional laplacian. However, in the present setting, there are two possible choices for the eigenbasis: the eigenfunctions of the laplacian, or the ones of the bilaplacian, described in Lemma B.2. These two choices seem to lead to different operators, and in particular, they are incompatible with one another.

Therefore it seems better to consider the so-called "restricted fractional laplacian": for $\psi \in H^{2\alpha}(\Omega)$ such that $\psi|_{\partial\Omega} = 0$ and $\partial_z \psi|_{\partial\Omega} = 0$ if $\alpha > \frac{3}{2}$, extend $\psi$ by zero outside $\Omega$, and define

$$(-\Delta)^\alpha \psi := C_\alpha \mathrm{PV} \int_{\mathbb{T}\times\mathbb{R}} \frac{\Delta\psi(x', z') - \Delta\psi(x, z)}{|(x, z) - (x', z')|^{2\alpha}} \, \mathrm{d}x' \, \mathrm{d}z'.$$

The equation for $\psi$ then becomes

$$(-\Delta)^\alpha \psi = \partial_x \theta \quad \text{in } \Omega, \qquad \psi|_{\Omega^c} = 0.$$

The main advantage of this choice is to be compatible with the end cases $\alpha = 1$ (IPM) and $\alpha = 2$ (Stokes-transport). However, due to the nonlocal nature of the fractional laplacian, having a description of the boundary layer formation seems much more involved.

***Schemes of proofs.*** Here we explain the main steps and difficulties of the proofs of Theorems 1.1, 1.2 and 1.3.

***Rewriting of the equation.*** Since perturbations of $\rho_s(z) = 1 - z$ are considered, it is natural to introduce the perturbation $\theta$ as

$$\rho = \rho_s + \theta,$$

with initial perturbation $\theta_0 = \rho_0 - \rho_s$. Substituting this expression in (1-1a) and recalling that stratified states do not contribute to the velocity field in the Stokes equation, we obtain the following equation for the perturbation $\theta$:

$$\begin{cases} \partial_t\theta + \boldsymbol{u}\cdot\nabla\theta = u_z, \\ -\Delta\boldsymbol{u} + \nabla p = -\theta\boldsymbol{e}_z, \\ \operatorname{div}\boldsymbol{u} = 0, \\ \boldsymbol{u}|_{\partial\Omega} = \boldsymbol{0}, \\ \theta|_{t=0} = \theta_0. \end{cases}$$

We note that we used the notation $\boldsymbol{u} = (u_x, u_z)$ and that $x$ and $z$ indices always denote the horizontal and vertical components and never derivatives with respect to $x$ or $z$.

The Stokes equation can be simplified by introducing the stream function of the divergence-free velocity field $\boldsymbol{u}$ through $\boldsymbol{u} = \nabla^\perp\psi = (-\partial_z\psi, \partial_x\psi)$. Substituting it in the Stokes equation and considering the curl of this equation, we get

$$\begin{cases} \partial_t\theta + \boldsymbol{u}\cdot\nabla\theta = u_z, \\ \Delta^2\psi = \partial_x\theta, \\ \boldsymbol{u} = \nabla^\perp\psi, \\ \psi|_{\partial\Omega} = \partial_{\boldsymbol{n}}\psi|_{\partial\Omega} = 0, \\ \theta|_{t=0} = \theta_0. \end{cases}$$

Notice that this writing is consistent with the previous observation that any $z$-dependent perturbation of the density does not affect the velocity field.

Once the steady states of (1-1a) are identified as the stratified density profiles, i.e., functions depending only on $z$, it is natural to decompose the perturbation $\theta(t, x, z)$ as the sum of its horizontal average $\bar\theta(t, z)$ and its complement $\theta'(t, x, z)$ with zero horizontal average, following [Elgindi 2017] and others:

$$\theta(t, x, z) = \bar\theta(t, z) + \theta'(t, x, z), \quad \bar\theta(t, z) = \frac{1}{2\pi}\int_0^{2\pi}\theta(t, x, z)\,\mathrm{d}x.$$

We note that contrary to [Elgindi 2017; Castro et al. 2019a], $\bar\theta$ denotes the average rather than the fluctuation, as this seems a more natural notation. In particular our notation is comparable to the standard notation used for the Reynolds-averaged Navier–Stokes equations.

This decomposition is actually orthogonal in any Sobolev space $H^m$ and one can project the transport equation onto the two appropriate complementary subspaces, leading to

$$\begin{cases} \partial_t\theta' + (\boldsymbol{u}\cdot\nabla\theta')' = (1-\partial_z\bar\theta)u_z, & \theta'|_{t=0} = \theta'_0, \\ \partial_t\bar\theta + \overline{\boldsymbol{u}\cdot\nabla\theta'} = 0, & \bar\theta|_{t=0} = \bar\theta_0, \\ \Delta^2\psi = \partial_x\theta', & \psi|_{\partial\Omega} = 0, \\ \boldsymbol{u} = \nabla^\perp\psi, & \partial_{\boldsymbol{n}}\psi|_{\partial\Omega} = 0. \end{cases} \tag{1-7}$$

Although more complicated at first sight, this equation allows us to distinguish the evolution of $\theta'$ and of the average perturbation $\bar\theta$. This is needed since the whole perturbation cannot be expected to decay in Sobolev spaces due to its pure transport. Only the average-free part $\theta'$ is decaying.

***Toy problem on the torus.*** In order to get an intuition of the decay of $\|\theta'\|_{L^2(\Omega)}$ and to highlight the specific difficulties of our work, we will first explain the strategy in the case when the problem is set on the torus, in order to avoid the issues associated with the boundary conditions. More precisely, we consider the following linear problem for $\theta'$ on the torus $\mathbb{T}^2$:

$$\begin{cases} \partial_t\theta' = (1-G)\partial_x\psi + S, \\ \Delta^2\psi = \partial_x\theta', \\ \theta'|_{t=0} = \theta'_0, \end{cases} \tag{1-8}$$

where $G$ is a given small function of $t$ and $z$, whose finality is to be replaced by $\partial_z\bar{\theta}$. The source term $S$, which will include the nonlinearities of the system, will be omitted in this short presentation for simplicity. Note that (1-8) differs from our original system through the periodic boundary conditions on $\psi$ in the vertical variable. The choice of periodic boundary conditions simplifies the analysis in several ways, which we will detail below.

For any $s \geq 0$, applying $\Delta^{s/2}$ to the first equation of (1-8) and projecting on $\Delta^{s/2}\theta'$, we obtain, after several integrations by parts in the right-hand side,

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\Delta^{\frac{s}{2}}\theta'\|_{L^2}^2 &= -\int_{\mathbb{T}^2}\Delta^{\frac{s}{2}}((1-G)\psi)\Delta^{\frac{s}{2}+2}\psi \\ &= -\int_{\mathbb{T}^2}\Delta^{\frac{s}{2}+1}((1-G)\psi)\Delta^{\frac{s}{2}+1}\psi \\ &\leq -(1-\overline{C}\|G\|_{H^{s+2}})\|\Delta^{\frac{s}{2}+1}\psi\|_{L^2}^2, \end{aligned} \tag{1-9}$$

where $\overline{C}$ is a universal constant. As a consequence, if $\overline{C}\|G(t)\|_{H^{s+2}} < 1$, then the $H^s$ norm of $\theta'$ is nonincreasing, and whence uniformly bounded.

Then, the decay of $\|\theta'(t)\|_{L^2}$ is deduced by using the following Gagliardo–Nirenberg interpolation inequality, which in the case of the torus can be proved simply by Fourier analysis:

$$\|\partial_x^{-1}\Delta^2\phi\|_{L^2}^2 \lesssim \frac{1}{K}\|\Delta\phi\|_{L^2}^2 + K^2\|\partial_x^{-3}\Delta^4\phi\|_{L^2}^2, \tag{1-10}$$

where $\partial_x^{-1} f$ denotes the antiderivative of $f$ with null horizontal average, and $K > 0$ is an arbitrary positive constant.

More precisely, combining (1-9) and (1-10) with $\phi = \Delta^{r/2}\psi$ for some $r \geq 0$ leads to

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\|\Delta^{\frac{r}{2}}\theta'\|_{L^2}^2 &\lesssim -\|\Delta^{\frac{r}{2}+1}\psi\|_{L^2}^2 \\ &\lesssim K^3\|\Delta^{\frac{r}{2}+4}\partial_x^{-3}\psi\|_{L^2}^2 - K\|\partial_x^{-1}\Delta^{\frac{r}{2}+2}\psi\|_{L^2}^2 \\ &\lesssim K^3\|\Delta^{\frac{r}{2}+2}\partial_x^{-2}\theta'\|_{L^2}^2 - K\|\Delta^{\frac{r}{2}}\theta'\|_{L^2}^2, \end{aligned}$$

recalling that $\Delta^2\psi = \partial_x\theta'$. Taking $K \simeq (1+t)^{-1}$ we deduce

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\Delta^{\frac{r}{2}}\theta'\|_{L^2}^2 + \frac{3}{1+t}\|\Delta^{\frac{r}{2}}\theta'\|_{L^2}^2 \lesssim \frac{1}{(1+t)^3}\|\Delta^{\frac{r}{2}+2}\partial_x^{-2}\theta'\|_{L^2}^2. \tag{1-11}$$

Note that here the factor 3 could be made arbitrarily large by taking a larger multiplicative constant in $K$, but any constant strictly larger than 2 is sufficient for the argument.

Since $\Delta^{r/2+2}\partial_x^{-2}\theta'$ is uniformly bounded in $L^2(\mathbb{T}^2)$ by $\|\partial_x^{-2}\theta_0'\|_{H^{r+4}}$, this integrates into

$$\forall t \geq 0, \quad \|\Delta^{\frac{r}{2}}\theta'(t)\|_{L^2} \lesssim \frac{\|\partial_x^{-2}\theta_0'\|_{H^{r+4}}}{1+t}. \tag{1-12}$$

Note that the index of regularity $r$ is arbitrary. Hence, plugging this estimate back into (1-11) and using an induction argument, we find that, for any $\alpha \geq 0,\ r \geq 0$,

$$\forall t \geq 0, \quad \|\Delta^{\frac{r}{2}}\theta'(t)\|_{L^2} \leq (1+t)^{-\frac{\alpha}{2}}\|\partial_x^{-\alpha}\theta_0'\|_{H^{r+2\alpha}}. \tag{1-13}$$

Let us emphasize that when $G = 0$, this estimate can be proved directly from the Fourier representation formula

$$\theta'(t, x, z) = \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^2 \\ k_x\neq 0}} \hat{\theta}_{\boldsymbol{k}}(0)\exp\left(-\frac{k_x^2}{|\boldsymbol{k}|^4}t\right)\exp(i\boldsymbol{k}\cdot(x, z)).$$

Hence the decay rate can be expected to be somewhat optimal. Moreover, in the case of the torus, the rate of decay can be as large as desired, the cost being the regularity required on $\theta_0'$. Note that, for $r = 0$, $\alpha = 2$, we find the decay rate announced in Theorem 1.1.

***Difficulties with Dirichlet boundary conditions.*** Let us now explain the main differences between (1-8) on $\mathbb{T}^2$ and the original system (1-7) on $\Omega = \mathbb{T} \times (0, 1)$. The strategy will be identical. We first prove a uniform bound for $\theta'$ in $H^4(\Omega)$, and then use interpolation inequalities together with the energy estimate to obtain the decay estimate (1-12). However the derivation of the different bounds will be substantially more involved.

We shall prove the uniform $H^4(\Omega)$ bound for $\theta'$ directly from the equation without spectral analysis. More precisely, the estimate (1-9) remains valid for $s = 0$ since $\psi|_{\partial\Omega} = \partial_{\boldsymbol{n}}\psi|_{\partial\Omega} = 0$. Higher-order uniform estimates in $H^s(\Omega)$ fail in general due to nonvanishing terms on the boundary. The question is therefore when the integration by parts done in (1-9) can be performed. The traces of $\theta'$ and $\partial_{\boldsymbol{n}}\theta'$ being zero, the traces of $\Delta^2\psi$ and $\partial_{\boldsymbol{n}}\Delta^2\psi$ are also vanishing (see Section 2.1) so integrations by parts in (1-9) can be done for $s = 4$ provided $G = 0$. Therefore a uniform $H^4$ bound can be deduced when $G = 0$. When $G$ is nonzero, some traces no longer vanish. The strategy will be to treat them perturbatively, i.e., not performing integration by parts on $\Delta^2(G\psi)\Delta^3\psi$. A similar interpolation argument (Lemma 2.7) allows us to then deduce the analogue of (1-12), i.e., that $\|\theta'\|_{L^2}$ is bounded by $(1+t)^{-1}$.

However, note that the higher decay (1-13) for $r > 0$ and $\alpha > 2$ does not hold in general, as Theorem 1.3 shows. Indeed, the decay rate is prescribed by the boundary layer part of the solution, for which we have $\|\Delta^{r/2}\theta^{\mathrm{BL}}\|_{L^2} \propto (1+t)^{-1+r/4-1/8}$. Hence the $H^s$ norm of $\theta'$ for $s \geq 5$ is not expected to be bounded.

Finally, let us mention that proving some time integrability on the velocity field is crucial in order to obtain the convergence of $\bar{\theta}$. As a consequence, the linear decay from (1-12) is not entirely sufficient to complete the proof of Theorem 1.1. In previous works, this higher decay on the velocity field was obtained either thanks to high-regularity bounds, or by taking advantage of the Fourier representation of the solution. Since none of these tools are available here, we rely on a different argument, involving bounds on the time derivative of $\theta'$.

**Remark 1.5** (about the spectral decomposition). Since the equation is no longer set on the torus, but rather in the domain $\Omega = \mathbb{T} \times (0, 1)$ endowed with boundary conditions, we can no longer perform a (discrete) Fourier transform in the vertical variable. However it is possible to analyze explicitly the eigenfunctions of the operator

$$L : \theta \in L^2 \mapsto \partial_x \psi \in L^2, \quad \text{where } \Delta^2 \psi = \partial_x \theta,\ \psi|_{\partial\Omega} = \partial_{\boldsymbol{n}} \psi|_{\partial\Omega} = 0, \tag{1-14}$$

and show that the eigenvalues $(\lambda_{\boldsymbol{k}})_{\boldsymbol{k} \in \mathbb{Z} \times \mathbb{N}}$ of the operator $L$ behave asymptotically as $k_x^2/|\boldsymbol{k}|^4$ (see Lemma B.2), so that the estimate (1-12) remains true. Details on the spectral analysis are presented in [Leblond 2023].

***Bootstrap.*** The last step of the proof consists in bringing the previous linear analysis into the full nonlinear system. Intuitively, the strategy is the following: denote by $(0, T^*)$ the maximal time interval over which $\|\theta'\|_{L^2} \le B(1+t)^{-1}$ and $\|\theta'\|_{H^4} \le B$ are valid with a constant $B$. In fact more estimates need to be included in the bootstrap argument for technical reasons; see (2-9). On this time interval, the quadratic terms can be treated perturbatively, provided $\|\theta_0\|_{H^4}$ is sufficiently small. Hence the bootstrap estimates hold with a constant which is better than $B$, and thus $T^* = \infty$. It follows that $\theta'$ converges towards zero in $L^2$, and that the time derivative of $\bar{\theta}$ is integrable. Hence $\bar{\theta}$ has a limit in $L^2$ as $t \to \infty$. This is the main part of the proof which is detailed in Section 2.3.

***Identification of the limit.*** Since $\theta'$ converges to zero in any $H^m$ for $m < 4$ as $t \to \infty$ and $\bar{\theta}$ has a limit in $L^2$ as $t \to \infty$, the whole density $\rho = \rho_s + \theta = \rho_s + \bar{\theta} + \theta'$ converges to some limit $\rho_\infty = \rho_s + \bar{\theta}_\infty$ in $L^2$ and $\rho_\infty$ depends only on $z$. The term $\partial_z \theta$ is small compared to $\partial_z \rho_s = -1$, and so is its limit $\partial_z \bar{\theta}_\infty$. Whence $\rho_\infty$ is strictly decreasing with respect to $z$, as is $\rho_s$. The transport of the density by the divergence-free field $\boldsymbol{u}$ ensures that the level sets of $\rho$ are preserved by the time evolution, and by strong convergence this is also the case for the limit $\rho_\infty$. According to rearrangement theory, $\rho_\infty$ is therefore a rearrangement of $\rho_0$. One can show that there exists a unique decreasing vertical rearrangement of $\rho_0$; hence $\rho_\infty$ is uniquely determined. This part of the proof is detailed in Section 2.4.

***Linear boundary layers for system* (1-5).** Let us now give a sketch of the proof of Theorem 1.2. We start with rather simple observations:

- First, it follows from the equation that $\partial_t \theta|_{\partial\Omega} = \partial_t \partial_{\boldsymbol{n}} \theta|_{\partial\Omega} = 0$. Therefore, for all $t \ge 0$,

$$\theta|_{\partial\Omega}(t) = \theta_0|_{\partial\Omega}, \quad \partial_{\boldsymbol{n}} \theta|_{\partial\Omega}(t) = \partial_{\boldsymbol{n}} \theta_0|_{\partial\Omega}.$$

- Taking the horizontal average of the evolution equation, we find that $\partial_t \bar{\theta} = 0$, and thus $\bar{\theta}(t) = \bar{\theta}_0$. Hence we focus on the long-time behavior of $\theta'$.

- Let us denote by $(b_{\boldsymbol{k}})_{\boldsymbol{k} \in \mathbb{Z} \times \mathbb{N}^*}$ the basis of eigenvectors of the operator $L$ defined in (1-14) (see Lemma B.2). Then we can always write

$$\theta'(t) = \sum_{\boldsymbol{k} \in \mathbb{Z}^* \times \mathbb{N}} \exp(-\lambda_{\boldsymbol{k}} t) \hat{\theta}'_{\boldsymbol{k}}(t=0) b_{\boldsymbol{k}},$$

with $(\hat{\theta}'_{\boldsymbol{k}}(t=0))_{\boldsymbol{k} \in \mathbb{Z}^* \times \mathbb{N}} \in \ell^2$. We recall that $\lambda_{\boldsymbol{k}}$ behaves asymptotically as $|k_x|^2/|\boldsymbol{k}|^4$. It then follows from Lebesgue's dominated convergence theorem that $\theta'(t) \to 0$ in $L^2$ as $t \to \infty$.

Therefore $\theta'(t)$ vanishes in $L^2$ while keeping a constant — and nonzero — value on the boundary. As a consequence, it is reasonable to expect that boundary layers are formed in the vicinity of $z=0$ and $z=1$ as $t\to\infty$. We then plug the ansatz (1-6) into (1-5) and identify the profiles $\Theta^0_{\text{top}}$, $\Theta^0_{\text{bot}}$. The role of $\Theta^0_{\text{top}}$ (resp. of $\Theta^0_{\text{bot}}$) is to lift the trace of $\theta_0'$ at the top boundary $z=1$ (resp. at the bottom boundary $z=0$). We find that these two profiles satisfy an ODE, with boundary conditions given by

$$\Theta^0_{\text{top}}(x, Z{=}0) = \theta_0'(x, z{=}1), \quad \partial_Z \Theta^0_{\text{top}}(x, Z{=}0) = 0,$$
$$\Theta^0_{\text{bot}}(x, Z{=}0) = \theta_0'(x, z{=}0), \quad \partial_Z \Theta^0_{\text{bot}}(x, Z{=}0) = 0.$$

In a similar way, the next-order boundary layer terms $\Theta^1_{\text{top}}$ and $\Theta^1_{\text{bot}}$ lift the traces of $\partial_{\boldsymbol{n}}\theta_0$ on $\partial\Omega$. Hence the first step is to construct explicitly the boundary layer profiles in terms of $\theta_0$. By construction, the remainder $\theta' - \theta^{\text{BL}}$ vanishes on the boundary, together with its normal derivative. We can then apply the decay analysis presented above to the remainder $\theta' - \theta^{\text{BL}}$, and we find that $\|(\theta' - \theta^{\text{BL}})(t)\|_{L^2} \lesssim (1+t)^{-1}$.

***Boundary layers for system* (1-1a).** We now turn towards Theorem 1.3. Note that the boundary layer term in Theorem 1.3 is smaller than in (1-6). This is directly linked to the fact that under the assumptions of Theorem 1.3, $\theta = \rho - \rho_s$ vanishes on the boundary, together with its normal derivative. Therefore, the boundary layer term $\theta^{\text{BL}}$ (or rather $\Delta^2\theta^{\text{BL}}$) now lifts the traces of $\Delta^2\theta'$ and $\partial_{\boldsymbol{n}}\Delta^2\theta'$. The overall strategy is the same as the one described above: we first identify the boundary-layer part of the solution by rigorously constructing the boundary layer profiles $\Theta^j_{\text{bot}}$ and $\Theta^j_{\text{top}}$. We then prove some decay estimates on the remainder $\theta^{\text{rem}} = \theta' - \theta^{\text{BL}}$, noticing that $\Delta^2\theta^{\text{rem}}$ satisfies assumptions that are very close to the ones of Theorem 1.1. Note that the higher decay we obtain on $\theta^{\text{rem}}$ is the main reason behind the strong regularity requirements on $\rho_0$.

However, there are several new conceptual and technical difficulties compared with Theorem 1.2. The main one lies in the fact that the traces $\Delta^2\theta'|_{\partial\Omega}$ and $\partial_{\boldsymbol{n}}\Delta^2\theta'|_{\partial\Omega}$ are not constant with respect to time. They merely have a finite limit as $t\to\infty$. Hence we need to find an asymptotic expansion in powers of $(1+t)^{-1/4}$ for $\Delta^2\theta'|_{\partial\Omega}$ and $\partial_{\boldsymbol{n}}\Delta^2\theta'|_{\partial\Omega}$ as $t \gg 1$. The main boundary layer profiles $\Theta^0_{\text{bot}}$ and $\Theta^0_{\text{top}}$ will lift the first term in this expansion (i.e., the long-time limit of $\Delta^2\theta'|_{\partial\Omega}$), whereas the next-order profiles $\Theta^j_{\text{top}}$, $\Theta^j_{\text{bot}}$ for $j\geq 2$ will lift the lower-order terms. Furthermore, in order to prove that $\Delta^2\theta'|_{\partial\Omega}$ converges in $H^s$ as $t\to\infty$ for some sufficiently large $s$, we shall need high-regularity bounds on $\theta'$. Eventually, the proof of Theorem 1.3 involves two nested bootstrap arguments: one on $\theta'$, which allows us to construct the boundary layer term $\theta^{\text{BL}}$ on the interval on which the bootstrap assumption is satisfied, and a second one on the remainder $\theta' - \theta^{\text{BL}}$, on a possibly smaller interval.

***Notation.*** Throughout the paper, we write $A \lesssim B$ whenever there exists a universal positive constant $C$ such that $A \leq CB$.

## 2. Long-time stability of stratified profiles: proof of Theorem 1.1

This section is devoted to the proof of Theorem 1.1. The proof follows the steps highlighted in the Introduction: we decompose $\theta$ into $\theta = \bar\theta + \theta'$, and we prove that $\theta'$ vanishes in $L^2$ with algebraic decay, while $\bar\theta$ converges in $L^2$ towards a profile $\bar\theta_\infty(z)$. To that end, we first study the linearized Stokes-transport system around a solution $\theta$ close to an affine profile. Thanks to a crucial interpolation inequality (see Lemma 2.7), which somehow replaces the spectral decomposition in the periodic setting, we quantify the $L^2$ decay of solutions of the linearized equation with a source term (see Proposition 2.6). We then use a bootstrap argument to show that the decay predicted by the linear analysis persists for the nonlinear evolution. This allows us to prove that $\theta'(t) \to 0$ and that $\bar{\theta} \to \bar{\theta}_\infty$ in $H^s(\Omega)$ as $t \to \infty$ for all $s < 4$. Eventually, we identify the asymptotic profile $\bar{\theta}_\infty$ in terms of the initial data.

The organization of this section is the following. We start in Section 2.1 with some preliminary remarks concerning the traces of $\theta$ and $\partial_{\boldsymbol{n}}\theta$. We then turn towards the analysis of the linearized system in Section 2.2. The bootstrap argument is presented in Section 2.3. Eventually, we prove in Section 2.4 that $\rho_\infty$ is the rearrangement of the initial data $\rho_0$.

**2.1.** ***Vanishing traces for $\theta'$ and $\partial_{\boldsymbol{n}}\theta'$.*** We prove here the following preliminary result:

**Lemma 2.1.** *Let $\theta_0 \in H^m(\Omega)$ with $m \geq 3$, and let $\theta \in L^\infty_{\mathrm{loc}}(\mathbb{R}_+, H^m)$ be the solution of* (1-7). *Assume that $\theta_0 = \partial_{\boldsymbol{n}}\theta_0 = 0$ on $\partial\Omega$. Then, for all $t \geq 0$,*

$$\theta(t)|_{\partial\Omega} = \partial_{\boldsymbol{n}}\theta(t)|_{\partial\Omega} = 0.$$

*If additionally $\partial_z^2\bar{\theta}_0 = 0$ on $\partial\Omega$, then $\partial_z^2\bar{\theta}(t)|_{\partial\Omega} = 0$ for all $t \geq 0$.*

**Remark 2.2.** If $\rho_0 \in H_0^m(\Omega)$ then the solution $\rho(t)$ of (1-1a) belongs to $H_0^m(\Omega)$ for all times. Indeed, the solution of the transport equation can be written as

$$\rho(t) = \rho_0(\boldsymbol{X}(t)^{-1}),$$

where $\boldsymbol{X} : \mathbb{R}_+ \times \Omega \to \Omega$ is the characteristic function associated to $\boldsymbol{u}$, defined as the solution of the ordinary differential equation

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \boldsymbol{u}(t, \boldsymbol{X}(t)), \\ \boldsymbol{X}(0) = \mathrm{Id}_\Omega. \end{cases}$$

We recall that $\boldsymbol{X}(t)$ is a diffeomorphism of $\Omega$ for all times $t \in \mathbb{R}_+$. Since $\boldsymbol{u}(t) \in H_0^1(\Omega)$ due to the homogeneous Dirichlet condition, the boundary $\partial\Omega$ is stable for the characteristic function at all times $t > 0$. In other words, $\boldsymbol{X}(t)|_{\partial\Omega} = \mathrm{Id}_{\partial\Omega}$, and consequently $\boldsymbol{X}(t)^{-1}|_{\partial\Omega} = \mathrm{Id}_{\partial\Omega}$. It follows that if $\rho_0 \in H_0^1(\Omega)$, then $\rho(t)|_{\partial\Omega} = 0$ for all $t \geq 0$. The claim for higher values of $m$ follows easily by induction.

Note that the assumptions of Lemma 2.1 are different since $\rho_0 = 1 - z + \theta_0$ does not vanish on the boundary.

*Proof.* We recall (see Theorem A.1) that $\theta \in C(\mathbb{R}_+, H^m) \cap C^1(\mathbb{R}_+, H^{m-1})$. Therefore, taking the trace of (1-1a), we get

$$\partial_t\theta|_{\partial\Omega} + \boldsymbol{u}|_{\partial\Omega} \cdot \nabla\theta|_{\partial\Omega} = u_z|_{\partial\Omega},$$

where $\boldsymbol{u}|_{\partial\Omega} = \boldsymbol{0}$. Hence $\partial_t\theta|_{\partial\Omega} = 0$ and the trace of $\theta$ is constant in time, equal to 0. Since horizontal derivatives preserve this property, we even have $\partial_x^\ell\theta|_{\partial\Omega} = 0$ for any $\ell$. Let us now consider the normal derivative. We recall that $\partial_{\boldsymbol{n}}$ coincides (up to a sign) with $\partial_z$. Applying one vertical derivative to the equation,

$$\partial_t \partial_z \theta + \partial_z \boldsymbol{u} \cdot \nabla\theta + \boldsymbol{u} \cdot \nabla \partial_z \theta = \partial_z u_z,$$

where

$$\partial_z \boldsymbol{u}|_{\partial\Omega} \cdot \nabla\theta|_{\partial\Omega} = \partial_z u_x|_{\partial\Omega}\, \partial_x \theta|_{\partial\Omega} + \partial_z u_z|_{\partial\Omega}\, \partial_z\theta|_{\partial\Omega}.$$

We recall that $\partial_x\theta|_{\partial\Omega} = 0$ and we use the divergence-free condition to observe $\partial_z u_z|_{\partial\Omega} = -\partial_x u_x|_{\partial\Omega} = 0$. In the end we get $\partial_z\theta|_{\partial\Omega} = 0$ for all times; hence $\theta \in H^2_0(\Omega)$. Trying to go further, applying the same ideas, we get

$$\partial_t \partial_z^2 \theta|_{\partial\Omega} = \partial_z^2 u_z|_{\partial\Omega}.$$

However, $\partial_z^2 u_z$ does not vanish on $\partial\Omega$, and therefore we cannot iterate the argument. Nevertheless, we get

$$\partial_t \partial_z^2 \bar{\theta}|_{\partial\Omega} = \frac{1}{2\pi} \int_{\mathbb{T}} \partial_z^2 u_z|_{\partial\Omega} = -\frac{1}{2\pi} \int_{\mathbb{T}} \partial_x \partial_z u_x|_{\partial\Omega} = 0.$$

Note that for higher orders of derivation, we cannot infer any cancellation in general. □

**Definition 2.3.** In the rest of the paper, we will set

$$G(t, z) = \partial_z \bar{\theta}(t, z).$$

Under the assumptions of Lemma 2.1, we infer that $G|_{\partial\Omega} = \partial_z G|_{\partial\Omega} = 0$.

**2.2. *Study of the linearized system.*** This subsection is concerned with the study of the linear system

$$\begin{cases} \partial_t\theta' = (1-G)\partial_x\psi + S \quad \text{in } (0,+\infty)\times\Omega, \\ \Delta^2\psi = \partial_x\theta' \quad \text{in } (0,+\infty)\times\Omega, \\ \psi|_{\partial\Omega} = \partial_{\boldsymbol{n}}\psi|_{\partial\Omega} = 0, \quad \theta'|_{t=0} = \theta'_0, \end{cases} \tag{2-1}$$

which is satisfied by $(\theta', \psi)$ in the first place, with $G = \partial_z\bar{\theta}$ and $S = -(\nabla^\perp\psi \cdot \nabla\theta')'$. It will also be satisfied for various derivatives of $(\theta', \psi)$ with different source terms $S$. The term $G$ will always be $\partial_z\bar{\theta}$.

Our goal is to analyze the long-time behavior of $\theta'$, under suitable decay assumptions on $S$. For later purposes, we have decomposed our results into several separate statements, whose proofs are postponed to the end of the section. The first one is a uniform $L^2$ bound on the solutions when the source term is time integrable:

**Lemma 2.4** (uniform $L^2$ bound on solutions of the linearized system). *Let $G \in L^\infty(\mathbb{R}_+, H^2)$, $S \in L^\infty(\mathbb{R}_+, L^2)$, and $\theta'_0 \in L^2$. Let $\theta' \in L^\infty(\mathbb{R}_+, L^2)$ be the unique solution of* (2-1). *Assume that $S$ can be decomposed as $S = S_\perp + S_\parallel$ satisfying for some $\sigma, \delta > 0$ and any $t \geq 0$*

$$\int_\Omega S_\perp(t,x)\theta'(t,x)\,\mathrm{d}x = 0, \quad \|S_\parallel(t)\|_{L^2} \lesssim \frac{\sigma}{(1+t)^{1+\delta}}. \tag{2-2}$$

*Thus, there exists a universal constant $\gamma_0 \in (0,1)$ such that if*

$$\|G\|_{H^2} \leq \gamma_0, \tag{2-3}$$

*then*

$$\|\theta'\|_{L^2} \leq \|\theta'_0\|_{L^2} + C_\delta\sigma.$$

**Remark 2.5.** The term $S_\perp$ will often have the structure $S_\perp = \boldsymbol{u}\cdot\nabla\theta'$: indeed, provided $\boldsymbol{u}$ and $\theta'$ have sufficient regularity, the divergence-free and homogeneous Dirichlet conditions ensure that

$$\int_\Omega (\boldsymbol{u}\cdot\nabla\theta')\theta' = \frac{1}{2}\int_\Omega \boldsymbol{u}\cdot\nabla|\theta'|^2 = -\frac{1}{2}\int_\Omega \operatorname{div}\boldsymbol{u}|\theta'|^2 + \frac{1}{2}\int_{\partial\Omega}\boldsymbol{u}\cdot\boldsymbol{n}|\theta'|^2 = 0.$$

Our second result, which is at the core of Theorem 1.1, gives a quantitative algebraic decay on $\theta'$:

**Proposition 2.6.** *There exists a universal constant $\gamma_0 \in (0,1)$ such that the following result holds. Let $T > 0$, $G \in L^\infty(\mathbb{R}_+, H^2)$, $S \in L^\infty(\mathbb{R}_+, L^2)$, and $\theta_0' \in L^2$ such that $\partial_x^{-2}\theta_0' \in H^4$. Let $\theta' \in C(\mathbb{R}_+, H^2)$ be the unique solution of* (2-1). *Assume that $\theta'$ and $\partial_{\boldsymbol{n}}\theta'$ vanish on $\partial\Omega$, and that $S$ decomposes into $S = S_\perp + S_\parallel + S_\Delta$ with, for some $\sigma, \delta > 0$ and all $t \in [0,T]$,*

$$\int_\Omega S_\perp(t,\boldsymbol{x})\theta'(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0, \quad \|S_\parallel(t)\|_{L^2} \le \frac{\sigma}{(1+t)^{1+\delta}}, \quad \|S_\Delta(t)\|_{L^2} \lesssim \frac{\|\Delta\psi\|_{L^2}}{(1+t)^{\frac{1}{2}}}. \tag{2-4}$$

*Assume moreover that $G$ satisfies* (2-3), *and that there exist $A, \alpha \ge 0$ such that for all $t \in [0,T]$*

$$\|\Delta^2\partial_x^{-2}\theta'(t)\|_{L^2} \le \frac{A}{(1+t)^\alpha}. \tag{2-5}$$

*Then*

$$\|\theta'(t)\|_{L^2} \lesssim \frac{\|\theta_0'\|_{L^2} + A + \sigma}{(1+t)^{\min(1+\alpha,\delta)}} \quad \forall t \in [0,T].$$

In order to prove this quantitative decay, we shall need to analyze the structure of the dissipation term

$$-\int_\Omega \partial_x\psi\theta' = \int_\Omega |\Delta\psi|^2.$$

In previous works for different but related models [Castro et al. 2019a], at this stage, an explicit spectral decomposition of the solution was used, relying on Fourier series. Note that such a spectral decomposition is also available for the operator $\Delta^{-2}\partial_x^2$ (see Lemma B.2). However, since we cannot interpolate for an arbitrary regularity, we choose here to use a different approach. We replace this spectral analysis with the following result, which can be seen as an interpolation lemma. It is noteworthy that in spite of its deceitfully simple form (and proof), this lemma provides the correct scaling for the solutions.

**Lemma 2.7.** *For any $\ell \ge 0$, and for all $\psi \in H^{8+\ell-2}(\Omega)$ satisfying*

$$\Delta^2\psi|_{\partial\Omega} = \partial_{\boldsymbol{n}}\Delta^2\psi|_{\partial\Omega} = 0\,,$$

*we have for all $K > 0$*

$$\|\partial_x^{\ell-1}\Delta^2\psi\|_{L^2}^2 \lesssim \frac{1}{K}\|\partial_x^\ell\Delta\psi\|_{L^2}^2 + K^2\|\partial_x^{\ell-3}\Delta^4\psi\|_{L^2}^2.$$

*Proof.* Since $\Delta^2\psi$ and $\partial_{\boldsymbol{n}}\Delta^2\psi$ vanish on the boundary $\partial\Omega$, we have after three integrations by parts

$$\begin{aligned}\|\partial_x^{\ell-1}\Delta^2\psi\|_{L^2}^2 &= -\int_\Omega \partial_x^\ell\Delta^2\psi\,\partial_x^{\ell-2}\Delta^2\psi = \int_\Omega \nabla\partial_x^\ell\Delta\psi\cdot\nabla\partial_x^{\ell-2}\Delta^2\psi \\ &= -\int_\Omega \partial_x^\ell\Delta\psi\,\partial_x^{\ell-2}\Delta^3\psi \le \|\partial_x^\ell\Delta\psi\|_{L^2}\|\partial_x^{\ell-2}\Delta^3\psi\|_{L^2}.\end{aligned}$$

On another hand, we also have by integrations by parts

$$\begin{aligned}\|\partial_x^{\ell-2}\Delta^3\psi\|_{L^2}^2 &= -\int_\Omega \partial_x^{\ell-1}\Delta^3\psi\,\partial_x^{\ell-3}\Delta^3\psi = \int_\Omega \nabla\partial_x^{\ell-1}\Delta^2\psi\cdot\nabla\partial_x^{\ell-3}\Delta^3\psi\\ &= -\int_\Omega \partial_x^{\ell-1}\Delta^2\psi\,\partial_x^{\ell-3}\Delta^4\psi \le \|\partial_x^{\ell-1}\Delta^2\psi\|_{L^2}\|\partial_x^{\ell-3}\Delta^4\psi\|_{L^2}.\end{aligned}$$

Hence, using the second bound in the first inequality, we obtain

$$\|\partial_x^{\ell-1}\Delta^2\psi\|_{L^2}^2 \le \|\partial_x^{\ell}\Delta\psi\|_{L^2}\,\|\partial_x^{\ell-1}\Delta^2\psi\|_{L^2}^{\frac{1}{2}}\,\|\partial_x^{\ell-3}\Delta^4\psi\|_{L^2}^{\frac{1}{2}}.$$

Gathering the similar terms on the left-hand side and applying Young's inequality yields, for any constant $K > 0$,

$$\begin{aligned}\|\partial_x^{\ell-1}\Delta^2\psi\|_{L^2}^2 &\le \|\partial_x^{\ell}\Delta\psi\|_{L^2}^{\frac{4}{3}}\,\|\partial_x^{\ell-3}\Delta^4\psi\|_{L^2}^{\frac{2}{3}}\\ &\lesssim (K^{-\frac{2}{3}}\|\partial_x^{\ell}\Delta\psi\|_{L^2}^{\frac{4}{3}})^{\frac{3}{2}} + (K^{\frac{2}{3}}\|\partial_x^{\ell-3}\Delta^4\psi\|_{L^2}^{\frac{2}{3}})^3\\ &\lesssim \frac{1}{K}\|\partial_x^{\ell}\Delta\psi\|_{L^2}^2 + K^2\|\partial_x^{\ell-3}\Delta^4\psi\|_{L^2}^2.\end{aligned}$$

□

Let us now turn towards the proof of Lemma 2.4 and Proposition 2.6.

*Proof of Lemma 2.4.* The energy estimate in (2-1) writes

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\theta'\|_{L^2}^2 = \int_\Omega (1-G)\partial_x\psi\theta' + \int_\Omega S\theta'.$$

A few integrations by parts provide, since $\psi|_{\partial\Omega} = \partial_{\boldsymbol{n}}\psi|_{\partial\Omega} = 0$,

$$\int_\Omega (1-G)\partial_x\psi\theta' = -\int_\Omega (1-G)\psi\Delta^2\psi = -\int_\Omega \Delta((1-G)\psi)\Delta\psi. \tag{2-6}$$

Using the Sobolev embeddings $H^2 \subset L^\infty$ and $H^2 \subset W^{1,4}$, we get

$$\int_\Omega (1-G)\partial_x\psi\theta' \le -(1 - C\|G\|_{H^2})\|\Delta\psi\|_{L^2}^2.$$

At this point we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\theta'\|_{L^2}^2 \le -(1-C\gamma_0)\|\Delta\psi\|_{L^2}^2 + \int_\Omega S\theta' = -(1-C\gamma_0)\|\Delta\psi\|_{L^2}^2 + \int_\Omega S_\parallel\theta'. \tag{2-7}$$

So if $\gamma_0$ is small enough, in a universal way, the first term in the right-hand side is nonpositive. Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\theta'(t)\|_{L^2} \le \|S_\parallel(t)\|_{L^2} \le \frac{\sigma}{(1+t)^{1+\delta}}$$

and since $\delta > 0$ this inequality integrates as

$$\|\theta'\|_{L^2} \le \|\theta_0'\|_{L^2} + C_\delta\sigma.$$

□

*Proof of Proposition 2.6.* Back to (2-7) and plugging the decomposition of $S$ we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\theta'\|_{L^2}^2 + (1-C\gamma_0)\|\Delta\psi\|_{L^2}^2 \le (\|S_\parallel\|_{L^2} + \|S_\Delta\|_{L^2})\|\theta'\|_{L^2}.$$

Then assumption (2-4) and Young's inequality provide

$$\|S_\parallel\|_{L^2}\|\theta'\|_{L^2} \le \frac{\sigma}{(1+t)^{\frac{1}{2}+\delta}}\frac{\|\theta'\|_{L^2}}{(1+t)^{\frac{1}{2}}} \lesssim \frac{\sigma^2}{(1+t)^{1+2\delta}} + \frac{\|\theta'\|_{L^2}^2}{1+t},$$
$$\|S_\Delta\|_{L^2}\|\theta'\|_{L^2} \lesssim \|\Delta\psi\|_{L^2}\frac{\|\theta'\|_{L^2}}{(1+t)^{\frac{1}{2}}} \lesssim \gamma_0\|\Delta\psi\|_{L^2}^2 + \frac{1}{\gamma_0}\frac{\|\theta'\|_{L^2}^2}{1+t}.$$

Hence if $\gamma_0$ is small enough, the dissipative term $\gamma_0\|\Delta\psi\|^2$ can be absorbed, and we have, for some $c_0 \in (0, 1)$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\theta'\|_{L^2}^2 + c_0\|\Delta\psi\|_{L^2}^2 \lesssim \frac{\sigma^2}{(1+t)^{1+2\delta}} + \frac{\|\theta'\|_{L^2}^2}{1+t}.$$

We now use the interpolation lemma, Lemma 2.7, with $\ell = 0$, recalling that $\Delta^2\psi|_{\partial\Omega} = \partial_x\theta'|_{\partial\Omega} = 0$ and $\partial_{\boldsymbol{n}}\Delta^2\psi|_{\partial\Omega} = \partial_{\boldsymbol{n}}\partial_x\theta'|_{\partial\Omega} = 0$. Choosing $K = \kappa/c_0(1+t)^{-1}$ with $\kappa > 0$ arbitrary large independently of the data, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\theta'\|_{L^2}^2 + \frac{\kappa}{1+t}\|\theta'\|_{L^2}^2 \lesssim \kappa^3\frac{\|\Delta^2\partial_x^{-2}\theta'\|_{L^2}^2}{(1+t)^3} + \frac{\sigma^2}{(1+t)^{1+2\delta}}.$$

Plugging assumption (2-5) provides

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\theta'\|_{L^2}^2 + \frac{\kappa}{1+t}\|\theta'\|_{L^2}^2 \lesssim \frac{(\kappa^3A^2+\sigma^2)}{(1+t)^{\min(3+2\alpha,1+2\delta)}},$$

which for a suitable choice of $\kappa$ integrates into

$$\|\theta'(t)\|_{L^2} \lesssim \frac{\|\theta_0'\|_{L^2} + A + \sigma}{(1+t)^{\min(1+\alpha,\delta)}} \quad \forall t \in [0, T]. \qquad \square$$

**Remark 2.8** (stability for more general stationary profiles). Let us now explain how our results can be generalized to other stably stratified profiles $\rho_s$. Let $\rho_s \in H^6(0, 1)$ such that $\sup \partial_z\rho_s < 0$. The linear evolution equation on $\theta'$ can be written as

$$\partial_t\theta' = -\partial_z\rho_s\partial_x\psi.$$

Multiplying the above equation by $-\theta'/\partial_z\rho_s$, we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \frac{1}{-\partial_z\rho_s}\theta'(t,x)^2\,\mathrm{d}x = \int_\Omega \partial_x\psi\theta' = -\int_\Omega |\Delta\psi|^2.$$

Since $\partial_z\rho_s$ is bounded from above and below by negative constants, $-\int_\Omega(\theta')^2/\partial_z\rho_s$ is equivalent to the $L^2$ norm, and the main linear estimate remains the same. However, additional commutators stem from the nonlinear terms when we use the weight $-1/\partial_z\rho_s$. For instance,

$$\int_\Omega \boldsymbol{u}\cdot\nabla\theta'\frac{\theta'}{-\partial_z\rho_s} = \frac{1}{2}\int_\Omega(\theta')^2\boldsymbol{u}\cdot\nabla\frac{1}{\partial_z\rho_s}.$$

These commutators will enter the terms $S_\parallel$ and $S_\perp$ from Lemma 2.4 and Proposition 2.6. Also, the constant $\gamma_0$ involved in the smallness condition (2-3) on $G$ will now depend on $\partial_z\rho_s$, but the result still holds. We leave the details to the reader, and stick to the case of a linearly stratified profile for simplicity.

Let us now consider the case of nonstably stratified profiles. First, if $\partial_z \rho_s > 0$, we have the opposite sign in front of the dissipative term $\int_\Omega |\Delta\psi|^2$: at the linearized level, the perturbation grows. For the nonlinear equation, starting from a small perturbation, $\|\theta'(t)\|_{L^2}$ will have a transient growth for small times, until its norm becomes of order 1 and the nonlinear term can no longer be neglected. In fact, when $\partial_z \rho_s$ is constant — say $\partial_z \rho_s = 1$ — the equation satisfied by $\theta$ is

$$\partial_t \theta + \boldsymbol{u} \cdot \nabla \theta + u_2 = 0,$$

and therefore

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\theta(t)\|_{L^2}^2 = -\int_\Omega \boldsymbol{u}_2 \rho = \int_\Omega (\Delta\psi)^2 \geq 0.$$

In this case, the $L^2$ norm of $\theta$ is increasing on the whole interval $[0, +\infty)$, but remains bounded since the $L^2$ norm of $\rho$ is conserved.

Let us indicate a few facts about the long-time behavior of solutions in the general case $\rho_0 \in L^\infty$ (see Section 3.6 in [Leblond 2023] for a proof of these results). The velocity $\boldsymbol{u}$ belongs to $L^2([0, +\infty), H_0^1(\Omega))$. Furthermore the $\omega$-limit set in $H^{-1}$ of $\rho_0$ is nonempty and contained in the set of stratified rearrangements of $\rho_0$. However it is not known in general whether this $\omega$-limit set is a singleton.

We now go back to the case where $\rho_0$ is a small perturbation of a stratified state $\rho_s$. When $\partial_z \rho_s$ is not of constant sign, we cannot conclude a priori, even at the linearized level. Indeed, if a density profile is not monotonous, then we cannot guarantee the proper sign in front of the integral $\partial_z \rho |\Delta\psi|^2$. Therefore, if a profile admits a nonmonotonous function in any of its neighborhoods, in arbitrary high regularity, and in particular in $H^6$, the proof of our stability result does not hold. In particular, when $\partial_z \rho_s \leq 0$ and $\partial_z \rho_s$ vanishes at a single point $z_0 \in (0, 1)$, this point is also an inflection point and therefore $\rho_s(z) = \rho_s(z_0) + O((z - z_0)^3)$ in a neighborhood of $z_0$. Perturbing $\rho_s$ by a function of the type $\varepsilon(z - z_0)^2 \chi(z - z_0)$ with a cut-off function $\chi \in C_0^\infty(\mathbb{R})$ breaks the monotony. Thus, even in the case when $\rho_s$ is monotonous, but has a vanishing derivative at single point, we cannot conclude.

**2.3. *Bootstrap argument.*** The purpose of this subsection is to prove, thanks to a bootstrap argument, that under the assumptions of Theorem 1.1, the solution $\theta'$ of (1-7) enjoys the same decay rates as the ones predicted by the linear analysis (see Proposition 2.6). More precisely, we shall prove the following result:

**Proposition 2.9.** *Let $\theta_0 \in H^6(\Omega)$ such that $\theta_0|_{\partial\Omega} = \partial_{\boldsymbol{n}} \theta_0|_{\partial\Omega} = 0$. There exists $\varepsilon_0 > 0$ such that if $\|\theta_0\|_{H^6} \leq \varepsilon \leq \varepsilon_0$ the solution of* (1-7) *satisfies*

$$\|\partial_x^3 \theta'(t)\|_{L^2} \lesssim \frac{\varepsilon}{1+t}, \quad \|\partial_x \theta'(t)\|_{H^4} \lesssim \varepsilon, \quad \|\bar{\theta}(t)\|_{H^5} \lesssim \varepsilon, \quad \forall t > 0. \tag{2-8}$$

**Remark 2.10.** The interplay between horizontal derivatives of $\theta$ and the considered regularities is consistent with the operator $\Delta^{-2}\partial_x^2$ from the linearized system

$$\partial_t \theta' = \partial_x \psi = \Delta^{-2}\partial_x^2 \theta'.$$

Note that $\Delta^{-2}$ denotes the operator solving the bilaplacian $\Delta^2\psi = f$ equation endowed with the boundary condition $\psi|_{\partial\Omega} = \partial_{\boldsymbol{n}}\psi|_{\partial\Omega} = 0$.

A proof of Proposition 2.9 is provided in the rest of this section. Remarks motivating the necessity of the bootstrap hypothesis and the method in general are included throughout. We also present our understanding of the obstacle to the iteration of this method to higher regularity on the perturbation.

***Bootstrap assumption and general argument.*** Let $0 < B < 1$. For some $C_0 > 1$, to be chosen later, let $\theta_0 \in H_0^2 \cap H^6$ such that $\|\theta_0\|_{H^6} \leq B/C_0$. In particular $\|\partial_x^3 \theta_0'\|_{L^2} \leq B/C_0$, $\|\partial_x \theta_0'\|_{H^4} \leq B/C_0$, and $\|\partial_z \bar{\theta}_0\|_{H^2} \leq B/C_0$. Let us note $\psi_0 := \psi|_{t=0}$. We also have, according to Lemma B.1, with universal positive constants gathered under the same notation $C$,

$$\|\psi_0\|_{H^4} \leq C\|\partial_x \theta_0'\|_{L^2} \leq CB/C_0,$$

and therefore,

$$\begin{aligned}\|\partial_t \partial_x \theta'|_{t=0}\|_{L^2} &\leq \|1 - \partial_z \bar{\theta}_0\|_{L^\infty} \|\partial_x^2 \psi_0\|_{L^2} + \|\partial_x (\nabla^\perp \psi_0 \cdot \nabla \theta_0')\|_{L^2} \\ &\leq (1 + \|\partial_z \bar{\theta}_0\|_{H^2}) \|\partial_x^2 \psi_0\|_{L^2} + \|\partial_x (\nabla^\perp \psi_0 \cdot \nabla \theta_0')\|_{L^2} \\ &\leq C(B/C_0 + (B/C_0)^2).\end{aligned}$$

Up to a choice of $C_0 > 1$ large enough, we find that all the bounds here above are strictly smaller than $B$. Therefore, by continuity of the Sobolev norms of $\theta$, ensured by Theorem A.1, there exists a maximal time $T^* \in \mathbb{R}_+ \cup \{+\infty\}$ such that the following inequalities are satisfied on $[0, T^*)$:

$$\begin{aligned}\|\partial_x^3 \theta'(t)\|_{L^2} &\leq \frac{B}{1+t}, \qquad & \|\partial_x \theta'(t)\|_{H^4} &\leq B, \\ \|\partial_z \bar{\theta}(t)\|_{H^2} &\leq B, & \|\partial_t \partial_x \theta'(t)\|_{L^2} &\leq \frac{B}{(1+t)^2}.\end{aligned} \tag{2-9}$$

We recall that these decay rates follow the behavior of the linearized system; see Proposition 2.6.

Let us assume by contradiction that $T^* < +\infty$. We show below by a bootstrap argument that hypothesis (2-9), combined with Lemma 2.4 and Proposition 2.6, actually leads to an improvement of the inequalities, satisfied with some new constant $0 < \underline{B} < B$, which contradicts the maximality of $T^*$. Whence $T^* = +\infty$ and inequalities (2-9) hold for all times.

**Remark 2.11.** Let us wait a little on the choice of the constant $B$. We will choose $B \leq \gamma_0$, so that assumption (2-3) is satisfied on $(0, T^*)$.

***Preliminary bounds.*** Throughout the proof we require estimates on $\theta'$ and $\psi$ derived from the bootstrap hypothesis (2-9). For the sake of readability, we introduce the following short-hand notation:

$$\underbrace{\|f\|^\alpha \|g\|^\beta}_{\alpha r + \beta r'} \quad \text{when } \|f\| \lesssim \frac{B}{(1+t)^r} \text{ and } \|g\| \lesssim \frac{B}{(1+t)^{r'}}.$$

First, from an integration by parts, since $\theta'|_{\partial\Omega} = \partial_{\boldsymbol{n}} \theta'|_{\partial\Omega} = 0$ (see Lemma 2.1)

$$\|\partial_x^2 \theta'\|_{H^2}^2 \lesssim \int_\Omega \partial_x^2 \Delta\theta' \partial_x^2 \Delta\theta' = -\int_\Omega \partial_x^3 \theta' \partial_x \Delta^2 \theta' \leq \|\partial_x^3 \theta'\|_{L^2} \|\partial_x \theta'\|_{H^4}.$$

We deduce, by assumption (2-9), for all $t \in (0, T^*)$

$$\|\partial_x^2 \theta'(t)\|_{H^2} \lesssim \underbrace{\|\partial_x^3 \theta'(t)\|_{L^2}^{\frac{1}{2}} \|\partial_x \theta'(t)\|_{H^4}^{\frac{1}{2}}}_{\frac{1}{2}\times 1 + \frac{1}{2}\times 0} \lesssim \frac{B}{(1+t)^{\frac{1}{2}}}. \tag{2-10}$$

We also get by interpolation, for any $0 \le m \le 4$, for all $t \in (0, T^*)$

$$\|\partial_x\theta'(t)\|_{H^m} \lesssim \underbrace{\|\partial_x\theta'(t)\|_{L^2}^{1-\frac{m}{4}}\|\partial_x\theta'(t)\|_{H^4}^{\frac{m}{4}}}_{(1-\frac{m}{4})\times 1+\frac{m}{4}\times 0} \lesssim \frac{B}{(1+t)^{1-\frac{m}{4}}}. \tag{2-11}$$

We will frequently use *Agmon's inequality* in dimension 2, namely

$$\forall f \in H_0^1 \cap H^2(\Omega), \quad \|f\|_{L^\infty} \lesssim \|f\|_{L^2}^{\frac{1}{2}}\|f\|_{H^2}^{\frac{1}{2}},$$

together with the following direct consequence:

$$\forall f \in H_0^2 \cap H^4(\Omega), \quad \|\nabla f\|_{L^\infty} \lesssim \|f\|_{L^2}^{\frac{1}{2}}\|f\|_{H^4}^{\frac{1}{2}}.$$

We infer, in particular, for all $t \in (0, T^*)$

$$\begin{aligned} \|\partial_x^2\theta'(t)\|_{L^\infty} &\lesssim \underbrace{\|\partial_x^2\theta'(t)\|_{L^2}^{\frac{1}{2}}\|\partial_x^2\theta'(t)\|_{H^2}^{\frac{1}{2}}}_{\frac{1}{2}\times 1+\frac{1}{2}\times\frac{1}{2}} \lesssim \frac{B}{(1+t)^{\frac{3}{4}}}, \\ \|\nabla\partial_x\theta'(t)\|_{L^\infty} &\lesssim \underbrace{\|\partial_x\theta'(t)\|_{L^2}^{\frac{1}{2}}\|\partial_x\theta'(t)\|_{H^4}^{\frac{1}{2}}}_{\frac{1}{2}\times 1+\frac{1}{2}\times 0} \lesssim \frac{B}{(1+t)^{\frac{1}{2}}}. \end{aligned} \tag{2-12}$$

We also need estimates on $\psi$. Any Sobolev norm of order larger than 4 inherits the decays from $\theta'$ thanks to Lemma B.1, providing, for $t \in (0, T^*)$,

$$\begin{aligned} \|\partial_x^2\psi(t)\|_{H^4} &\lesssim \|\partial_x^3\theta'(t)\|_{L^2} \lesssim \frac{B}{1+t}, \\ \|\partial_x\psi(t)\|_{H^6} &\lesssim \|\partial_x^2\theta'(t)\|_{H^2} \lesssim \frac{B}{(1+t)^{\frac{1}{2}}}. \end{aligned} \tag{2-13}$$

We also need higher-order decays on $\partial_x\psi$ in $L^2(\Omega)$. We access this quantity thanks to the control of $\partial_t\theta'$ by rewriting

$$\partial_x\psi = \frac{\partial_t\theta' + (\boldsymbol{u}\cdot\nabla\theta')'}{1-G}.$$

We know that $\|G\|_{L^\infty} \lesssim \|G\|_{H^2} \le B$ so it is enough to have $B$ smaller than a universal constant to ensure that the inverse of $(1-G)$ is well-defined, which allows us to estimate $\partial_x\psi$ and $\partial_x^2\psi$ in $L^2$. We illustrate the computation for $\partial_x\psi$ since the same reasoning applies for $\partial_x^2\psi$ with a few extra terms:

$$\begin{aligned} \|\partial_x\psi(t)\|_{L^2} &\lesssim \|\partial_t\theta'(t)\|_{L^2} + \|\boldsymbol{u}\cdot\nabla\theta'(t)\|_{L^2} \\ &\lesssim \|\partial_t\theta'(t)\|_{L^2} + \|\partial_z\psi(t)\|_{L^\infty}\|\partial_x\theta'(t)\|_{L^2} + \|\partial_x\psi(t)\|_{L^2}\|\partial_z\theta'(t)\|_{L^\infty} \\ &\lesssim \underbrace{\|\partial_t\theta'(t)\|_{L^2}}_{2} + \underbrace{\|\partial_x\theta'(t)\|_{L^2}^2}_{2\times 2} + \|\partial_x\psi(t)\|_{L^2}\frac{B}{(1+t)^{\frac{1}{2}}}. \end{aligned}$$

Hence for $B > 0$ small enough once again we get

$$\|\partial_x^2\psi(t)\|_{L^2} \lesssim \frac{B}{(1+t)^2}.$$

By interpolation and Agmon inequalities we deduce, in the same fashion as above, the following decay estimates, with the latest valid for $0 \le m \le 4$:

$$\begin{aligned} &\|\partial_x^2\psi(t)\|_{H^2} \lesssim \frac{B}{(1+t)^{\frac{3}{2}}}, \quad \|\partial_x^2\psi(t)\|_{L^\infty} \lesssim \frac{B}{(1+t)^{\frac{7}{4}}}, \quad \|\nabla\partial_x^2\psi(t)\|_{L^\infty} \lesssim \frac{B}{(1+t)^{\frac{3}{2}}}, \\ &\|\nabla^2\psi(t)\|_{L^\infty} \lesssim \frac{B}{(1+t)^{\frac{5}{4}}}, \quad \|\partial_x^2\psi(t)\|_{H^m} \lesssim \|\partial_x^2\psi(t)\|_{L^2}^{1-\frac{m}{4}}\|\partial_x^2\psi(t)\|_{H^4}^{\frac{m}{4}} \lesssim \frac{B}{(1+t)^{2-\frac{m}{4}}}. \end{aligned} \tag{2-14}$$

**Remark 2.12.** Let us come back briefly to the derivation of estimates (2-14), which ensure that $\psi$ decays *faster* than $\theta'$. Note that such a fast decay is necessary to close the estimates: indeed, $\|\boldsymbol{u}(t)\|_{W^{1,\infty}}$ should be time integrable in order that $\theta(t)$ converges as $t \to \infty$. Formally, one needs to take an antiderivative in space of the equation, i.e., apply the operator $\Delta^{-2}\partial_x^2$ to (1-7). However, because of the nonlinear term, this is a rather tedious operation. Therefore we rather derive estimates on $\partial_t\theta'$, and use (1-7) in order to infer estimates on $\psi$. Note that the two operations (taking a time derivative or an antispace derivative) are equivalent at main order, since the linear operator is $\partial_t - \Delta^{-2}\partial_x^2$. This idea, although simple, seems to us to be new.

***$H^6$ bound on the solution and $H^4$ bound on $G$.*** In our nonlinear bootstrap argument, we shall need some high Sobolev bound on the solution. In order to lighten the proof of the bootstrap as much as possible, we isolate in the present subsection this technical step.

**Lemma 2.13.** *Let $\theta_0 \in H^6(\Omega)$ such that $\theta_0|_{\partial\Omega} = \partial_{\boldsymbol{n}}\theta_0|_{\partial\Omega} = 0$. Let $T^*$ be the maximal time on which the assumptions* (2-9) *are satisfied. Then, for all $t \in (0, T^*)$,*

$$\|\theta'(t)\|_{H^6} \lesssim B(1+t)^{\frac{1}{2}}, \quad \|G(t)\|_{H^4} \le B/C_0 + CB^2, \quad \|\partial_t G(t)\|_{L^\infty} \lesssim \frac{B^2}{(1+t)^2}.$$

*Proof.* We cannot estimate $\theta'$ in $H^6$ directly from its evolution equation since it requires an assumption on $G = \partial_z\bar{\theta} \in H^6$, and therefore on $\theta$ in $H^7$. To get around this, we directly perform an estimate from the whole perturbed evolution equation, namely

$$\partial_t\theta + \boldsymbol{u}\cdot\nabla\theta = u_z.$$

For any derivative of order 6 (and less) of the previous equation, we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\partial^6\theta\|_{L^2}^2 + \int_\Omega [\partial^6, \boldsymbol{u}\cdot\nabla]\theta\,\partial^6\theta = \int_\Omega \partial^6\partial_x\psi\,\partial^6\theta,$$

where the commutator comes from the incompressibility assumption and the no-slip boundary condition. Hence we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\theta\|_{H^6} \lesssim \|\partial_x\psi\|_{H^6} + \|[\partial^6, \boldsymbol{u}\cdot\nabla]\theta\|_{L^2}.$$

The first term is dealt with thanks to the bilaplacian regularization (Lemma B.1) and estimate (2-10),

$$\|\partial_x\psi\|_{H^6} \lesssim \|\partial_x^2\theta'\|_{H^2} \lesssim \frac{B}{(1+t)^{\frac{1}{2}}}.$$

Notice that this $\frac{1}{2}$-algebraic decay, issued from the linear system, is critical to prove the $\frac{1}{2}$-algebraic growth control of $\theta$ in $H^6(\Omega)$. Concerning the nonlinear term, we rely on the following *tame estimate*, valid for any $m \in \mathbb{N}$,

$$\forall f, g \in H^m \cap L^\infty(\Omega), \quad \|fg\|_{H^m} \lesssim \|f\|_{L^\infty}\|g\|_{H^m} + \|f\|_{H^m}\|g\|_{L^\infty}, \tag{2-15}$$

which leads to

$$\forall f \in H^m \cap W^{1,\infty}(\Omega),\, g \in H^{m-1} \cap L^\infty(\Omega), \quad \|[\partial^m, f]g\|_{L^2} \lesssim \|\nabla f\|_{L^\infty}\|g\|_{H^{m-1}} + \|f\|_{H^m}\|g\|_{L^\infty}. \tag{2-16}$$

Hence, we decompose the nonlinear commutator into

$$[\partial^6, \boldsymbol{u} \cdot \nabla]\theta = -[\partial^6, \partial_z\psi\partial_x]\theta + [\partial^6, \partial_x\psi\partial_z]\theta.$$

Each part can be estimated, thanks to (2-16), as follows:

$$\begin{aligned}\|[\partial^6, \partial_z\psi]\partial_x\theta\|_{L^2} &\lesssim \|\nabla\partial_z\psi\|_{L^\infty}\|\partial_x\theta\|_{H^5} + \|\partial_z\psi\|_{H^6}\|\partial_x\theta\|_{L^\infty} \\ &\lesssim \underbrace{\|\nabla^2\psi\|_{L^\infty}}_{\frac{5}{4}}\|\theta\|_{H^6} + \underbrace{\|\partial_x\theta'\|_{H^3}\|\partial_x\theta'\|_{L^\infty}}_{\frac{1}{4}+\frac{3}{4}} \lesssim \frac{B}{(1+t)^{\frac{5}{4}}}\|\theta\|_{H^6} + \frac{B^2}{1+t},\end{aligned}$$

and

$$\begin{aligned}\|[\partial^6, \partial_x\psi]\partial_z\theta\|_{L^2} &\lesssim \|\nabla\partial_x\psi\|_{L^\infty}\|\partial_z\theta\|_{H^5} + \|\partial_x\psi\|_{H^6}\|\partial_z\theta\|_{L^\infty} \\ &\lesssim \underbrace{\|\nabla\partial_x\psi\|_{L^\infty}}_{\frac{3}{2}}\|\theta\|_{H^6} + \underbrace{\|\partial_x^2\theta'\|_{H^2}}_{\frac{1}{2}}\|\nabla\theta\|_{L^\infty} \lesssim \frac{B}{(1+t)^{\frac{3}{2}}}\|\theta\|_{H^6} + \frac{B^2}{1+t} + \frac{B^2}{(1+t)^{\frac{1}{2}}},\end{aligned}$$

where we observed in particular that

$$\|\nabla\theta\|_{L^\infty} \le \|\nabla\theta'\|_{L^\infty} + \|G\|_{L^\infty} \lesssim \frac{B}{(1+t)^{\frac{1}{2}}} + B.$$

In the end, gathering and summing up all these bounds provides

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\theta\|_{H^6} \lesssim \frac{B}{(1+t)^{\frac{5}{4}}}\|\theta\|_{H^6} + \frac{B^2}{(1+t)^{\frac{1}{2}}},$$

and we get

$$\|\theta\|_{H^6} \lesssim \|\theta_0\|_{H^6} + B(1+t)^{\frac{1}{2}} \lesssim B^2(1+t)^{\frac{1}{2}}.$$

Eventually, let us prove decaying bounds on $\partial_t G$ and uniform bounds on $G$. We recall that $G = \partial_z\bar{\theta}$ depends only on the variables $t$ and $z$. From the evolution equation and one integration by parts we observe (omitting the factors $\frac{1}{2\pi}$ for clarity)

$$\partial_t\bar{\theta} = -\overline{\boldsymbol{u} \cdot \nabla\theta'} = \int_{\mathbb{T}}(\partial_z\psi\partial_x\theta' - \partial_x\psi\partial_z\theta') = -\partial_z\int_{\mathbb{T}}\partial_x\psi\theta',$$

so we can write

$$\partial_t G = -\partial_z^2\int_{\mathbb{T}}\partial_x\psi\theta'.$$

The same arguments as above lead to

$$\|\partial_t G\|_{L^2(0,1)} \lesssim \underbrace{\|\partial_x\psi\|_{L^\infty}\|\theta'\|_{H^2}}_{\frac{7}{4}+\frac{1}{2}} + \underbrace{\|\partial_x\psi\|_{H^2}\|\theta'\|_{L^\infty}}_{\frac{3}{2}+\frac{3}{4}} \lesssim \frac{B^2}{(1+t)^{2+\frac{1}{4}}}.$$

Using the $H^6$ estimate, we also have

$$\|\partial_t G\|_{H^4(0,1)} \lesssim \underbrace{\|\partial_x \psi\|_{L^\infty} \|\theta'\|_{H^6}}_{\frac{7}{4}-\frac{1}{2}} + \underbrace{\|\partial_x \psi\|_{H^6} \|\theta'\|_{L^\infty}}_{\frac{1}{2}+\frac{3}{4}} \lesssim \frac{B^2}{(1+t)^{1+\frac{1}{4}}}.$$

Since the right-hand side of the above inequality is time-integrable, we infer that

$$\|G\|_{H^4(0,1)} \le \|G(0)\|_{H^4(0,1)} + \int_0^t \|\partial_t G\|_{H^4(0,1)} \le B/C_0 + CB^2.$$

Moreover, for all $t \in (0, T^*)$,

$$\|\partial_t G\|_{L^\infty} \lesssim \underbrace{\|\partial_t G\|_{L^2}^{\frac{3}{4}} \|\partial_t G\|_{H^4}^{\frac{1}{4}}}_{\frac{3}{4}\times\frac{9}{4}+\frac{1}{4}\times\frac{5}{4}} \lesssim \frac{B^2}{(1+t)^2}. \qquad \square$$

**Remark 2.14.** In the estimate of $\partial^6\theta$, it would be tempting to proceed to the same computations as in (2-6) in order to exhibit a dissipative term, which would allow us to ignore its contribution as for lower-order derivatives. Doing so requires to control the boundary integrals, which do not vanish a priori in this case,

$$\int_\Omega \partial^6 \partial_x \psi \, \partial^6 \theta = -\int_\Omega \partial^6 \psi \, \Delta^2 \partial^6 \psi = -\|\Delta \partial^6 \psi\|_{L^2}^2 + \int_{\partial\Omega} (\partial_{\boldsymbol{n}} \partial^6 \psi \, \Delta \partial^6 \psi - \partial^6 \psi \, \partial_{\boldsymbol{n}} \partial^6 \Delta \psi).$$

For instance, trying to bound the integral involving the higher order of $z$-derivatives on $\psi$ provides at best

$$\left| \int_{\partial\Omega} \partial^6 \psi \, \partial_z \Delta \partial^6 \psi \right| \lesssim \|\partial_x \psi\|_{H^{6+1/2+\delta}} \|\partial_x^{-1} \psi\|_{H^{9+1/2+\delta}} \lesssim B^2 (1+t)^{\frac{1}{4}+\frac{\delta}{2}}.$$

This estimate ensures no better growth control than $\|\theta'\|_{H^6} \lesssim (1+t)^{3/4}$, which is not enough to close the bootstrap and get the control by $(1+t)^{1/2}$.

***Improvements of the bootstrap bounds.*** We now improve the uniform bound on $\theta'$ and $\partial_x\theta'$ in $H^4(\Omega)$, relying on the linear analysis from Section 2.2. Since $\|\theta'\|_{H^4} \le \|\partial_x \theta'\|_{H^4}$, it is enough to treat $\partial_x\theta'$. Also we have according to Lemma B.1 the inequality

$$\|\partial_x \theta'\|_{H^4} \lesssim \|\Delta^2 \partial_x \theta'\|_{L^2}$$

since $\partial_x\theta'$ belongs in particular to $H_0^2 \cap H^4(\Omega)$ as detailed in Section 2.1, so it is enough to deal with $\partial_x \Delta^2 \theta'$ in $L^2(\Omega)$.

**Lemma 2.15** (uniform bound for $\|\partial_x\theta'\|_{H^4}$). *As long as the bootstrap hypothesis* (2-9) *holds we have*

$$\|\partial_x \Delta^2 \theta'(t)\|_{L^2} \le B/C_0 + CB^2. \tag{2-17}$$

*Proof.* In view of the application of Lemma 2.4 to $\Delta^2\partial_x\theta'$, we observe that its evolution is governed by the equation

$$\partial_t \Delta^2 \partial_x \theta' = (1-G) \partial_x \Delta^2 \partial_x \psi - [\Delta^2 \partial_x, G] \partial_x \psi - \Delta^2 \partial_x (\boldsymbol{u} \cdot \nabla \theta')',$$

which is of the form (2-1) with $\Delta^2\partial_x\psi = \partial_x^2\theta'$ and $\partial_z\Delta^2\partial_x\psi = \partial_z\partial_x^2\theta'$ vanishing on the boundary $\partial\Omega$ and with

$$S = \underbrace{-[\Delta^2\partial_x, G]\partial_x\psi - [\Delta^2\partial_x, \boldsymbol{u}\cdot\nabla]\theta'}_{S_\parallel}\underbrace{-\boldsymbol{u}\cdot\nabla\Delta^2\partial_x\theta'}_{S_\perp}.$$

We already know that $\|G\|_{H^2}$ satisfies the smallness assumption (2-9) for $B > 0$ small enough. We show that $S_\parallel$ presents an algebraic decay strictly larger than 1, as in (2-2). To do so we apply the tame estimate (2-16) to the two commutator terms. Let us emphasize that we need to be thorough by substituting $\boldsymbol{u} = \nabla^\perp\psi$ such that the transport operator can be written as

$$\boldsymbol{u}\cdot\nabla = -\partial_z\psi\partial_x + \partial_x\psi\partial_z.$$

Hence the nonlinear term presents formally only a vertical derivative of order 1. This makes a difference in the estimates and allows us to reach more optimal decay rates.

On the one hand, we get for the perturbation due to $G$, using (2-14),

$$\|[\Delta^2, G]\partial_x^2\psi\|_{L^2} \le \|G\|_{H^4}\underbrace{\|\partial_x^2\psi\|_{L^\infty}}_{\frac{7}{4}} + \|\nabla G\|_{L^\infty}\underbrace{\|\partial_x^2\psi\|_{H^3}}_{\frac{5}{4}} \lesssim \frac{B^2}{(1+t)^{\frac{5}{4}}}.$$

Note that we used here the uniform $H^4$ bound on $G$ from Lemma 2.13. On the other hand the contribution of $[\Delta^2\partial_x, \boldsymbol{u}\cdot\nabla]\theta'$ splits into four terms as follows:

$$[\Delta^2\partial_x, \boldsymbol{u}\cdot\nabla]\theta' = -\Delta^2(\partial_x\partial_z\psi\partial_x\theta') + \Delta^2(\partial_x^2\psi\partial_z\theta') - [\Delta^2, \partial_z\psi]\partial_x^2\theta' + [\Delta^2, \partial_x\psi]\partial_z\partial_x\theta'.$$

We estimate each term accordingly. We have, for instance, using the bootstrap assumption (2-9), the preliminary bounds (2-12) and (2-14), and Lemma 2.13,

$$\begin{aligned}\|\Delta^2(\partial_x^2\psi\partial_z\theta')\|_{L^2} &\lesssim \|\partial_x^2\psi\|_{H^4}\|\partial_z\theta'\|_{L^\infty} + \|\partial_x^2\psi\|_{L^\infty}\|\partial_z\theta'\|_{H^4}\\ &\lesssim \underbrace{\|\partial_x^3\theta'\|_{L^2}\|\nabla\theta'\|_{L^\infty}}_{1+\frac{1}{2}} + \underbrace{\|\partial_x^2\psi\|_{L^\infty}\|\theta'\|_{H^5}}_{\frac{7}{4}-\frac{1}{4}} \lesssim \frac{B^2}{(1+t)^{\frac{3}{2}}}.\end{aligned}$$

The limiting decay comes from one of the commutators, which we estimate, thanks to (2-16) together with the bounds (2-14), (2-11) and (2-12):

$$\|[\Delta^2, \partial_z\psi]\partial_x^2\theta'\|_{L^2} \lesssim \underbrace{\|\nabla\partial_z\psi\|_\infty\|\partial_x^2\theta'\|_{H^3}}_{\frac{5}{4}+0} + \underbrace{\|\psi\|_{H^5}\|\partial_x^2\theta\|_{L^\infty}}_{\frac{3}{4}+\frac{3}{4}} \lesssim \frac{B^2}{(1+t)^{\frac{5}{4}}}.$$

Gathering these estimates provides

$$\|S_\parallel\|_{L^2} \lesssim \frac{B^2}{(1+t)^{\frac{5}{4}}},$$

and Lemma 2.4 applies, ensuring

$$\|\Delta^2\partial_x\theta'\|_{L^2} \le \|\Delta^2\partial_x\theta_0'\|_{L^2} + CB^2. \qquad \square$$

**Lemma 2.16** (decay of $\|\partial_x^3\theta'\|_{L^2}$). *As long as the bootstrap hypothesis* (2-9) *holds*, *we have*

$$\|\partial_x^3\theta'\|_{L^2} \lesssim \frac{B/C_0 + B^2}{1+t}.$$

*Proof.* Note that $\partial_x^3\theta'$ satisfies (2-1) with the source term

$$S = S_\parallel = -\partial_x^3(\boldsymbol{u}\cdot\nabla\theta').$$

We can bound the whole term $S = S_\parallel$ as follows:

$$\begin{aligned}\|S\|_{L^2} &\le \|\partial_x^3(\partial_z\psi\,\partial_x\theta')\|_{L^2}+\|\partial_x^3(\partial_x\psi\,\partial_z\theta')\|_{L^2}\\ &\lesssim \|\partial_x^3\partial_z\psi\|_{L^2}\|\partial_x\theta'\|_{L^\infty}+\|\partial_z\psi\|_{L^\infty}\|\partial_x^4\theta'\|_{L^2}+\|\partial_x^4\psi\|_{L^2}\|\partial_z\theta'\|_{L^\infty}+\|\partial_x\psi\|_{L^\infty}\|\partial_x^3\partial_z\theta'\|_{L^2}\\ &\lesssim \underbrace{\|\partial_x^2\psi\|_{H^2}\|\partial_x\theta'\|_{L^\infty}}_{\frac32+\frac34>2}+\underbrace{\|\nabla\psi\|_{L^\infty}\|\partial_x^2\theta'\|_{H^2}}_{\frac32+\frac12=2}+\underbrace{\|\partial_x^2\psi\|_{H^2}\|\nabla\theta'\|_{L^\infty}}_{\frac32+\frac12=2}+\underbrace{\|\partial_x\psi\|_{L^\infty}\|\partial_x^2\theta'\|_{H^2}}_{\frac74+\frac12>2}\lesssim\frac{B^2}{(1+t)^2}.\end{aligned}$$

Assumption (2-4) is satisfied with $\delta = 1$. Additionally, the norm of $(\Delta^2\partial_x^{-2})\partial_x^3\theta' = \Delta^2\partial_x\theta'$ is bounded according to (2-17), so assumption (2-5) is satisfied with $A = \|\Delta^2\partial_x\theta_0'\|_{L^2} + CB^2$ and $\alpha = 0$. Moreover, the traces of $\partial_x^3\theta'$ and $\partial_{\boldsymbol{n}}\partial_x^3\theta'$ vanish as a direct consequence of Lemma 2.1. Therefore $\min(1+\alpha,\delta) = 1$ and Proposition 2.6 provides

$$\|\partial_x^3\theta'\|_{L^2} \lesssim (\|\partial_x^3\theta_0'\|_{L^2} + \|\Delta^2\partial_x\theta'\|_{L^\infty((0,t),L^2)} + B^2)\frac{1}{1+t}.$$

Using inequality (2-17), we obtain the desired estimate. $\square$

**Lemma 2.17** (stronger decay of $\|\partial_t\partial_x\theta'(t)\|_{L^2}$). *Under assumptions* (2-9) *we have*, *for all* $t \in (0, T^*)$,

$$\|\partial_t\partial_x\theta'\|_{L^2} \lesssim \frac{B/C_0 + B^2}{(1+t)^2}.$$

*Proof.* The pair $(\partial_t\partial_x\theta', \partial_t\partial_x\psi)$ satisfies (2-1) with

$$S = -\partial_t G\partial_x^2\psi - \partial_t\partial_x(\boldsymbol{u}\cdot\nabla\theta').$$

Note that $\partial_t\partial_x\theta'$ and $\partial_{\boldsymbol{n}}\partial_t\partial_x\theta'$ vanish on the boundary, from Lemma 2.1. In order to apply Proposition 2.6, we have to bound $(\Delta^2\partial_x^{-2})\partial_t\partial_x\theta'$ in $L^2(\Omega)$. Going back to (1-7), we have

$$\Delta^2\partial_x^{-1}\partial_t\theta' = \Delta^2((1-G)\psi) - \Delta^2\partial_x^{-1}(\boldsymbol{u}\cdot\nabla\theta')),$$

the norm of which can be estimated as

$$\begin{aligned}\|\Delta^2\partial_x^{-1}\partial_t\theta'\|_{L^2} &\le (1+\|G\|_{H^4})\|\Delta^2\psi\|_{L^2} + \|\Delta^2\partial_x^{-1}(\boldsymbol{u}\cdot\nabla\theta')\|_{L^2}\\ &\lesssim \|\partial_x\theta'\|_{L^2} + \underbrace{\|\psi\|_{H^5}\|\nabla\theta'\|_{L^\infty}}_{\frac34+\frac12} + \underbrace{\|\nabla\psi\|_{L^\infty}\|\theta'\|_{H^5}}_{\frac32-\frac14} \lesssim \frac{B/C_0+B^2}{1+t} + \frac{B^2}{(1+t)^{\frac54}}.\end{aligned}$$

Hence assumption (2-5) is satisfied with $\alpha = 1$ and $A = C(B/C_0 + B^2)$.

Let us set $S_\perp := \boldsymbol{u}\cdot\nabla\partial_t\partial_x\theta'$, indeed orthogonal to $\partial_t\partial_x\theta'$ in $L^2(\Omega)$. We further define

$$\begin{aligned}S_\parallel &:= -\partial_t G\partial_x^2\psi - \partial_x\boldsymbol{u}\cdot\nabla\partial_t\theta',\\ S_\Delta &:= \partial_t\partial_x\boldsymbol{u}\cdot\nabla\theta' - \partial_t\boldsymbol{u}\cdot\nabla\partial_x\theta',\end{aligned}$$

so that $S_\perp + S_\parallel + S_\Delta = S$. Let us now check that $S_\parallel$ and $S_\Delta$ satisfy the assumptions of Proposition 2.6.

The first term in $S_\parallel$ can be bounded directly as follows, using Lemma 2.13:

$$\|\partial_t G \partial_x^2 \psi\|_{L^2} \le \underbrace{\|\partial_t G\|_{L^\infty} \|\partial_x^2 \psi\|_{L^2}}_{2+2} \lesssim \frac{B^2}{(1+t)^4}.$$

The second requires, for instance, a bound on $\partial_t \theta'$ in $H^1(\Omega)$, obtained directly from the evolution equation:

$$\begin{aligned}\|\partial_t \theta'\|_{H^1} &\lesssim (1+\gamma_0)\|\partial_x \psi\|_{H^1} + \|\boldsymbol{u}\cdot\nabla\theta'\|_{H^1}\\ &\lesssim (1+\gamma_0)\|\partial_x \psi\|_{H^1} + \|\boldsymbol{u}\|_{H^1}\|\nabla\theta'\|_{L^\infty} + \|\boldsymbol{u}\|_{L^\infty}\|\nabla\theta'\|_{H^1}\\ &\lesssim (1+\gamma_0)\underbrace{\|\partial_x\psi\|_{H^1}}_{\frac{7}{4}} + \underbrace{\|\psi\|_{H^2}\|\nabla\theta'\|_{L^\infty}}_{\frac{3}{2}+\frac{1}{2}} + \underbrace{\|\nabla\psi\|_{L^\infty}\|\theta'\|_{H^2}}_{\frac{3}{2}+\frac{1}{2}} \lesssim \frac{B}{(1+t)^{\frac{7}{4}}}.\end{aligned}$$

Hence

$$\|\partial_x \boldsymbol{u}\cdot\nabla\partial_t\theta'\|_{L^2} \lesssim \|\partial_x \boldsymbol{u}\|_{L^\infty}\|\partial_t\theta'\|_{H^1} \lesssim \underbrace{\|\nabla\partial_x\psi\|_{L^\infty}\|\partial_t\theta'\|_{H^1}}_{\frac{3}{2}+\frac{7}{4}} \lesssim \frac{B^2}{(1+t)^{\frac{13}{4}}},$$

and $S_\parallel$ satisfies the assumption (2-4) with $\sigma = CB^2$ and $\delta = \frac{9}{4}$. Continuing our computations,

$$\|\partial_t\partial_x \boldsymbol{u}\cdot\nabla\theta'\|_{L^2} \lesssim \|\nabla\partial_t\partial_x\psi\|_{L^2}\underbrace{\|\nabla\theta'\|_{L^\infty}}_{\frac{1}{2}} \lesssim \frac{B\|\Delta\partial_t\partial_x\psi\|_{L^2}}{(1+t)^{\frac{1}{2}}},$$

and the same consideration applies for

$$\|\partial_t \boldsymbol{u}\cdot\nabla\partial_x\theta'\|_{L^2} \lesssim \|\nabla\partial_t\psi\|_{L^2}\underbrace{\|\nabla\partial_x\theta'\|_{L^\infty}}_{\frac{1}{2}} \lesssim \frac{B\|\Delta\partial_t\partial_x\psi\|_{L^2}}{(1+t)^{\frac{1}{2}}}.$$

Hence

$$S_\Delta = \partial_t\partial_x \boldsymbol{u}\cdot\nabla\theta' - \partial_t\boldsymbol{u}\cdot\nabla\partial_x\theta'$$

indeed satisfies assumption (2-4). Finally Proposition 2.6 applies with $\min(1+\alpha, \delta) = 2$ and we obtain

$$\|\partial_t\partial_x\theta'\|_{L^2} \lesssim \frac{B/C_0 + B^2}{(1+t)^2}. \qquad \square$$

***Conclusion.*** Let us close the bootstrap argument. Assuming $\|\theta_0\|_{H^6} \le B/C_0$, we had, by continuity-in-time of the Sobolev norms of $\theta$ ensured by Theorem A.1, existence of a maximal time $T^* \in \mathbb{R}_+ \cup \{+\infty\}$ such that (2-9) is satisfied for any $t \in [0, T^*)$, reported here:

$$\begin{aligned}\|\partial_x^3\theta'\|_{L^2} &\le \frac{B}{1+t}, & \|\partial_x\theta'\|_{H^4} &\le B,\\ \|G\|_{H^2} &\le B, & \|\partial_t\partial_x\theta'\|_{L^2} &\le \frac{B}{(1+t)^2}.\end{aligned} \tag{2-18}$$

These decay estimates induce, as shown in Lemmas 2.13, 2.15, 2.16, and 2.17 that (2-18) holds for another constant $\underline{B}$ defined as

$$\underline{B} = \frac{CB}{C_0} + CB^2,$$

where $C > 0$ is universal. By choosing $C_0$ large enough and $B$ small enough, we have $\underline{B} < B$ and inequalities (2-18) are strictly satisfied for any $t \in [0, T^*)$. Therefore $T^*$ must be $+\infty$; otherwise the continuity of $t \mapsto \|\theta(t)\|_{H^6}$ would imply the existence of a larger validity time interval for (2-18). In the end, these bounds are valid for all times, and setting $\varepsilon_0 := B/C_0$ closes the demonstration of Proposition 2.9.

**Remark 2.18** (generalization at any order). Motivated by the fact that the perturbed subproblem

$$\begin{cases} \partial_t \partial_x^\ell \theta' = (1-G)\partial_x^{\ell+1}\psi, \\ \Delta^2 \partial_x^{\ell-1}\psi = \partial_x^\ell \theta', \\ \partial_x^{\ell-1}\psi|_{\partial\Omega} = \partial_{\boldsymbol{n}}\partial_x^{\ell-1}\psi|_{\partial\Omega} = 0 \end{cases}$$

is stable under horizontal derivation we could expect to propagate arbitrarily high horizontal regularity on $\theta'$. Nevertheless, our proof relies on the control

$$\|\theta\|_{H^6} \lesssim (1+t)^{\frac{1}{2}},$$

which we can obtain thanks to the classical divergence-free condition on $\boldsymbol{u}$ canceling the extra-derivative term. Let us try to do the same on $\partial_x^\ell \theta$; we write

$$\partial_t \partial_x^\ell \theta + \sum_{k=0}^{\ell-1} C_{\ell,k} \partial_x^{\ell-k}\boldsymbol{u} \cdot \nabla \partial_x^k \theta + \boldsymbol{u} \cdot \nabla \partial_x^\ell \theta = \partial_x^{\ell+1}\psi,$$

and multiply by $\partial_x^\ell \theta'$. Then the estimation does not close, even though one of its terms does not contribute, just as in the initial equation. Indeed, we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_x^\ell \theta\|_{H^6}^2 + \sum_{k=0}^{\ell-1} C_{\ell,k} \int_\Omega \partial^6(\partial_x^{\ell-k}\boldsymbol{u} \cdot \nabla \partial_x^k \theta)\partial^6 \partial_x^\ell \theta + \underbrace{\frac{1}{2}\int_\Omega \boldsymbol{u} \cdot \nabla |\partial^6 \partial_x^\ell \theta|^2}_{=0} \le \|\partial_x^\ell \psi\|_{H^6}\|\partial_x^\ell \theta\|_{H^6}.$$

Note that crossed derivatives integrands, such as

$$\int_\Omega \partial_x^\ell \boldsymbol{u} \cdot \nabla \partial^6 \theta \partial^6 \partial_x^\ell \theta$$

do not lead to a vanishing integral. Hence deriving an estimate on $\partial_x^\ell \theta'$ in $H^6$ requires first the derivation of an estimate of $\theta$ in $H^{6+\ell}$, in the spirit of Lemma 2.13. We will derive such estimates in Section 4 (see Lemma 4.3), at the price of much stronger and more complicated bootstrap assumptions (see (4-3)).

**2.4.** ***Convergence as $t \to \infty$ and identification of the asymptotic profile.*** Regarding the asymptotic behavior of the density for the Stokes-transport system without any assumption on the type of initial data, we can only say that if $\rho$ converges toward some $\rho_\infty$ in $H^{-1}$, this limit profile is stratified. Indeed, the energy balance (1-2) ensures that $\boldsymbol{u} \in L^2(\mathbb{R}_+, H^1)$, and since $\boldsymbol{u}$ is also $\mathrm{Lip}(\mathbb{R}_+, H^1)$ by linearity of the Stokes system, we infer that $\|\boldsymbol{u}(t)\|_{H^1} \to 0$ as $t \to \infty$, but without any information about its decay rate. At least we have

$$\|\nabla p + \rho \boldsymbol{e}_z\|_{H^{-1}} \lesssim \|\boldsymbol{u}\|_{H^1} \xrightarrow[t\to\infty]{} 0.$$

The $H^{-1}$ convergence of $\rho$ leads to the existence of $p_\infty$ such that

$$\nabla p_\infty = -\rho_\infty \boldsymbol{e}_z.$$

Observing that $\partial_x p_\infty = 0$ and that the domain $\Omega = \mathbb{T} \times (0, 1)$ is convex ensures that $p_\infty$ and $\rho_\infty$ are both independent of the horizontal coordinate $x$.

In the context of a small perturbation of the stationary profile $\rho_s(z) = 1 - z$, we obtained explicit decay rates for Sobolev norms of $\boldsymbol{u}$. We show that these decays are sufficient to ensure the strong convergence of $\rho$ toward a limit profile $\rho_\infty$. Moreover, the smallness of the perturbation $\theta$ does not affect the vertical monotonicity of the whole density $\rho$, from which we deduce that $\rho_\infty$ is exactly the vertical rearrangement of $\rho_0$.

**Proposition 2.19.** *Under the assumptions of Theorem 1.1, the whole density $\rho$ converges in $H^m$ for any $m < 4$ towards its vertical decreasing rearrangement.*

*Proof.* The proof is divided in the following steps:

***Convergence.*** It is enough to show that $\partial_t \rho$ belongs to $L^1(\mathbb{R}_+, H^m)$ for $m < 4$, which implies the strong convergence of $\rho(t)$ in $H^m$ and existence of a limit $\rho_\infty$. Let us estimate $\partial_t \rho$ in $H^m$ for any $0 \le m \le 4$, using the tame estimates (2-15)

$$\begin{aligned} \|\partial_t \rho\|_{H^m} &= \|\boldsymbol{u} \cdot \nabla \rho\|_{H^m} \le \|\partial_z \psi \partial_x \rho\|_{H^m} + \|\partial_x \psi \partial_z \rho\|_{H^m} \\ &\lesssim \|\nabla \psi\|_{L^\infty} \|\partial_x \rho\|_{H^m} + \|\psi\|_{H^{m+1}} \|\partial_x \rho\|_{L^\infty} + \|\partial_x \psi\|_{L^\infty} \|\rho\|_{H^{m+1}} + \|\partial_x \psi\|_{H^m} \|\partial_z \rho\|_{L^\infty}. \end{aligned}$$

Recalling that $\partial_z \rho = -1 + G + \partial_z \theta'$ is bounded in $H^5(\Omega)$, that $\partial_x \rho = \partial_x \theta'$ decays as $(1+t)^{-1+m/4}$ for $m \le 4$, as well as the decay estimates (2-13) and (2-14), we find

$$\|\partial_t \rho\|_{H^m} \lesssim \frac{\|\rho_0\|_{H^6}^2}{(1+t)^{2-\frac{m}{4}}},$$

which is integrable for any $m < 4$, hence the convergence.

***Stratified limit.*** Since $\rho$ converges, so do $\theta' = (\rho - \rho_s)'$ and $\bar{\theta} = \overline{\rho - \rho_s}$. We obtained in (2-8) that $\theta'$ vanishes in $H^m$ for $m < 4$, and therefore the limit $\rho_\infty$ is stratified. Hence $\rho_\infty$ can be written as the sum of $\rho_s$ and the limit $\bar{\theta}_\infty$ of $\bar{\theta}$. In view of (2-8) this limit satisfies in particular $\|\partial_z \bar{\theta}_\infty\|_{L^\infty} \le C\varepsilon_0$, with the notation of Theorem 1.1. At least for $\varepsilon_0 > 0$ such that $C\varepsilon_0 < 1 = -\partial_z \rho_s$, we know that $\sup_{(0,1)} \partial_z \rho_\infty < 0$, which means that $\rho_\infty$ is strictly decreasing with respect to $z$.

***Rearrangement.*** The divergence-free character of the velocity field $\boldsymbol{u}$ ensures that all $L^q$ norms and the cumulative distribution function of $\rho(t)$ are preserved along time, in the sense

$$\forall \lambda \ge 0, \quad |\{\rho(t) > \lambda\}| = |\{\rho_0 > \lambda\}|. \tag{2-19}$$

This property transfers to the limit state $\rho_\infty$ by $L^q$ strong convergence of $\rho$. According to rearrangement theory such as that developed in [Lieb and Loss 2001, Chapter 3], we say that two maps are rearrangements of each other if they have the same level sets, in the sense of (2-19). Adapting slightly the construction of [Lieb and Loss 2001], we know there exists a unique vertical decreasing rearrangement of $\rho_0 : \Omega \to \mathbb{R}_+$, which can be defined as

$$\rho_0^*(z) := \int_0^\infty \mathbf{1}_{0 \le z \le |\{\rho_0 > \lambda\}|} \, \mathrm{d}\lambda.$$

In the end, we know that $\rho_\infty$ is a decreasing rearrangement of $\rho_0$; therefore it is $\rho_0^*$ by uniqueness. □

**Remark 2.20.** Note that the above argument extends immediately to the settings investigated by Elgindi [2017] and Castro, Córdoba and Lear [Castro et al. 2019a] for the incompressible porous media problem, as mentioned in the Introduction.

Notice that we actually have $\|\partial_z\theta\|_{L^\infty} \lesssim \varepsilon_0$ for all times. Therefore the total density has a strictly negative vertical derivative, for all $x \in \mathbb{T}$ and for all times $t \in \mathbb{R}_+$, since

$$\partial_z\rho(t, x, \cdot) = -1 + \partial_z\theta(t, x, \cdot),$$

and the density reordering is essentially horizontal. This is a rare case in which we can describe the asymptotic profile. This intuition of having heavy fluids sinking under the lighter ones prompts to wonder if, at least in a weak sense, the density profile should always converge toward the vertical rearrangement of the initial datum, unless it is already stratified. This question remains open, both for the Stokes-transport equation and for the incompressible porous media.

## 3. Formation of large-time boundary layers in the linear setting: proof of Theorem 1.2

The purpose of this section is to prove Theorem 1.2. We consider the linear problem

$$\begin{cases} \partial_t\theta = \partial_x\psi \quad \text{in } (0, +\infty) \times \Omega, \\ \Delta^2\psi = \partial_x\theta \quad \text{in } \Omega, \quad \psi|_{\partial\Omega} = \partial_{\boldsymbol{n}}\psi|_{\partial\Omega} = 0, \\ \theta(t=0) = \theta_0, \end{cases} \tag{3-1}$$

with $\theta_0 \in H^6(\Omega)$ arbitrary. The difference with the linear analysis of Section 2.2, and in particular with Proposition 2.6, lies in the fact that we do not assume that $\theta_0$ and $\partial_{\boldsymbol{n}}\theta_0$ vanish on the boundary. As a consequence, as explained in the sketch of the proof in the Introduction, boundary layers are created as $t \to \infty$ close to $z = 0$ and $z = 1$, and the purpose of this section is precisely to describe the mechanism driving the apparition of these boundary layers. We will therefore decompose $\theta$ as the sum of an interior term decaying like $t^{-1}$ in $L^2$, and some boundary layer terms which lift the traces of $\theta$ and $\partial_{\boldsymbol{n}}\theta$ on the boundary. This will lead us to Theorem 1.2. We will then return to our nonlinear system (1-7) in Section 4.

In fact, we will prove a more precise version of Theorem 1.2:

**Proposition 3.1.** *Let $\theta_0 \in H^s(\Omega)$ for some $s > 0$ sufficiently large. Let $\theta \in C(\mathbb{R}_+, H^s)$ be the unique solution of* (3-1). *There exists a boundary layer profile $\theta^{\mathrm{BL}}$ of the form*

$$\theta^{\mathrm{BL}} = \sum_{j=0}^{4} (1+t)^{-\frac{j}{4}} (\Theta^j_{\mathrm{bot}}(x, Z_{\mathrm{bot}}) + \Theta^j_{\mathrm{top}}(x, Z_{\mathrm{top}})),$$

*with $Z_{\mathrm{bot}} = z(1+t)^{1/4}$ and $Z_{\mathrm{top}} = (1-z)(1+t)^{1/4}$, such that $\theta^{\mathrm{int}} = \theta - \theta^{\mathrm{BL}}$ satisfies, for all $t \geq 0$,*

$$\|\partial_x^2\theta^{\mathrm{int}}(t)\|_{L^2} \lesssim \frac{\|\theta_0\|_{H^s}}{1+t}, \quad \|\theta^{\mathrm{int}}(t)\|_{H^4} \lesssim \|\theta_0\|_{H^s}, \quad \|\Delta^{-2}\partial_x^2\theta^{\mathrm{int}}(t)\|_{L^2} \lesssim \frac{\|\theta_0\|_{H^s}}{(1+t)^2}.$$

*Furthermore, there exists a constant $c > 0$ such that $\|\Theta^j_a(\,\cdot\,, Z)\|_{H^4(\mathbb{T})} \lesssim \|\theta_0\|_{H^s} \exp(-cZ^{4/5})$ for all $Z > 0$.*

The organization of this section is the following. After motivating the ansatz (1-6), we formally derive the equation satisfied by the boundary layer profiles. We then construct the boundary layer part of the solution, denoted by $\theta^{\mathrm{BL}}$, and we establish some properties. Eventually, we prove that $\theta - \theta^{\mathrm{BL}}$ satisfies the assumptions of Proposition 2.6, and we conclude.

**3.1.** ***Motivation for the ansatz and derivation of the boundary layer equations.*** We recall (see page 1966) that a simple spectral decomposition suggests that the solution $\theta$ has strong variations close to the boundaries, and that $\theta(t)|_{\partial\Omega} = \theta_0|_{\partial\Omega}$, $\partial_{\boldsymbol{n}}\theta(t)|_{\partial\Omega} = \partial_{\boldsymbol{n}}\theta_0|_{\partial\Omega}$. Hence we take an ansatz of the form

$$\theta(t) \simeq \theta^{\mathrm{int}} + \Theta^0_{\mathrm{bot}}(x, (1+t)^\alpha z) + \Theta^0_{\mathrm{top}}(x, (1+t)^\alpha (1-z)) \\ + (1+t)^{-\alpha}\Theta^1_{\mathrm{bot}}(x, (1+t)^\alpha z) + (1+t)^{-\alpha}\Theta^1_{\mathrm{top}}(x, (1+t)^\alpha (1-z)) + \text{l.o.t.}$$

for some $\alpha > 0$ to be determined, where

$$\theta^{\mathrm{int}}|_{\partial\Omega} = \partial_z \theta^{\mathrm{int}}|_{\partial\Omega} = 0,$$
$$\Theta^j_{\mathrm{bot}}(x, Z) \to 0 \quad \text{and} \quad \Theta^j_{\mathrm{top}}(x, Z) \to 0 \quad \text{as } Z \to \infty.$$

The role of $\Theta^0_{\mathrm{top}}$ (resp. of $\Theta^0_{\mathrm{bot}}$) is to lift the trace of $\theta_0$ at the top boundary $z = 1$ (resp. at the bottom boundary $z = 0$). Hence we take

$$\Theta^0_{\mathrm{top}}(x, Z{=}0) = \theta_0(x, z{=}1), \quad \partial_Z \Theta^0_{\mathrm{top}}(x, Z{=}0) = 0,$$
$$\Theta^0_{\mathrm{bot}}(x, Z{=}0) = \theta_0(x, z{=}0), \quad \partial_Z \Theta^0_{\mathrm{bot}}(x, Z{=}0) = 0.$$

In a similar way, the next-order boundary layer terms $\Theta^1_{\mathrm{top}}$ and $\Theta^1_{\mathrm{bot}}$ lift the traces of $\partial_{\boldsymbol{n}}\theta_0$ on $\partial\Omega$, i.e.,

$$\Theta^1_{\mathrm{top}}(x, Z{=}0) = 0, \quad \partial_Z \Theta^1_{\mathrm{top}}(x, Z{=}0) = -\partial_z \theta_0(x, z{=}1),$$
$$\Theta^1_{\mathrm{bot}}(x, Z{=}0) = 0, \quad \partial_Z \Theta^1_{\mathrm{bot}}(x, Z{=}0) = \partial_z \theta_0(x, z{=}0).$$

Similarly, we assume that

$$\psi(t) \simeq \psi^{\mathrm{int}} + (1+t)^{-4\alpha}\Psi^0_{\mathrm{bot}}(x, (1+t)^\alpha z) + (1+t)^{-4\alpha}\Psi^0_{\mathrm{top}}(x, (1+t)^\alpha (1-z)) \\ + (1+t)^{-5\alpha}\Psi^1_{\mathrm{bot}}(x, (1+t)^\alpha z) + (1+t)^{-5\alpha}\Psi^1_{\mathrm{bot}}(x, (1+t)^\alpha (1-z)) + \text{l.o.t.},$$

where

$$\partial_Z^4 \Psi^j_a = \partial_x \Theta^j_a,$$
$$\Psi^j_a = \partial_Z \Psi^j_a = 0 \quad \text{on } Z = 0,$$
$$\Psi^j_a \to 0 \quad \text{as } Z \to \infty,\ a \in \{\text{top}, \text{bot}\}.$$

Plugging these ansatz into (3-1), we find that at main order

$$\alpha(1+t)^{-1} Z \partial_Z \Theta^0_a = (1+t)^{-4\alpha}\partial_x \Psi^0_a.$$

Consequently, identifying the powers of $(1+t)$, we take $\alpha = \frac{1}{4}$, which is precisely the ansatz (1-6). Hence the equation for $\Psi^0_a$, $a \in \{\text{top}, \text{bot}\}$, becomes

$$\begin{cases} \frac{1}{4} Z \partial_Z^5 \Psi^0_a = \partial_x^2 \Psi^0_a & \text{in } \mathbb{T} \times (0, +\infty), \\ \Psi^0_{a|Z=0} = \partial_Z \Psi^0_{a|Z=0} = 0, \\ \partial_Z^4 \Psi^0_{a|Z=0} = \gamma^0_a(x), \quad \partial_Z^5 \Psi^0_{a|Z=0} = 0, \\ \lim_{Z\to\infty} \Psi^0_a(x, Z) = 0, \end{cases} \tag{3-2}$$

where $\gamma^0_{\text{bot}}(x) = \partial_x\theta_0(x, z=0)$, $\gamma^0_{\text{top}}(x) = \partial_x\theta_0(x, z=1)$. Note that the above boundary conditions are redundant: indeed, if $\partial_Z \Psi^0_{a|Z=0} = 0$, then it follows from the equation (after one differentiation with respect to $Z$) that $\partial^5_Z \Psi^0_{a|Z=0} = 0$. Hence in the following subsection we will drop the condition $\partial^5_Z \Psi^0_{a|Z=0} = 0$.

In a similar fashion, the equation for $\Psi^1_a$, $a \in \{\text{top}, \text{bot}\}$, is

$$\begin{cases} -\partial^4_Z \Psi^1_a + Z\partial^5_Z \Psi^1_a = 4\partial^2_x \Psi^1_a & \text{in } \mathbb{T} \times (0, +\infty), \\ \Psi^1_{a|Z=0} = \partial_Z \Psi^1_{a|Z=0} = 0, \\ \partial^4_Z \Psi^1_{a|Z=0} = 0, \quad \partial^5_Z \Psi^1_{a|Z=0} = \gamma^1_a(x), \\ \lim_{Z\to\infty} \Psi^1_a(x, Z) = 0, \end{cases} \tag{3-3}$$

where $\gamma^1_{\text{bot}}(x) = \partial_x\partial_z\theta_0(x, z=0)$, $\gamma^1_{\text{top}}(x) = -\partial_x\partial_z\theta_0(x, z=1)$. Once again, the condition $\partial^4_Z \Psi^1_{a|Z=0} = 0$ is redundant and is automatically satisfied when one takes the trace of the equation at $Z = 0$, using the other boundary conditions. We now turn towards the well-posedness of (3-2) and (3-3).

**3.2.** ***Construction of the main profiles.*** The well-posedness of (3-2) and (3-3) stems from the following result:

**Lemma 3.2.** *Let $m \geq m_0 > 0$ and let $S \in C([0, +\infty))$, $\delta > 0$, such that*

$$\|S\|^2 := \int_0^1 \frac{S(Z)^2}{Z^2}\, \mathrm{d}Z + \int_0^\infty S(Z)^2 \exp(\delta Z^{\frac{4}{5}})\, \mathrm{d}Z < +\infty.$$

*Consider the ODE*

$$Z\partial^5_Z \Psi(Z) = -m\Psi(Z) + S(Z) \quad \text{in } (0, +\infty), \qquad \lim_{Z\to\infty} \Psi(Z) = 0, \tag{3-4}$$

*endowed with one of the following four boundary conditions*:

(i) $\Psi(0) = \partial_Z \Psi(0) = \partial^4_Z \Psi(0) = 0$.

(ii) $\Psi(0) = \partial^3_Z \Psi(0) = \partial^4_Z \Psi(0) = 0$.

(iii) $\Psi(0) = \partial^2_Z \Psi(0) = \partial^3_Z \Psi(0) = 0$.

(iv) $\Psi(0) = \partial_Z \Psi(0) = \partial^2_Z \Psi(0) = 0$.

*Then there exists a constant $c > 0$ depending only on $m_0$ and $\delta$ such that* (3-4) *endowed with one of the four previous conditions has a unique solution $\Psi \in H^5_{\text{loc}}(\mathbb{R}_+)$ such that, for all $k \in \{0, \dots, 5\}$,*

$$\int_0^\infty |\partial^k_Z \Psi(Z)|^2 \exp(cZ^{\frac{4}{5}})\, \mathrm{d}Z \leq C\|S\|^2 < +\infty.$$

*As a consequence, for $k \leq 4$, there exists a constant $C$ such that*

$$|\partial^k_Z \Psi(Z)| \leq C\|S\| \exp\Big(-\frac{c}{4} Z^{\frac{4}{5}}\Big) \quad \forall Z > 0.$$

The proof of Lemma 3.2 is postponed to Appendix C, and relies on the use of the Lax–Milgram lemma in weighted Sobolev spaces. As a corollary, we have the following result:

**Corollary 3.3.** *For all $j \in \{0, 1\}$, there exists a unique solution $\chi_j \in C^\infty(0, +\infty)$ of the ODE*

$$Z\partial_Z^5\chi_j - j\partial_Z^4\chi_j + 4\chi_j = 0 \quad \text{on } (0, +\infty),$$

*endowed with the boundary conditions*

- $\chi_0(0) = \partial_Z\chi_0(0) = 0,\ \partial_Z^4\chi_0(0) = 1,$
- $\partial_Z\chi_1(0) = \partial_Z^4\chi_1(0) = 0,\ \partial_Z^5\chi_1(0) = 1,$

*and such that, for $j = 0, 1,\ 0 \le k \le 5$,*

$$\int_0^\infty |\partial_Z^K\chi_j(Z)| \exp(cZ^{\frac{4}{5}})\,\mathrm{d}Z < +\infty.$$

*Furthermore, $\partial_Z^5\chi_0(0) = \chi_1(0) = 0$.*

*Proof.* Let us start with $\chi_0$. Let $\eta \in C_c^\infty(\mathbb{R})$ such that $\eta \equiv 1$ in a neighborhood of zero. Then $\chi_0 - Z^4\eta/4!$ satisfies (3-4) with the boundary conditions (i) and with a $C^\infty$ and compactly supported source term. Hence the result follows from Lemma 3.2. The $C^\infty$ regularity of $\chi_0$ follows easily from the ODE (3-4) and from an induction argument. Differentiating the ODE and taking the trace at $Z = 0$, we obtain $\partial_Z^5\chi_0(0) = -4\partial_Z\chi_0(0) = 0$.

Concerning $\chi_1$, we first consider the solution of the ODE

$$Z\partial_Z^5\phi + 4\phi = 0 \quad \text{on } (0, +\infty),$$
$$\phi(0) = \partial_Z^3\phi(0) = 0, \quad \partial_Z^4\phi(0) = 1, \quad \phi(+\infty) = 0.$$

The existence, uniqueness, and exponential decay of $\phi$ follow from a lifting argument and Lemma 3.2 with boundary conditions (ii). We then set $\chi_1(Z) = -\int_Z^\infty \phi$ and note that $\partial_Z(Z\partial_Z^5\chi_1 - \partial_Z^4\chi_1 + 4\chi_1) = 0$. As a consequence, $Z\partial_Z^5\chi_1(Z) - \partial_Z^4\chi_1(Z) + 4\chi_1(Z) = \text{const.} = 0$ on $(0, +\infty)$, thanks to the decay properties of $\phi$ at infinity. Hence the existence, uniqueness and decay of $\chi_1$ follow. Taking the trace of the equation at $Z = 0$, we find that $\chi_1(0) = 0$. □

Let us now explain how we construct the boundary layer profiles $\Psi_a^j$ for $a \in \{\text{top}, \text{bot}\}$ and $j = 0, 1$ which satisfy (3-2) and (3-3). Taking the Fourier transform of (3-2) with respect to $x$ and dropping the index $a$, we infer that $\widehat{\Psi}_k^0$ satisfies

$$\tfrac{1}{4}Z\partial_Z^5\widehat{\Psi}_k^0 = -k^2\widehat{\Psi}_k^0,$$
$$\partial_Z^4\widehat{\Psi}_{k|Z=0}^0 = \hat{\gamma}_k^0, \quad \widehat{\Psi}_{k|Z=0}^0 = \partial_Z\widehat{\Psi}_{k|Z=0}^0 = 0.$$

Considering the function $\chi_0$ defined in Corollary 3.3, it is then easily checked that

$$\widehat{\Psi}_k^0 = |k|^{-2}\hat{\gamma}_k^0\,\chi_0(|k|^{\frac{1}{2}}Z)$$

is a solution of the problem. We infer that

$$\Psi_a^0(x, Z) := \sum_{k\in\mathbb{Z}\setminus\{0\}} |k|^{-2}\hat{\gamma}_{a,k}^0\,\chi_0(|k|^{\frac{1}{2}}Z)e^{ikx} \tag{3-5}$$

is a solution of (3-2). In a similar way,

$$\Psi_a^1(x,Z) := \sum_{k\in\mathbb{Z}\setminus\{0\}} |k|^{-\frac{5}{2}} \hat{\gamma}_{a,k}^1\, \chi_1(|k|^{\frac{1}{2}}Z)e^{ikx} \tag{3-6}$$

satisfies (3-3).

As a consequence, we have the following estimates, which follow easily from formulas (3-5) and (3-6):

**Corollary 3.4.** *Let $\gamma_a^0, \gamma_a^1 \in L^2(\mathbb{T})$. Then* (3-2) (*resp.* (3-3)) *has a unique solution $\Psi_a^0 \in H_x^{9/4}L_Z^2 \cap L_x^2 H_Z^{9/2}$* (*resp. $\Psi_a^1 \in H_x^{11/4}L_Z^2 \cap L_x^2 H_Z^{11/2}$*). *Furthermore, for all $m > \frac{9}{2}$,*

$$\|\Psi_a^0\|_{H_x^m L_z^2} \lesssim \|\gamma_a^0\|_{H^{m-9/4}} \lesssim \|\partial_x\theta_0\|_{H^{m-7/4}(\Omega)}, \qquad \|\Psi_a^0\|_{L_x^2 H_z^m} \lesssim \|\gamma_a^0\|_{H^{m/2-9/4}} \lesssim \|\partial_x\theta_0\|_{H^{m/2-7/4}(\Omega)},$$
$$\|\Psi_a^1\|_{H_x^m L_z^2} \lesssim \|\gamma_a^1\|_{H^{m-11/4}} \lesssim \|\partial_x\theta_0\|_{H^{m-5/4}(\Omega)}, \qquad \|\Psi_a^1\|_{L_x^2 H_z^m} \lesssim \|\gamma_a^1\|_{H^{m/2-11/4}} \lesssim \|\partial_x\theta_0\|_{H^{m/2-5/4}(\Omega)}.$$

*Additionally, the profiles $\Psi_a^0$ and $\Psi_a^1$ have exponential decay: there exists a universal constant $\bar{c} > 0$ such that, for any $Z_0 \geq 1$ and any $m \in \mathbb{N}$,*

$$\|\Psi_a^0\|_{H^m(\mathbb{T}\times(Z_0,+\infty))} \lesssim \|\theta_0\|_{H^1(\Omega)} \exp(-\bar{c}Z_0^{\frac{4}{5}}),$$
$$\|\Psi_a^1\|_{H^m(\mathbb{T}\times(Z_0,+\infty))} \lesssim \|\theta_0\|_{H^2(\Omega)} \exp(-\bar{c}Z_0^{\frac{4}{5}}).$$

**3.3.** ***Construction of an approximate solution.*** The idea is now to find a decomposition of $\theta$ as $\theta = \theta^{\mathrm{BL}} + \theta^{\mathrm{int}}$, where $\theta^{\mathrm{BL}}$ is a solution of

$$\partial_t\theta^{\mathrm{BL}} = \partial_x^2\Delta^{-2}\theta^{\mathrm{BL}} + S_r,$$

with a remainder term $S_r$ such that, for some $\delta > 0$,

$$\begin{aligned} S_r(t) &= O((1+t)^{-2}) && \text{in } L^2(\Omega),\\ S_r(t) &= O((1+t)^{-1-\delta}) && \text{in } H^4(\Omega),\\ \partial_t S_r(t) &= O((1+t)^{-3}) && \text{in } L^2(\Omega), \end{aligned}$$

and a boundary layer profile $\theta^{\mathrm{BL}}$ such that $\theta^{\mathrm{BL}}|_{\partial\Omega}(t=0) = \theta_0|_{\partial\Omega}$, $\partial_n\theta^{\mathrm{BL}}|_{\partial\Omega} = \partial_n\theta|_{\partial\Omega}$. Recall that the operator $\Delta^{-2}$ is endowed with homogeneous conditions for the trace and the normal derivative on the boundary of $\partial\Omega$.

As a consequence, the interior part $\theta^{\mathrm{int}}$ satisfies

$$\partial_t\theta^{\mathrm{int}} = \partial_x^2\Delta^{-2}\theta^{\mathrm{int}} - S_r$$

and the trace of $\theta^{\mathrm{int}}$ vanishes on $\partial\Omega$, together with its normal derivative. Thus we may apply Lemma 2.4 and Proposition 2.6, and we obtain $\|\theta^{\mathrm{int}}\|_{L^2} = O((1+t)^{-1})$, which will complete the proof of Theorem 1.2.

The main-order part of $\theta^{\mathrm{BL}}$ will be given by the profiles $\Theta_a^j$, $j = 0, 1$, $a \in \{\mathrm{top}, \mathrm{bot}\}$, constructed in Corollary 3.4. However, a few adjustments must be made in order to have a suitable decomposition:

- First, the profiles $\Theta_a^j$ must be truncated away from $z = 0$ and $z = 1$, so that their (exponentially small) trace does not pollute the opposite boundary. Since $\Theta_a^j$ has exponential decay, this introduces a remainder of order $\exp(-ct^{1/5})$, which will be included in $S_r$. More precisely, the error terms generated by this truncation will be dealt with thanks to the following lemma, whose proof is left to the reader:

**Lemma 3.5.** *Let $\Psi \in L^2(\mathbb{T} \times (0, +\infty))$ such that there exist $c, C > 0$ such that*

$$\int_{\mathbb{T}} \int_0^\infty |\Psi(x, Z)|^2 \exp(cZ^{\frac{4}{5}})\, \mathrm{d}Z\, \mathrm{d}x \le C < +\infty.$$

*Let $\zeta \in L^\infty(0, 1)$ such that $\operatorname{Supp} \zeta \subset \left(\frac{1}{4}, 1\right)$. Then there exists a constant $c' > 0$, depending only (and explicitly) on $c$, such that*

$$\|\Psi(x, (1+t)^{\frac{1}{4}} z)\zeta(z)\|_{L^2(\mathbb{T})} \lesssim C\|\zeta\|_\infty \exp(-c'(1+t)^{\frac{1}{5}}).$$

• More importantly, the main-order profiles $(\Theta_a^j, \Psi_a^j)$ for $j = 0, 1$ do not satisfy exactly

$$\Delta^2\big((1+t)^{-1}\Psi^j_{\mathrm{bot}}(x, (1+t)^{\frac{1}{4}} z)\big) = \Theta^j_{\mathrm{bot}}(x, (1+t)^{\frac{1}{4}} z).$$

Indeed, when constructing $\Psi_a^0$, we only kept the main-order terms in $\Delta^2$, i.e., the $z$-derivatives. It turns out that the term $2\partial_x^2\partial_z^2$ in the bilaplacian generates an error term in the equation which is not $O((1+t)^{-2})$. As a consequence, we introduce lower-order correctors, whose purpose is precisely to cancel this error term. We emphasize that the construction of such additional correctors is quite classical in multiscale problems. In order to determine the order at which the expansion can be stopped, we will rely on the following lemma, whose proof is postponed to the end of this section:

**Lemma 3.6.** • *Let $f \in H^4(\mathbb{T}, L^2(\mathbb{R}_+))$ such that there exist constants $c, C_1 > 0$ such that, for all $k \in \{0, \dots, 4\}$,*

$$|\partial_x^k f(x, Z)| \le C_1 \exp(-cZ^{\frac{4}{5}}) \quad \forall (x, Z) \in \mathbb{T} \times \mathbb{R}_+. \tag{3-7}$$

*Then there exists a constant $C$ depending only on $c$ such that*

$$\left\|\Delta^{-2}\big(f(x, (1+t)^{\frac{1}{4}} z)\chi(z)\big)\right\|_{L^2} \le \frac{CC_1}{(1+t)^{\frac{3}{4}}}.$$

*Furthermore, if*

$$\int_0^\infty Z^2 f(x, Z)\, \mathrm{d}Z = \int_0^\infty Z^3 f(x, Z)\, \mathrm{d}Z = 0 \quad \forall x \in \mathbb{T},$$

*this estimate becomes*

$$\left\|\Delta^{-2}\big(f(x, (1+t)^{\frac{1}{4}} z)\chi(z)\big)\right\|_{L^2} \le \frac{CC_1}{1+t}.$$

• *Let $f \in H^2(\mathbb{T}, L^2(\mathbb{R}_+))$ such that (3-7) holds for all $k \in \{0, 1, 2\}$. Then there exists a constant $C$ depending only on $c$ such that*

$$\left\|\Delta^{-2}\big(f(x, (1+t)^{\frac{1}{4}} z)\chi(z)\big)\right\|_{L^2} \le \frac{CC_1}{(1+t)^{\frac{1}{2}}}.$$

With the two above lemmas in mind, we define $\theta^{\mathrm{BL}}$ in the following way. Let $\chi \in C_c^\infty(\mathbb{R})$ be a cut-off function such that $\chi \equiv 1$ on $\left(-\frac{1}{4}, \frac{1}{4}\right)$, and $\operatorname{Supp} \chi \subset \left(-\frac{1}{2}, \frac{1}{2}\right)$. We look for $\theta^{\mathrm{BL}}$ in the form

$$\begin{aligned}\theta^{\mathrm{BL}}(t, x, z) &:= \sum_{j=0}^4 (1+t)^{-\frac{j}{4}} \Theta^j_{\mathrm{bot}}(x, (1+t)^{\frac{1}{4}} z)\chi(z) + \sum_{j=0}^4 (1+t)^{-\frac{j}{4}} \Theta^j_{\mathrm{top}}(x, (1+t)^{\frac{1}{4}}(1-z))\chi(z-1) \\ &=: \theta^{\mathrm{BL}}_{\mathrm{bot}} + \theta^{\mathrm{BL}}_{\mathrm{top}}\end{aligned}$$

and

$$\begin{aligned}\psi^{\mathrm{BL}}(t,x,z) &:= \sum_{j=0}^{4}(1+t)^{-1-\frac{j}{4}}\Psi^j_{\mathrm{bot}}(x,(1+t)^{\frac14}z)\chi(z)+\sum_{j=0}^{4}(1+t)^{-1-\frac{j}{4}}\Psi^j_{\mathrm{top}}(x,(1+t)^{\frac14}(1-z))\chi(z-1)\\ &=: \psi^{\mathrm{BL}}_{\mathrm{bot}}+\psi^{\mathrm{BL}}_{\mathrm{top}}.\end{aligned}$$

The profiles $\Theta^j_a$, $\Psi^j_a$ for $j=0,1$ and $a\in\{\mathrm{bot},\mathrm{top}\}$ were defined in the previous subsection, and we now proceed to define $\Theta^j_a$, $\Psi^j_a$ for $j\geq 2$. The reason why we stop the expansion at $j=4$ follows from Lemma 3.6, as we will see shortly.

We focus on the part near $z=0$, since the part near $z=1$ works identically. Setting $Z=(1+t)^{1/4}z$, we have

$$\frac{\partial}{\partial t}\theta^{\mathrm{BL}}_{\mathrm{bot}}=(1+t)^{-1}\sum_{j=0}^{4}(1+t)^{-\frac{j}{4}}\big[-\tfrac14 j\Theta^j_{\mathrm{bot}}(x,Z)+\tfrac14 Z\partial_Z\Theta^j_{\mathrm{bot}}(x,Z)\big]\chi(z).$$

For $j=0,1$, the bracketed term in the right-hand side is simply $\partial_x\Psi^j_{\mathrm{bot}}(x,(1+t)^{1/4}z)$. Similarly, we choose $\Psi^j_a$ for $j=2,3,4$ and $a\in\{\mathrm{bot},\mathrm{top}\}$ so that

$$\partial_x\Psi^j_a=-\tfrac14 j\Theta^j_a+\tfrac14 Z\partial_Z\Theta^j_a. \tag{3-8}$$

With this choice, we have

$$\partial_t\theta^{\mathrm{BL}}=\partial_x\psi^{\mathrm{BL}}.$$

There remains to choose $\Theta^j_a$ so that $\partial_x\psi^{\mathrm{BL}}=\Delta^{-2}\partial_x^2\theta^{\mathrm{BL}}+O((1+t)^{-2})$ in $L^2$. To that end, we observe

$$\begin{aligned}\Delta^2\psi^{\mathrm{BL}}_{\mathrm{bot}} &= \sum_{j=0}^{4}(1+t)^{-\frac{j}{4}}\partial_Z^4\Psi^j_{\mathrm{bot}}(x,Z)\chi(z)+2\sum_{j=0}^{4}(1+t)^{-\frac12-\frac{j}{4}}\partial_x^2\partial_Z^2\Psi^j_{\mathrm{bot}}(x,Z)\chi(z)\\ &\quad+\sum_{j=0}^{4}(1+t)^{-1-\frac{j}{4}}\partial_x^4\Psi^j_{\mathrm{bot}}(x,Z)\chi(z)\\ &\quad+\sum_{j=0}^{4}\sum_{k=0}^{3}\binom{k}{4}(1+t)^{-1+\frac{k-j}{4}}\partial_Z^k\Psi^j_{\mathrm{bot}}(x,Z)\chi^{(4-k)}(z)\\ &\quad+2\sum_{j=0}^{4}\sum_{k=0}^{1}\binom{k}{2}(1+t)^{-1+\frac{k-j}{4}}\partial_x^2\partial_Z^k\Psi^j_{\mathrm{bot}}(x,Z)\chi^{(2-k)}(z).\end{aligned}$$

The last two terms are handled by Lemma 3.5 (anticipating that $\Psi^j_a$ will have exponential decay for $j=2,3,4$).

We obtain

$$\begin{aligned}\Delta^2\psi^{\mathrm{BL}}_{\mathrm{bot}} &= \partial_x\theta^{\mathrm{BL}}_{\mathrm{bot}}+O(\exp(-c'(1+t)^{\frac15}))\\ &\quad+(1+t)^{-\frac12}[-\partial_x\Theta^2_{\mathrm{bot}}+\partial_Z^4\Psi^2_{\mathrm{bot}}+2\partial_x^2\partial_Z^2\Psi^0_{\mathrm{bot}}](x,Z)\chi(z)\\ &\quad+(1+t)^{-\frac34}[-\partial_x\Theta^3_{\mathrm{bot}}+\partial_Z^4\Psi^3_{\mathrm{bot}}+2\partial_x^2\partial_Z^2\Psi^1_{\mathrm{bot}}](x,Z)\chi(z)\\ &\quad+(1+t)^{-1}[-\partial_x\Theta^4_{\mathrm{bot}}+\partial_Z^4\Psi^4_{\mathrm{bot}}+2\partial_x^2\partial_Z^2\Psi^2_{\mathrm{bot}}+\partial_x^4\Psi^0](x,Z)\chi(z)\\ &\quad+\sum_{j\geq 5}(1+t)^{-\frac{j}{4}}\Phi^j_{\mathrm{bot}}(x,Z)\chi(z)\end{aligned}\tag{3-9}$$

for some functions $\Phi^j_{\rm bot}$ depending on the profiles $\Psi^j_{\rm bot}$ (for instance $\Phi^5 = 2\partial_x^2\partial_Z^2\Psi^3_{\rm bot} + \partial_x^4\Psi^1_{\rm bot}$). Thanks to Lemma 3.6, the inverse bilaplacian of the last term has a size of order $(1+t)^{-2}$ in $L^2$. Hence it will be included in the remainder $S_r$. Note that the reason why we need to stop the expansion in $\theta^{\rm BL}$ at $j=4$ is dictated by the above formula and by Lemma 3.6. If we stop the expansion for a lower $j$, then the remainder may be greater than $(1+t)^{-2}$ in $L^2$.

Therefore we focus on the terms of order $(1+t)^{-j/4}$ with $j = 2, 3, 4$. We treat the cases $j=2$ and $j=3$ simultaneously, and we will focus on the case $j=4$ later.

• Construction of $\Psi^j_a$ for $j=2,3$: Remembering (3-8), we choose $\Theta^j_a$ and $\Psi^j_a$ for $a\in\{\text{bot},\text{top}\}$ and $j=2,3$ so that

$$\partial_x\Psi^j_a = -\tfrac14 j\Theta^j_a + \tfrac14 Z\partial_Z\Theta^j_a,\quad -\partial_x\Theta^j_a + \partial_Z^4\Psi^j_a + 2\partial_x^2\partial_Z^2\Psi^{j-2}_a = 0,$$

endowed with the boundary conditions

$$\lim_{Z\to\infty}\Psi^j_a = 0,\quad \Psi^j_a(Z{=}0) = \partial_Z\Psi^j_a(Z{=}0) = \Theta^j_a(Z{=}0) = \partial_Z\Theta^j_a(Z{=}0) = 0.$$

As before, we note that the boundary conditions at $Z=0$ are redundant. Eliminating $\Theta^j_a$ from the equation, we find that $\Psi^j_a$ satisfies

$$\begin{aligned} Z\partial_Z^5\Psi^j_a - j\partial_Z^4\Psi^j_a &= 4\partial_x^2\Psi^j_a + S^j_a,\\ \Psi^j_a(Z{=}0) &= \partial_Z\Psi^j_a(Z{=}0) = 0,\\ \partial_Z^4\Psi^j_a &= -2\partial_x^2\partial_Z^2\Psi^{j-2}_a = -\frac1j S^j_a \quad \text{at } Z=0,\\ \partial_Z^5\Psi^j_a &= -2\partial_x^2\partial_Z^3\Psi^{j-2}_a = -\frac{1}{j-1}\partial_Z S^j_a \quad \text{at } Z=0,\\ \lim_{Z\to\infty}\Psi^j_a &= 0, \end{aligned}\tag{3-10}$$

where $S^j_a = -2(Z\partial_Z - j)\partial_x^2\partial_Z^2\Psi^{j-2}_a$. Therefore

$$\partial_Z^j S^j_a = -2Z\partial_Z^{j+3}\partial_x^2\Psi^{j-2}_a = -8\partial_x^4\partial_Z^{j-2}\Psi^{j-2}_a.$$

As a consequence, $\partial_Z^j\Psi^j_a$ is a solution of

$$\begin{cases} Z\partial_Z^5\partial_Z^j\Psi^j_a = 4\partial_x^2\partial_Z^j\Psi^j_a - 8\partial_x^4\partial_Z^{j-2}\Psi^{j-2}_a,\\ \partial_Z^j\Psi^j_a = 0 \quad \text{at } Z=0,\\ \partial_Z^4\Psi^j_a = -2\partial_x^2\partial_Z^2\Psi^{j-2}_a,\quad \partial_Z^5\Psi^j_a = -2\partial_x^2\partial_Z^3\Psi^{j-2}_a \quad \text{at } Z=0,\\ \lim_{Z\to\infty}\partial_Z^j\Psi^j_a = 0. \end{cases}$$

Note that the boundary condition $\partial_Z^j\Psi^j_a(Z{=}0) = 0$ follows from the identity

$$\partial_x\partial_Z^j\Psi^j_a = \tfrac14 Z\partial_Z^{j+1}\Theta^j_a.$$

Taking the Fourier transform with respect to $x$, we observe that $\widehat{\partial_Z^j\Psi_a}(k)$ satisfies (3-4) with nonhomogeneous boundary conditions of type (iii) (for $j=2$) or (iv) (for $j=3$). Using the Fourier representations

(3-5) and (3-6) for $\Psi^0$ and $\Psi^1$, we anticipate that $\Psi_a^2$ and $\Psi_a^3$ can be written as

$$\begin{aligned} \Psi_a^2(x,Z) &= \sum_{k\in\mathbb{Z}\setminus\{0\}} |k|^{-1}\hat{\gamma}_{a,k}^0\, \chi_2(|k|^{\frac12}Z)e^{ikx}, \\ \Psi_a^3(x,Z) &= \sum_{k\in\mathbb{Z}\setminus\{0\}} |k|^{-\frac32}\hat{\gamma}_{a,k}^1\, \chi_3(|k|^{\frac12}Z)e^{ikx}, \end{aligned} \tag{3-11}$$

with $\chi_2, \chi_3 \in C^\infty((0,+\infty))$ decaying like $\exp(-\bar{c}Z^{4/5})$. The precise construction of $\chi_2$ and $\chi_3$ will be performed below. We obtain the following result:

**Lemma 3.7.** *Let $a \in \{\text{top}, \text{bot}\}$ and $\gamma_a^0, \gamma_a^1 \in L^2(\mathbb{T})$. Consider the solutions $\Psi_a^0, \Psi_a^1$ of* (3-2), (3-3) *given by Corollary 3.4.*

*Then there exist unique solutions $\Psi_a^2 \in H_x^{5/4}L_Z^2 \cap L_x^2H_Z^{5/2}$, $\Psi_a^3 \in H_x^{7/4}L_Z^2 \cap L_x^2H_Z^{7/2}$ of* (3-10). *Furthermore, for any $m \in \mathbb{N}$,*

$$\begin{aligned} \|\Psi_a^2\|_{H_x^mL_Z^2} &\lesssim \|\partial_x\theta_0\|_{H^{m-3/4}(\Omega)}, & \|\Psi_a^2\|_{L_x^2H_Z^m} &\lesssim \|\partial_x\theta_0\|_{H^{m/2-3/4}(\Omega)}, \\ \|\Psi_a^3\|_{H_x^mL_Z^2} &\lesssim \|\partial_x\theta_0\|_{H^{m-1/4}(\Omega)}, & \|\Psi_a^3\|_{L_x^2H_Z^m} &\lesssim \|\partial_x\theta_0\|_{H^{m/2-1/4}(\Omega)}. \end{aligned}$$

*Additionally, the profiles $\Psi_a^2$ and $\Psi_a^3$ have exponential decay: for any $Z_0 \geq 1$, for any $m \in \mathbb{N}$,*

$$\begin{aligned} \|\Psi_a^2\|_{H^m(\mathbb{T}\times(Z_0,+\infty))} &\lesssim \|\theta_0\|_{H^1(\Omega)} \exp(-\bar{c}Z_0^{\frac45}), \\ \|\Psi_a^3\|_{H^m(\mathbb{T}\times(Z_0,+\infty))} &\lesssim \|\theta_0\|_{H^2(\Omega)} \exp(-\bar{c}Z_0^{\frac45}). \end{aligned}$$

*Proof.* In view of (3-11), it is sufficient to construct $\chi_2$ and $\chi_3$. We first construct the solution $\phi_j$ of

$$\begin{cases} Z\partial_Z^5\phi_j(Z) = -4\phi_j(Z) - 8\partial_Z^{j-2}\chi_{j-2}, \\ \phi_j(0) = 0, \quad \partial_Z^{4-j}\phi_j(0) = -2\partial_Z^2\chi_{j-2}(0), \quad \partial_Z^{5-j}\phi_j(0) = -2\partial_Z^3\chi_{j-2}(0), \\ \lim_{Z\to\infty}\phi_j(Z) = 0. \end{cases}$$

Note that after a suitable lifting, $\phi_j$ satisfies (3-4) with the boundary conditions (iii) from Lemma 3.2 (for $j=2$) or (iv) (for $j=3$). Hence the existence and uniqueness of $\phi_j$ (and its exponential decay) follow from Lemma 3.2. Now, define $\chi_j$ as

$$\partial_Z^j\chi_j = \phi_j, \quad \partial_Z^k\chi_j(+\infty) = 0 \quad \text{for } 0 \leq k \leq j-1.$$

It follows that $\chi_j$ decays like $\exp(-\bar{c}Z^{4/5})$. Furthermore, by construction

$$\partial_Z^j[Z\partial_Z^5\chi_j - j\partial_Z^4\chi_j + 4\chi_j - 2(Z\partial_Z - j)\partial_Z^2\chi_{j-2}] = 0.$$

We deduce that $Z\partial_Z^5\chi_j - j\partial_Z^4\chi_j + 4\chi_j + 2(Z\partial_Z - j)\partial_Z^2\chi_{j-2}$ is a polynomial of order at most $j-1$, which has exponential decay at infinity. Therefore, the following equality holds:

$$Z\partial_Z^5\chi_j - j\partial_Z^4\chi_j + 4\chi_j - 2(Z\partial_Z - j)\partial_Z^2\chi_{j-2} = 0.$$

Taking the trace of the above identity at $Z=0$, we infer that $\chi_j(0)=0$. In a similar way, we also find that $\chi_j'(0)=0$. Now, defining $\Psi_a^j$ by (3-11), we obtain that $\Psi_a^j$ satisfies (3-10). The Sobolev estimates are then a consequence of the Fourier representation formula. □

• Construction of $\Psi_a^4$: The definitions of $\Psi_a^4$ and $\Theta_a^4$ are similar. We choose $\Psi_a^4$ such that

$$\begin{cases} Z\partial_Z^5\Psi_a^4 - 4\partial_Z^4\Psi_a^4 = 4\partial_x^2\Psi_a^4 + S_a^4, \\ \Psi_a^4(0) = \partial_Z\Psi_a^4(0) = 0, \quad \partial_Z^4\Psi_a^4 = -\frac{1}{4}S_a^4, \quad \partial_Z^5\Psi_a^4 = -\frac{1}{3}\partial_Z S_a^4 \quad \text{at } Z = 0, \\ \lim_{Z\to\infty}\Psi_a^4 = 0, \end{cases}$$

where

$$S_a^4 = -(Z\partial_Z - 4)(2\partial_x^2\partial_Z^2\Psi_a^2 + \partial_x^4\Psi_a^0).$$

Therefore the Fourier transform of $\partial_z^4\Psi_a^4$, after a suitable lifting, is a solution of (3-4). The main difference with the construction of $\Psi_a^j$ for $j \le 3$ lies in the fact that $\partial_Z^4\Psi_a^4$ is not fully determined. Indeed, we lack a boundary condition on $\partial_Z^k\Psi_a^4$ for some $k \ge 6$. Once again, this phenomenon (a high-order corrector is underdetermined) is quite common in multiscale problems. In fact it turns out that $\Psi_a^4$ could be determined in a unique fashion if we were looking for a higher-order expansion (see Remark 3.8). In this case, we should choose $\Psi_{\text{bot}}^4$ so that $\partial_Z^4\Theta_{\text{bot}}^4|_{Z=0}$ lifts the trace of $\Delta^2\theta|_{z=0}$. In the present case, since we merely wish to close the first-order expansion, we simply further require that $\partial_Z^8\Psi_a^4|_{|Z=0} = 0$, so that the lifted Fourier transform of $\partial_Z^4\Psi_a^4$ satisfies the boundary conditions (i) of Lemma 3.2. We conclude that $\Psi_a^4$ is well-defined and satisfies the same estimates as $\Psi_a^j$ for $j \le 3$. The details of the proof are left to the reader.

**3.4.** ***Estimate of the remainder and conclusion.*** At this stage, we have constructed $\theta^{\text{BL}}$ such that, for all $t \ge 0$,

$$\theta^{\text{BL}}(t)|_{\partial\Omega} = \theta(t)|_{\partial\Omega} = \theta_0|_{\partial\Omega},$$
$$\partial_{\boldsymbol{n}}\theta^{\text{BL}}(t)|_{\partial\Omega} = \partial_{\boldsymbol{n}}\theta(t)|_{\partial\Omega} = \partial_{\boldsymbol{n}}\theta_0|_{\partial\Omega}$$

and

$$\partial_t\theta^{\text{BL}} = \Delta^{-2}\partial_x^2\theta^{\text{BL}} + \Delta^{-2}\partial_x T_r + O(\exp(-c(1+t)^{\frac{1}{5}}) \quad \text{in } L^2,$$

where $T_r = T_{\text{top}} + T_{\text{bot}}$ and

$$T_{\text{bot}} := \Big[\sum_{j\ge 5}(1+t)^{-\frac{j}{4}}\Phi_{\text{bot}}^j(x, (1+t)^{\frac{1}{4}}z)\chi(z)\Big],$$

with a similar expression for $T_{\text{top}}$. According to Lemma 3.6,

$$\|\Delta^{-2}\partial_x T_r\|_{L^2} \lesssim \|\theta_0\|_{H^s}(1+t)^{-2}, \quad \|\partial_t\Delta^{-2}\partial_x T_r\|_{L^2} \lesssim \|\theta_0\|_{H^s}(1+t)^{-3}.$$

Furthermore,

$$\|\Delta^{-2}\partial_x T_r\|_{H^4} \lesssim \|\partial_x T_r\|_{L^2} \lesssim \|\theta_0\|_{H^s}(1+t)^{-\frac{5}{4}}$$

for some finite (and computable) index $s > 0$. Therefore $\theta^{\text{int}} = \theta - \theta^{\text{BL}}$ solves

$$\partial_t\theta^{\text{int}} = \partial_x^2\Delta^{-2}\theta^{\text{int}} - \Delta^{-2}\partial_x T_r + O(\exp(-c(1+t)^{\frac{1}{5}}),$$

and $\theta^{\text{int}} = \partial_{\boldsymbol{n}}\theta^{\text{int}} = 0$ on $\partial\Omega$. We first apply Lemma 2.4 to $\Delta^2\theta^{\text{int}}$ and find that $\|\Delta^2\theta^{\text{int}}(t)\|_{L^2} \lesssim \|\theta_0\|_{H^s}$ for all $t \ge 0$, for some finite $s$. From there, we apply Proposition 2.6 to $\partial_x^2\theta^{\text{int}}$ with $\alpha = 0$, and we deduce that $\|\partial_x^2\theta^{\text{int}}(t)\|_{L^2} \lesssim \|\theta_0\|_{H^s}(1+t)^{-1}$. As in Section 2, estimates on $\psi^{\text{int}}$ can be obtained by deriving bounds on $\partial_t\theta^{\text{int}}$. More precisely, applying Proposition 2.6 to $\partial_t\theta^{\text{int}}$, we find that $\|\partial_t\theta^{\text{int}}(t)\|_{L^2} \lesssim \|\theta_0\|_{H^s}(1+t)^{-2}$, and therefore $\|\partial_x^2\Delta^{-2}\theta^{\text{int}}\|_{L^2} \lesssim \|\theta_0\|_{H^s}(1+t)^{-2}$. This completes the proof of Theorem 1.2.

**Remark 3.8** (construction of an approximation at any order). Since $\Delta^2\theta$ solves the same equation as $\theta$, one can easily iterate this construction. More precisely, if $\theta_0 \in H^{4k}$, it can be proved that there exist sequences of profiles $(\Theta^j_{\mathrm{bot}}, \Theta^j_{\mathrm{top}})_{0\le j\le 4k}$ such that the following result holds:

$$\theta(t,x) = \sum_{j=1}^{4k} (1+t)^{-\frac{j}{4}} [\Theta^j_{\mathrm{bot}}(x, (1+t)^{\frac14} z)\chi(z) + \Theta^j_{\mathrm{top}}(x, (1+t)^{\frac14}(1-z))\chi(z-1)] + \theta^j_{\mathrm{rem}}(t)$$

and

$$\|\theta^j_{\mathrm{rem}}(t)\|_{L^2} \lesssim \frac{1}{(1+t)^k}, \quad \|\theta^j_{\mathrm{rem}}(t)\|_{H^{4k}} \lesssim 1.$$

For instance, the role of $\Theta^{4j}_{\mathrm{bot}}$ is to lift the trace of $\Delta^{2j}\theta$ at $z=0$, the one of $\Theta^{4j+1}_{\mathrm{top}}$ is to lift the one of $\partial_z \Delta^{2j}\theta$ at $z=1$, etc.

The details of the construction are very similar to the ones of the profiles $\Theta^j_a$ for $0 \le j \le 3$ above and are left to the reader.

**3.5.** ***Proof of Lemma 3.6.*** We first define a function $f_1$ such that

$$\partial_Z^4 f_1 = f,$$

and $\partial_Z^k f_1(+\infty) = 0$ for $0 \le k \le 3$. Note that the exponential decay assumption on $f$ ensures that $f_1$ exists, and $f_1 \in W^{4,\infty} \cap H^4$. Moreover, for $0 \le m_1, m_2 \le 4$,

$$|\partial_x^{m_1}\partial_Z^{m_2} f_1(x,Z)| \le C\exp(-cZ^{\frac15}),$$

with possibly different constants $C$ and $c$. Setting $Z = (1+t)^{1/4}z$, we infer that

$$\begin{aligned}\Delta^2\big((1+t)^{-1} f_1(x,Z)\chi(z)\big) &= f(x,Z)\chi(z) + 2(1+t)^{-\frac12}\partial_x^2\partial_Z^2 f_1(x,Z)\chi(z)\\ &\quad + (1+t)^{-1}\partial_x^4 f_1(x,Z)\chi(z) + O(\exp(-ct^{\frac15})) \quad \text{in } L^2,\end{aligned}$$

where the term $O(\exp(ct^{1/5}))$ stems from the commutator-involving derivatives of $\chi$ (see Lemma 3.5). Note that $\partial_x^2\partial_Z^2 f_1$ satisfies the same decay assumptions as $f$, and therefore we can lift it by another corrector $f_2$ such that

$$\partial_Z^4 f_2 = -2\partial_x^2\partial_Z^2 f_1,$$

i.e., $\partial_Z^2 f_2 = -2\partial_x^2 f_1$. Therefore

$$\Delta^2\big(\big((1+t)^{-1} f_1(x,Z) + (1+t)^{-\frac32} f_2(x,Z)\big)\chi(z)\big) = f(x,Z)\chi(z) + O((1+t)^{-1}) \quad \text{in } H^{-2}.$$

The only remaining issue lies in the fact that $f_1$, $f_2$ and their normal derivatives do not vanish on the boundary. Hence we set $a_i(x) = f_i(x,0)$, $b_i(x) = \partial_Z f_i(x,0)$, and we add a corrector

$$f_3(t,x,z) := -\sum_{i=1,2} (1+t)^{-\frac{i-1}{2}} (a_i(x) + z(1+t)^{\frac14} b_i(x))\chi(z).$$

Now

$$\Delta^2\big(\big((1+t)^{-1} f_1(x,Z) + (1+t)^{-\frac32} f_2(x,Z)\big)\chi(z) + (1+t)^{-1} f_3\big) = f(x,Z)\chi(z) + O((1+t)^{-\frac34}) \quad \text{in } H^{-2},$$

and for $k = 0, 1$

$$\partial_z^k\big(((1+t)^{-1}f_1(x,Z)+(1+t)^{-\frac{3}{2}}f_2(x,Z))\chi(z)+(1+t)^{-1}f_3\big)|_{\partial\Omega}=0.$$

It follows that

$$\big((1+t)^{-1}f_1(x,Z)+(1+t)^{-\frac{3}{2}}f_2(x,Z)\big)\chi(z)+(1+t)^{-1}f_3=\Delta^{-2}(f(x,Z)\chi(z))+O((1+t)^{-\frac{3}{4}})\quad\text{in } H^2.$$

Let us now prove that when $\int_0^\infty Z^2 f(\,\cdot\,,Z)\,\mathrm{d}Z=\int_0^\infty Z^3 f(\,\cdot\,,Z)\,\mathrm{d}Z=0$, we gain an additional factor $(1+t)^{-1/4}$. It can be easily checked that

$$f_1|_{Z=0}=\frac{1}{6}\int_0^\infty Z^3 f(\,\cdot\,,Z)\,\mathrm{d}Z=0,$$
$$\partial_Z f_1|_{Z=0}=-\frac{1}{2}\int_0^\infty Z^2 f(\,\cdot\,,Z)\,\mathrm{d}Z=0.$$

Hence, with the notation above, $a_1=b_1=0$ and therefore $f_3=O((1+t)^{-1/4})$. With the same arguments, we infer that

$$\big((1+t)^{-1}f_1(x,Z)+(1+t)^{-\frac{3}{2}}f_2(x,Z)\big)\chi(z)+(1+t)^{-1}f_3=\Delta^{-2}(f(x,Z)\chi(z))+O((1+t)^{-1})\quad\text{in } H^2.\quad\square$$

**Remark 3.9.** Note that the first statement of Lemma 3.6 provides a better decay of the $H^{-2}$ norm, but the second one requires less horizontal derivatives on $f$. In the next section, we will also use the following variant: Assume that there exists a sequence $(\gamma_k)_{k\in\mathbb{Z}}$ such that

$$f(x,Z)=\sum_{k\in\mathbb{Z}\setminus\{0\}}\gamma_k e^{ikx}\varphi(|k|^{\frac{1}{2}}Z),$$

where $\varphi\in C^\infty(\mathbb{R})$ decays like $C_1\exp(-cZ^{4/5})$. Then, following the previous computations,

$$f_1(x,Z)=\sum_{k\in\mathbb{Z}\setminus\{0\}}|k|^{-2}\gamma_k e^{ikx}\varphi^{(-4)}(|k|^{\frac{1}{2}}Z),$$
$$f_2(x,Z)=2\sum_{k\in\mathbb{Z}\setminus\{0\}}|k|^{-1}\gamma_k e^{ikx}\varphi^{(-6)}(|k|^{\frac{1}{2}}Z),$$

where $\partial_Z^m\varphi^{(-m)}=\varphi$, and $\varphi^{(-m)}(+\infty)=0$. Hence

$$\|\Delta^{-2}(f(x,(1+t)^{\frac{1}{4}}z)\chi(z))\|_{L^2}\lesssim C_1\Big(\sum_{k\in\mathbb{Z}\setminus\{0\}}|k|^2|\gamma_k|^2\Big)^{\frac{1}{2}}(1+t)^{-\frac{3}{4}}.$$

## 4. Long-time boundary layers in the nonlinear setting: proof of Theorem 1.3

We now go back to the long-time analysis of (1-7) when $\theta_0'=\partial_{\boldsymbol{n}}\theta_0'=0$ on $\partial\Omega$. We recall (see Theorem 1.1) that in this case $\theta'(t)$ converges towards zero in $H^s$ for all $s<4$ as $t\to\infty$.

A natural question is to investigate whether the algebraic decay rate provided by Theorem 1.1 can be improved, possibly at the cost of a stronger regularity requirement on the initial data. In other words, if we assume that $\theta_0\in H^s$ with $s$ large, can we prove a uniform $H^s$ bound on a solution, and thereby a higher decay estimate on $\theta'$?

As explained in the Introduction and in Section 2, such a result does not follow immediately from an induction argument. Indeed, the traces of $\Delta^2\theta'$ and of $\partial_{\boldsymbol{n}}\Delta^2\theta'$ do not vanish on the boundary (even when the traces of $\Delta^2\theta'_0$ and of $\partial_{\boldsymbol{n}}\Delta^2\theta'_0$ do), and therefore we cannot apply Proposition 2.6 to $\Delta^2\theta'$.

However, it turns out that when $\partial_z^2\bar{\theta}_0|_{\partial\Omega} = 0$, we can use (a variant of) the linear analysis of Section 3 to analyze the long-time behavior of $\Delta^2\theta'$. In other words, in this case, there are boundary layers in the vicinity of the boundary, but they are driven by a linear mechanism. Theorem 1.3 will follow.

To that end, the strategy is to consider the equation satisfied by $\Delta^2\theta'$. As we have seen previously, the structure of the equation is overall the same. The main difference lies in the fact that the traces of $\Delta^2\theta'$ and $\partial_{\boldsymbol{n}}\Delta^2\theta'$ do not vanish on the boundary, which makes the situation rather close to the one described in Theorem 1.2. Following the methodology of the previous section, we may lift them thanks to a corrector which remains linear at main order. Modifying slightly the bootstrap argument from Section 2 in order to account for these boundary layers, we eventually prove Theorem 1.3, or rather the following more precise version:

**Proposition 4.1.** *There exists a universal constant $\varepsilon_0$ such that the following statement holds. Let $\theta_0 \in H^{14}(\Omega)$ such that $\theta_0|_{\partial\Omega} = \partial_{\boldsymbol{n}}\theta_0|_{\partial\Omega} = 0$, and $\partial_z^2\bar{\theta}_0|_{\partial\Omega} = 0$. Assume that $\|\theta_0\|_{H^{14}} \le \varepsilon_0$.*

*There exists a boundary layer profile $\theta^{\mathrm{BL}} \in L^\infty_{\mathrm{loc}}(\mathbb{R}_+, H^9(\Omega))$, given by*

$$\theta^{\mathrm{BL}} = \sum_{j=0}^{4} (1+t)^{-1-\frac{j}{4}} \big(\Theta^j_{\mathrm{bot}}(x, Z_{\mathrm{bot}}) + \Theta^j_{\mathrm{top}}(x, Z_{\mathrm{top}})\big),$$

*where $Z_{\mathrm{bot}} = z(1+t)^{1/4}$, $Z_{\mathrm{top}} = (1-z)(1+t)^{1/4}$, and $\Theta^j_a \in H^9(\mathbb{T}\times\mathbb{R}_+)$, such that the following estimates hold on $\theta^{\mathrm{rem}} := \theta' - \theta^{\mathrm{BL}}$ for all $t \ge 0$:*

$$\begin{aligned}
\|\partial_x^4\theta^{\mathrm{rem}}(t)\|_{L^2} &\lesssim \|\theta_0\|_{H^{14}}(1+t)^{-2},\\
\|\partial_x^2\Delta^2\theta^{\mathrm{rem}}(t)\|_{L^2} &\lesssim \|\theta_0\|_{H^{14}}(1+t)^{-1},\\
\|\Delta^4\theta^{\mathrm{rem}}(t)\|_{L^2} &\lesssim \|\theta_0\|_{H^{14}},\\
\|\partial_x^6\Delta^{-2}\theta^{\mathrm{rem}}(t)\|_{L^2} &\lesssim \|\theta_0\|_{H^{14}}(1+t)^{-3}.
\end{aligned}$$

**Remark 4.2.** Note that the assumptions of Proposition 4.1 are slightly weaker than the ones of Theorem 1.3. Indeed, we do not require that $\theta_0 \in H^3_0$, but rather that $\theta_0 \in H^2_0$ and $\partial_z^2\bar{\theta}_0 = 0$. According to Lemma 2.1, these properties are propagated by the equation. Using the notation of Section 2 and setting $G = \partial_z\bar{\theta}$, we infer that $G$ and $\partial_z G$ vanish at $z = 0$ and $z = 1$.

**4.1.** ***General strategy.*** Following the same strategy as in Section 3, we look for an ansatz for $\theta'$ as a sum of a boundary layer part $\theta^{\mathrm{BL}}$, whose role is to lift the trace of $\Delta^2\theta'$ and $\partial_{\boldsymbol{n}}\Delta^2\theta'$ on the boundary, and an interior part $\theta^{\mathrm{int}}$, which vanishes at a high order on the boundary, and for which we will therefore be able to prove better decay estimates. Let us give a few additional details on these two parts:

- As in the previous section, the boundary layer term will be defined as an asymptotic expansion in powers of $(1+t)^{-1/4}$, and the width of the boundary layers will also be $(1+t)^{-1/4}$. The different terms of the expansion will be constructed recursively: the main-order terms will lift the traces of $\Delta^2\theta'$ and

$\partial_{\boldsymbol{n}}\Delta^2\theta'$ (or rather, their limits as $t\to\infty$), and the next-order terms will correct error terms generated by the first-order ones. The precise construction of the boundary layer is the purpose of Section 4.4 below.

- In fact, $\Delta^2\theta^{\text{rem}}=\Delta^2(\theta'-\theta^{\text{BL}})$ is not identically zero on the boundary, but it is of order $(1+t)^{-1}$. Hence we construct additional small correctors $\theta_c$, $\sigma^{\text{NL}}_{\text{lift}}$, which handle the remaining traces and part of the error term.

- Thanks to the design of the boundary layer, the remaining interior part $\theta^{\text{int}}=\theta^{\text{rem}}-\theta_c-\sigma^{\text{NL}}_{\text{lift}}$ is such that

$$\theta^{\text{int}}=\partial_{\boldsymbol{n}}\theta^{\text{int}}=\Delta^2\theta^{\text{int}}=\partial_{\boldsymbol{n}}\Delta^2\theta^{\text{int}}=0\quad\text{on }\partial\Omega.$$

As a consequence, $\Delta^4\theta^{\text{int}}$ satisfies assumptions that are similar to those of Lemma 2.4, and it is reasonable to expect that $\|\Delta^4\theta^{\text{int}}\|_{L^2}$ remains uniformly bounded. Applying Proposition 2.6 first to $\partial_x^2\Delta^2\theta^{\text{int}}$, and then to $\partial_x^4\theta^{\text{int}}$, we infer that $\|\partial_x^2\Delta^2\theta^{\text{int}}\|_{L^2}=O((1+t)^{-1})$ and $\|\partial_x^4\theta^{\text{int}}\|_{L^2}=O((1+t)^{-2})$. We will use a bootstrap argument to propagate these bounds; the corresponding argument is described in Section 4.5.

Before constructing $\theta^{\text{BL}}$ and proving the decay estimates on $\theta^{\text{int}}$, some preliminary (and somewhat technical) steps are in order. The traces of $\Delta^2\theta'$ and $\partial_{\boldsymbol{n}}\Delta^2\theta'$ need to be decomposed as an asymptotic expansion in powers of $(1+t)^{-1/4}$, in order to identify the relevant boundary conditions for the terms in the expansion of $\theta^{\text{BL}}$. This is performed in Lemma 4.9 below, whose proof involves some high-regularity bounds on $\theta$. This is the main reason for the requirement $\theta_0\in H^{14}$ from Theorem 1.3. As a consequence, the organization of the rest of this section is the following. In Section 4.2, we prove some quantitative $H^s$ bounds ($s\le 14$) on $\theta'$ under our bootstrap assumption. In Section 4.3, we provide a decomposition of $\Delta^2\theta'|_{\partial\Omega}$ and $\partial_{\boldsymbol{n}}\Delta^2\theta'|_{\partial\Omega}$ under the bootstrap assumption. The main results of each section are given in the beginning of the corresponding section. The reader wishing to avoid the technicalities may jump to Section 4.4, in which we construct the boundary layer, using the decomposition of Section 4.3 together with arguments from Section 3. Eventually, we close the bootstrap argument in Section 4.5.

Let us now introduce the bootstrap assumption that will be used throughout this section. We shall decompose $\theta'$ as $\theta'=\theta^{\text{BL}}+\theta^{\text{rem}}$. As explained above, the remainder $\theta^{\text{rem}}$ does not satisfy $\Delta^2\theta^{\text{rem}}|_{\partial\Omega}=\partial_{\boldsymbol{n}}\Delta^2\theta^{\text{rem}}|_{\partial\Omega}=0$, and therefore $\theta^{\text{rem}}$ will be further decomposed into a sum of correctors and an interior term.

The term $\theta^{\text{BL}}$ will take the form

$$\theta^{\text{BL}}=\frac{1}{1+t}\Theta_{\text{top}}(x,(1+t)^{\frac14}(1-z))+\frac{1}{1+t}\Theta_{\text{bot}}(x,(1+t)^{\frac14}z)+\text{l.o.t.},\tag{4-1}$$

with boundary layer profiles $\Theta_{\text{top}}$, $\Theta_{\text{bot}}$ such that

$$\|\Theta_a\|_{H^9(\mathbb{T}\times\mathbb{R}_+)}\le B$$

for some constant $B>0$. Note that the amplitude of the boundary layer term $\theta^{\text{BL}}$ is $O((1+t)^{-1})$, whereas we recall that the amplitude of the boundary layer term in Section 3 was $O(1)$ (compare (4-1) with (1-6)). This is directly linked to the fact that $\theta'|_{\partial\Omega}=\partial_{\boldsymbol{n}}\theta'|_{\partial\Omega}=0$ in this section, while these quantities were nonzero in Section 3. However we keep the same notation for the sake of simplicity.

The remainder term $\theta^{\mathrm{rem}}$ will satisfy the bootstrap assumptions

$$\begin{aligned}\sup_{t\in[0,T]} (1+t)^2\|\partial_x^4\theta^{\mathrm{rem}}(t)\|_{L^2}+\|\Delta^4\theta^{\mathrm{rem}}(t)\|_{L^2}\le B,\\ \sup_{t\in[0,T]} (1+t)^3\|\partial_t\partial_x^4\theta^{\mathrm{rem}}(t)\|_{L^2}+(1+t)^3\|\partial_x^5\psi^{\mathrm{rem}}\|_{L^2}\le B,\end{aligned}\tag{4-2}$$

where $\psi^{\mathrm{rem}}=\Delta^{-2}\partial_x\theta^{\mathrm{rem}}$.

As a consequence, our bootstrap assumptions on $\theta'$ read as follows:

$$\begin{aligned}&\forall t\in(0,T),\ \forall k\in\{4,\dots,8\}, \qquad \|\partial_x^k\theta'\|_{L^2}\le B(1+t)^{-\frac{9}{8}}+B(1+t)^{\frac{k-8}{2}},\\ &\forall t\in(0,T),\ \forall k\in\{0,\dots,8\}, \qquad \|\partial_z^k\theta'\|_{L^2}\le B(1+t)^{-\frac{9}{8}+\frac{k}{4}}, \quad \|\partial_x^4\Delta^2\theta'\|_{L^2}\le B,\\ &\qquad\forall t\in(0,T), \quad \|\partial_x^5\psi(t)\|_{L^2}\le B(1+t)^{-\frac{17}{8}}.\end{aligned}\tag{4-3}$$

Note that these assumptions imply in particular that for $0\le k\le 3$

$$\|\psi\|_{W^{k,\infty}}\lesssim B(1+t)^{-\frac{17}{8}+\frac{k+1}{4}} \quad \forall t\in[0,T].\tag{4-4}$$

Let us prove inequality (4-4) in the case $k=3$ (the other cases are treated in a similar fashion and left to the reader.) By the Gagliardo–Nirenberg–Sobolev inequality,

$$\|\psi\|_{W^{3,\infty}}\lesssim\|\psi\|_{L^2}^{\frac{1}{5}}\|\psi\|_{H^5}^{\frac{4}{5}}+\|\psi\|_{L^2}.$$

By (4-3), $\|\psi\|_{L^2}\lesssim\|\partial_x^5\psi\|_{L^2}\lesssim B(1+t)^{-17/8}$, while

$$\|\partial_z^5\psi\|_{L^2}\lesssim\|\partial_z\partial_x\theta'\|_{L^2}\lesssim\|\partial_x^2\theta'\|_{L^2}^{\frac{1}{2}}\|\partial_z^2\theta'\|_{L^2}^{\frac{1}{2}}\lesssim B(1+t)^{-\frac{7}{8}}.$$

Estimate (4-4) follows.

We also infer from (4-3) some interpolated inequalities (which may be suboptimal when compared to the bootstrap assumption on $\partial_x^4\Delta^2\theta'$, depending on the values of $k,\ell$): for all $k,\ell\ge 0$ such that $k+\ell\le 8$, we have

$$\|\partial_x^k\partial_z^\ell\theta'\|_{L^2}\le\|\partial_x^{k+\ell}\theta'\|_{L^2}^{\frac{k}{k+\ell}}\|\partial_z^{k+\ell}\theta'\|_{L^2}^{\frac{\ell}{k+\ell}}\lesssim B(1+t)^{-\frac{9}{8}+\frac{\ell}{4}}+B(1+t)^{\frac{k}{2}+\frac{\ell}{4}-4\frac{k}{k+\ell}-\frac{9}{8}\frac{\ell}{k+\ell}}.\tag{4-5}$$

**4.2. *High regularity bounds under the bootstrap assumption.*** The purpose of this subsection is to prove the following estimates, which are the analogue of Lemma 2.13 in higher regularity (see also Remark 2.18):

**Lemma 4.3.** *Let $\theta=\theta'+\bar\theta$ be a solution of* (1-7), *and assume that $\theta_0\in H^{14}$ satisfies the assumptions of Theorem 1.3. Let $T>0$ be such that the bounds* (4-3) *hold on* $(0,T)$ *for some constant $B\in(0,1)$. Assume furthermore that $\|\theta_0\|_{H^{14}}\le B$. Then, for all $t\in[0,T]$,*

$$\begin{aligned}&\|\bar\theta(t)\|_{H^6}\lesssim B, \quad \|\bar\theta(t)\|_{H^9}\lesssim B(1+t)^{\frac{1}{2}}, \qquad \|\theta\|_{H^{14}}\lesssim B(1+t)^{\frac{5}{2}},\\ &\|\partial_x^8\Delta^2\theta'\|_{L^2}\lesssim B, \quad \|\partial_x^{10}\theta'\|_{L^2}\lesssim B(1+t)^{-1}, \quad \|\partial_x^{10}\psi\|_{L^2}\lesssim B(1+t)^{-2}.\end{aligned}$$

*Proof.* First, recalling that

$$\partial_t\bar\theta=-\overline{\nabla^\perp\psi\cdot\nabla\theta'}$$

and using the bootstrap assumptions (4-3) and (4-4) together with the tame estimates (2-15), we infer that

$$\|\partial_t \partial_z^6 \bar{\theta}\|_{L^2} \lesssim B^2(1+t)^{-\frac{5}{4}},$$

and thus $\|\bar{\theta}(t)\|_{H^6} \lesssim \|\theta_0\|_{H^6} + B^2 \lesssim B$. A similar argument also shows that $\|\bar{\theta}(t)\|_{H^7} \lesssim B + B^2 \ln(1+t)$. Let us then compute the equation satisfied by $\partial^{10}\theta$, where $\partial \in \{\partial_x, \partial_z\}$. We have

$$\partial_t \partial^{10}\theta + \boldsymbol{u} \cdot \nabla \partial^{10}\theta = \partial^{10}\partial_x \psi - [\partial^{10}, \boldsymbol{u} \cdot \nabla]\theta.$$

Multiplying by $\partial^{10}\theta$ and integrating by parts, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial^{10}\theta\|_{L^2} \leq 2\|\partial^{10}\partial_x\psi\|_{L^2} + 2\|[\partial^{10}, \boldsymbol{u} \cdot \nabla]\theta\|_{L^2}.$$

Using the bootstrap assumptions (4-3), we have

$$\|\partial^{10}\partial_x\psi\|_{L^2} \lesssim B(1+t)^{\frac{3}{8}} \quad \forall t \in [0, T].$$

As for the commutator term, using the tame estimates (2-16) together with the identity $\boldsymbol{u} = \nabla^\perp \psi$, we obtain, for any $k \geq 5$,

$$\begin{aligned} \|[\partial^k, \boldsymbol{u} \cdot \nabla]\theta\|_{L^2} &\lesssim \|\psi\|_{W^{2,\infty}}\|\theta\|_{H^k} + \|\partial_z\psi\|_{H^k}\|\partial_x\theta\|_\infty + \|\partial_x\psi\|_{H^k}\|\partial_z\theta\|_\infty \\ &\lesssim B(1+t)^{-\frac{11}{8}}\|\theta\|_{H^k} + \|\partial_x\theta\|_{H^{k-3}}\|\partial_x\theta\|_\infty + \|\partial_x^2\theta\|_{H^{k-4}}\|\partial_z\theta\|_\infty. \end{aligned} \tag{4-6}$$

In particular, using the estimates (4-5), we get, for $k = 10$,

$$\|[\partial^{10}, \boldsymbol{u} \cdot \nabla]\theta\|_{L^2} \lesssim B(1+t)^{-\frac{11}{8}}\|\theta\|_{H^{10}} + B^2(1+t)^{\frac{3}{8}}.$$

Assuming that $B < 1$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\theta\|_{H^{10}} \lesssim B(1+t)^{\frac{3}{8}} + B(1+t)^{-\frac{11}{8}}\|\theta\|_{H^{10}}.$$

The Gronwall lemma then ensures that

$$\|\theta(t)\|_{H^{10}} \lesssim \|\theta_0\|_{H^{10}} + B(1+t)^{\frac{11}{8}} \lesssim B(1+t)^{\frac{11}{8}}. \tag{4-7}$$

We then use the same strategy to estimate $\|\partial_x^2\Delta^4\theta\|_{L^2}$. The linear term in the right-hand side is now

$$\partial_x^3\Delta^4\psi = \partial_x^4\Delta^2\theta' = O(B) \quad \text{in } L^2.$$

The only difference in the treatment of the commutator term lies in the bound of terms of the form $\partial_z^2\psi\,\partial_x^3\partial_z^7\theta'$. For those, we use our first estimate on $\|\theta\|_{H^{10}}$ (4-7) together with the bootstrap assumptions (4-3) (see in particular (4-4), (4-5)), and we obtain

$$\|\partial_z^2\psi\,\partial_x^3\partial_z^7\theta'\|_{L^2} \lesssim \|\partial_z^2\psi\|_\infty\|\theta'\|_{H^{10}} \lesssim B^2(1+t)^{-\frac{11}{8}+\frac{11}{8}} \lesssim B^2.$$

It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_x^2\Delta^4\theta\|_{L^2} \lesssim B + B(1+t)^{-\frac{11}{8}}\|\partial_x^2\Delta^4\theta\|_{L^2},$$

and therefore $\|\partial_x^2\Delta^4\theta\|_{L^2} \lesssim B(1+t)$. The next step is to prove that $\sup_{t\in[0,T]}\|\partial_x^6\Delta^2\theta'\|_{L^2} \lesssim B$. To that end, we check that $\partial_x^6\Delta^2\theta'$ satisfies the assumptions of Lemma 2.4. The source term is

$S = \boldsymbol{u}\cdot\nabla\partial_x^6\Delta^2\theta' + [\partial_x^6\Delta^2, \boldsymbol{u}\cdot\nabla]\theta$. Classically, the first term is orthogonal to $\theta'$. It is therefore sufficient to bound the commutator. The terms involving $\bar\theta$ can be treated as perturbations of the dissipation term $\|\partial_x^7\Delta\theta'\|_{L^2}^2$, and therefore we focus on $[\partial_x^6\Delta^2, \boldsymbol{u}\cdot\nabla]\theta'$. First, note that

$$\|(\nabla^\perp\partial_x^6\Delta^2\psi)\cdot\nabla\theta'\|_{L^2} \le \|\nabla\theta'\|_\infty\|\nabla\partial_x^7\theta'\|_{L^2} \lesssim B(1+t)^{-\frac{3}{4}}\|\partial_x^7\Delta\theta'\|_{L^2}.$$

The other terms can be estimated thanks to the bootstrap assumptions together with the preliminary bounds on $\|\theta\|_{H^{10}}$ and $\|\partial_x^2\Delta^4\theta\|_{L^2}$. We obtain

$$\|[\partial_x^6\Delta^2, \boldsymbol{u}\cdot\nabla]\theta\|_{L^2} \lesssim B(1+t)^{-1-\delta}\|\partial_x^6\Delta^2\theta'\|_{L^2} + B^2(1+t)^{-1-\delta} + B(1+t)^{-\frac{1}{2}-\delta}\|\partial_x^7\Delta\theta'\|_{L^2}$$

for some $\delta > 0$. The details are left to the reader. Using a Cauchy–Schwarz inequality, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_x^6\Delta^2\theta'\|_{L^2}^2 + c\|\partial_x^7\Delta\theta'\|_{L^2}^2 \lesssim B^2(1+t)^{-1-\delta} + B(1+t)^{-1-\delta}\|\partial_x^6\Delta^2\theta'\|_{L^2}^2.$$

The Gronwall lemma then implies that $\sup_{t\in[0,T]}\|\partial_x^6\Delta^2\theta'\|_{L^2} \lesssim \|\theta_0\|_{H^{10}} + B^2 \lesssim B$.

We then follow the same strategy to obtain bounds on $\|\theta\|_{H^{12}}$, $\|\partial_x^4\Delta^4\theta\|_{L^2}$ and $\|\partial_x^8\Delta^2\theta\|_{L^2}$. We have

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\theta\|_{H^{12}} \lesssim \|\partial_x^2\theta'\|_{H^8} + \|[\partial^{12}, \boldsymbol{u}\cdot\nabla]\theta\|_{L^2}.$$

The first term in the right-hand side is bounded by $B(1+t)$. The commutator is estimated thanks to (4-6) together with the bootstrap assumptions and our preliminary bounds on derivatives up to order 10. We obtain $\|\theta(t)\|_{H^{12}} \lesssim B(1+t)^2$. We then write

$$\partial_t\partial_x^4\Delta^4\theta' + \boldsymbol{u}\cdot\nabla\partial_x^4\Delta^4\theta' = \partial_x^6\Delta^2\theta' - [\partial_x^4\Delta^4, \boldsymbol{u}\cdot\nabla]\theta.$$

The first term in the right-hand side is bounded by $CB$ in $L^2$. We then check that the nonlinear term can be treated perturbatively, using the bounds on $\theta'$ obtained so far, and we infer that $\|\partial_x^4\Delta^4\theta'(t)\|_{L^2} \lesssim B(1+t)$. Once again, we then use Lemma 2.4 in order to prove that $\|\partial_x^8\Delta^2\theta'(t)\|_{L^2} \lesssim B$ and that

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\theta(t)\|_{H^{14}} \lesssim B(1+t)^{-1-\delta}\|\theta(t)\|_{H^{14}} + B(1+t)^{\frac{3}{2}}.$$

The computations are very similar to the ones above, are left to the reader, and lead to the estimate of $\|\theta(t)\|_{H^{14}}$.

The last step is to prove additional decay on $\|\partial_x^{10}\theta'\|_{L^2}$ and $\|\partial_x^{11}\psi\|_{L^2}$. Setting $S = -\partial_x^{10}(\boldsymbol{u}\cdot\nabla\theta')$, we can decompose $S$ into $S = S_\parallel + S_\perp + S_\Delta$, with $S_\perp = \boldsymbol{u}\cdot\nabla\partial_x^{10}\theta'$, and

$$S_\Delta := -\sum_{k\le 6}\binom{10}{k}\partial_z\partial_x^{10-k}\psi\,\partial_x^{k+1}\theta + \sum_{k\le 5}\binom{10}{k}\partial_x^{11-k}\psi\,\partial_x^k\partial_z\theta,$$

$$S_\parallel := -\sum_{7\le k\le 9}\binom{10}{k}\partial_z\partial_x^{10-k}\psi\,\partial_x^{k+1}\theta + \sum_{6\le k\le 9}\binom{10}{k}\partial_x^{11-k}\psi\,\partial_x^k\partial_z\theta,$$

so that

$$\|S_\parallel\|_{L^2} \lesssim B^2(1+t)^{-2} + B(1+t)^{-1-\delta}\|\partial_x^{10}\theta\|_{L^2}, \quad \|S_\Delta\|_{L^2} \lesssim B(1+t)^{-\frac{1}{2}}\|\partial_x^{10}\Delta\psi\|_{L^2}.$$

Hence for $B$ sufficiently small, $S$ satisfies the assumptions of Proposition 2.6, and we obtain

$$\|\partial_x^{10}\theta'(t)\|_{L^2} \lesssim B(1+t)^{-1}.$$

Differentiating the equation on $\partial_x^9\theta$ with respect to time, we get

$$\partial_t\partial_t\partial_x^9\theta' = (1-G)\partial_t\partial_x^{10}\psi - \partial_t\partial_x^9(\boldsymbol{u}\cdot\nabla)\theta' - \partial_t G\partial_x^{10}\psi.$$

Estimating the norm of each term in the right-hand side and using Proposition 2.6, we obtain, for all $t \in [0, T]$,

$$\|\partial_t\partial_x^9\theta'(t)\|_{L^2} \lesssim \frac{1}{(1+t)^2}\Big(\sup_{s\in[0,T]}(1+s)\|\partial_s\partial_x^7\Delta^2\theta'(s)\|_{L^2} + B^2\Big).$$

Writing

$$\partial_t\partial_x^7\Delta^2\theta' = \partial_x^9\theta' - \partial_x^7\Delta^2(\boldsymbol{u}\cdot\nabla\theta),$$

we find that $\|\partial_t\partial_x^7\Delta^2\theta'\|_{L^2} \lesssim B(1+t)^{-1}$, and thus $\|\partial_t\partial_x^9\theta'\|_{L^2} \lesssim B(1+t)^{-2}$. Going back to the equation on $\partial_x^9\theta'$, we find that

$$\partial_x^{10}\psi = \partial_t\partial_x^9\theta' + \partial_x^9(\nabla^\perp\psi\cdot\nabla\theta) = O((1+t)^{-2}) \quad \text{in } L^2.$$

Finally, plugging these estimates into the equation on $\bar\theta$ leads to the desired bound on $\|\bar\theta\|_{H^9}$. □

Let us now prove a useful (albeit technical) result concerning the trace of $\partial_z^3\theta'$:

**Corollary 4.4.** *Under the assumptions of Lemma 4.3, for all $t \in [0, T]$,*

$$\|\partial_z^3\theta'|_{z=0}\|_{H^{33/4}(\mathbb{T})} \lesssim B(1+t)^{-\frac{1}{8}}.$$

*The same estimate holds for the trace at $z = 1$.*

*Proof.* Using Theorem 3.1 in Chapter 1 of [Lions and Magenes 1968],

$$\|\partial_z^3\theta'|_{z=0}\|_{H^s(\mathbb{T})} \lesssim \|\theta'\|^{\frac{1}{8}}_{H_x^\beta L_z^2}\|\partial_z^4\theta'\|^{\frac{7}{8}}_{H_x^\gamma L_z^2},$$

where $\frac{1}{8}\beta + \frac{7}{8}\gamma = s$. Taking $\beta = 10$ and $\gamma = 8$ and using the bounds of Lemma 4.3, we obtain the desired result. □

**4.3.** ***Decomposition of the traces of $\boldsymbol{\Delta^2\theta'}$ and $\boldsymbol{\partial_n\Delta^2\theta'}$.*** The role of the boundary layer is to lift the traces of $\Delta^2\theta'$ and $\partial_{\boldsymbol{n}}\Delta^2\theta'$ on the boundary. Therefore we first need to prove that these traces converge towards a (generically nontrivial) limit. In fact, we will even need to have a rather precise asymptotic expansion of the traces in powers of $(1+t)^{-1/4}$. This is the main purpose of this section.

The first result of this section concerns the long-time behavior of $\Delta^2\theta'|_{\partial\Omega}$ and $\partial_{\boldsymbol{n}}\Delta^2\theta'|_{\partial\Omega}$:

**Lemma 4.5** (long-time behavior of $\partial_z^k\Delta^2\theta'|_{\partial\Omega}$ and of $\partial_z^{2+k}G(t)|_{\partial\Omega}$, $k = 0, 1$). *For $k = 0, 1$, let*

$$\gamma^k_{\text{top}}(t, x) := \partial_z^k\Delta^2\theta'(t, x, z{=}1), \quad \gamma^k_{\text{bot}}(t, x) := \partial_z^k\Delta^2\theta'(t, x, z{=}0).$$

*Assume that $\theta_0 \in H^{14}(\Omega)$ and $\theta_0 = \partial_{\boldsymbol{n}}\theta_0 = 0$ on $\partial\Omega$, $\partial_z^2\bar\theta_0 = 0$ on $\partial\Omega$. Let $T > 0$, $B \in (0, 1)$ such that the bootstrap assumptions* (4-3) *hold on* $[0, T]$. *Assume furthermore that $\|\theta_0\|_{H^{14}} \leq B$. Then there exist* *universal constants $B_0$, $\delta > 0$ and functions $\gamma_{a,T}^0 \in H^9(\mathbb{T})$, $\gamma_{a,T}^1 \in H^8(\mathbb{T})$ such that if $B \le B_0$,*

$$\|\gamma_{a,T}^0\|_{H^9(\mathbb{T})} \lesssim \|\theta_0\|_{H^{14}} + B^2 \quad \text{and} \quad \|\gamma_a^0(t) - \gamma_{a,T}^0\|_{H^9(\mathbb{T})} \lesssim B^2 \frac{1}{(1+t)^\delta} \quad \forall t \in [0, T],$$
$$\|\gamma_{a,T}^1\|_{H^8(\mathbb{T})} \lesssim \|\theta_0\|_{H^{14}} + B^2 \quad \text{and} \quad \|\gamma_a^1(t) - \gamma_{a,T}^1\|_{H^8(\mathbb{T})} \lesssim B^2 \frac{1}{(1+t)^\delta} \quad \forall t \in [0, T].$$

*In a similar fashion, for $k = 2, 3$, $a \in \{\text{top}, \text{bot}\}$, there exists $g_{a,T}^k \in \mathbb{R}$ such that*

$$|g_{a,T}^k| \lesssim \|\theta_0\|_{H^{14}} + B^2 \quad \text{for } k \in \{2, 3\},$$
$$|g_{\text{bot},T}^2 - \partial_z^2 G(t, 0)| \lesssim \frac{B^2}{(1+t)^{\frac{3}{4}}} \quad \forall t \in [0, T),$$
$$|g_{\text{bot},T}^3 - \partial_z^3 G(t, 0)| \lesssim \frac{B^2}{(1+t)^{\frac{1}{2}}} \quad \forall t \in [0, T).$$

*The same estimates hold for $g_{\text{top},T}^k - \partial_z^k G(t, 1)$.*

The proof of Lemma 4.5 is postponed to the end of this section.

The second intermediate result of this section pushes further the decomposition of $\gamma_a^k(t)$. It holds under additional structural assumptions on $\theta^{\text{BL}}$ and $\theta^{\text{rem}} = \theta' - \theta^{\text{BL}}$. More precisely, let us assume that there exist profiles $\Theta_a^j$, $\Psi_a^j$ such that

$$\begin{aligned} \theta^{\text{BL}}(t, x, z) &= \sum_{j=0}^{4} (1+t)^{-1-\frac{j}{4}} \big(\Theta_{\text{bot}}^j(x, (1+t)^{\frac{1}{4}} z) + \Theta_{\text{top}}^j(x, (1+t)^{\frac{1}{4}}(1-z))\big), \\ \psi^{\text{BL}}(t, x, z) &= \sum_{j=0}^{4} (1+t)^{-2-\frac{j}{4}} \big(\Psi_{\text{bot}}^j(x, (1+t)^{\frac{1}{4}} z) + \Psi_{\text{top}}^j(x, (1+t)^{\frac{1}{4}}(1-z))\big), \end{aligned} \tag{4-8}$$

where there exists a universal constant $c > 0$ such that, for all $Z_0 \ge 0$,

$$\|\Theta_a^j\|_{H^9(\mathbb{T}\times(Z_0,+\infty))} + \|\Psi_a^j\|_{H^{11}(\mathbb{T}\times(Z_0,+\infty))} \lesssim (\|\theta_0\|_{H^{14}} + B^2) \exp(-c Z_0^{\frac{4}{5}}). \tag{4-9}$$

In the course of the proof, we shall also need the following assumption:

$$\partial_Z^2 \Theta_a^j|_{Z=0} \in H^7(\mathbb{T}), \quad \partial_Z^3 \Theta_a^j|_{Z=0} \in H^{\frac{15}{2}}(\mathbb{T}). \tag{4-10}$$

**Remark 4.6.** The profiles $\Theta_a^j$, $\Psi_a^j$ are not the same as the ones of Section 3. However we kept the same notation for convenience.

**Definition 4.7** (definition of $\gamma_a^{j,k}$). Let $\Theta_a^j$, $\Psi_a^j$ be the boundary layer profiles from (4-8), with $a \in \{\text{top}, \text{bot}\}$, $j \in \{0, \dots, 4\}$.

Let $\eta_{\text{bot}} = 1$ and $\eta_{\text{top}} = -1$. We then define the following coefficients:

$$\begin{aligned} \gamma_a^{0,2} &= 12 g_{a,T}^2 \partial_x \partial_Z^2 \Psi_a^0|_{Z=0}, \\ \gamma_a^{0,3} &= 8 g_{a,T}^2 \partial_x \partial_Z^2 \Psi_a^1|_{Z=0} - \tfrac{4}{3}\eta_a [\partial_Z^4 \{\Psi_a^0, \Theta_a^0\}_{x,Z}]'_{|Z=0}, \\ \gamma_a^{1,1} &= 40 \eta_a g_{a,T}^2 \partial_x \partial_Z^3 \Psi_a^0|_{Z=0}, \\ \gamma_a^{1,2} &= 20(\eta_a g_{a,T}^2 \partial_x \partial_Z^3 \Psi_a^1|_{Z=0} + g_{a,T}^3 \partial_x \partial_Z^2 \Psi_a^0|_{Z=0}) + 2[\partial_Z^5 \{\Psi_a^0, \Theta_a^0\}_{x,Z}]'_{|Z=0}, \end{aligned} \tag{4-11}$$

where $\{\cdot,\cdot\}_{x,Z}$ denotes the Poisson bracket

$$\{f, g\}_{x,Z} = \partial_x f\, \partial_Z G - \partial_Z f\, \partial_x g.$$

**Remark 4.8.** The bounds (4-9) ensure that $\|\gamma_a^{j,k}\|_{H^5(\mathbb{T})} \lesssim B^2$. We shall actually derive stronger regularity estimates in the course of the proof, as we construct explicitly the profiles $\Theta_a^j$ and $\Psi_a^j$.

**Lemma 4.9** (decomposition of $\gamma_a^k$). *Assume that $\theta_0 \in H^{14}(\Omega)$ and $\theta_0 = \partial_{\boldsymbol{n}}\theta_0 = 0$ on $\partial\Omega$, $\partial_z^2\bar{\theta}_0 = 0$ on $\partial\Omega$. Let $T > 0$, $B \in (0, 1)$ such that $\|\theta_0\|_{H^{14}(\Omega)} \le B$ and such that the bootstrap assumptions* (4-3) *hold on* $[0, T]$. *Assume furthermore that there exist profiles $\Theta_a^j$, $\Psi_a^j$ satisfying* (4-9) *and* (4-10) *such that $\theta^{\mathrm{rem}} = \theta' - \theta^{\mathrm{BL}}$ satisfies* (4-2), *where $\theta^{\mathrm{BL}}$ is defined by* (4-8). *Define the coefficients $\gamma_a^{j,k}$ by* (4-11).

*Then for $j = 0, 1$, $a \in \{\mathrm{top}, \mathrm{bot}\}$, there exists $\Gamma_{a,T}^j \in W^{1,\infty}((0, T); L^2(\mathbb{T}))$ such that for all $t \in [0, T]$,*

$$\gamma_a^0(t) = \gamma_{a,T}^0 + \gamma_a^{0,2}(1+t)^{-\frac{1}{2}} + \gamma_a^{0,3}(1+t)^{-\frac{3}{4}} + \Gamma_{a,T}^0(t) - \gamma_a^{0,2}(1+T)^{-\frac{1}{2}} - \gamma_a^{0,3}(1+T)^{-\frac{3}{4}},$$
$$\gamma_a^1(t) = \gamma_{a,T}^1 + \gamma_a^{1,1}(1+t)^{-\frac{1}{4}} + \gamma_a^{1,2}(1+t)^{-\frac{1}{2}} + \Gamma_{a,T}^1(t) - \gamma_a^{1,1}(1+T)^{-\frac{1}{4}} + \gamma_a^{1,2}(1+T)^{-\frac{1}{2}}.$$

*where, for all $t \in [0, T]$, for $j = 0, 1$, $\ell = 0, 1, 2$,*

$$\|\partial_t^\ell \Gamma_{a,T}^j(t)\|_{L^2(\mathbb{T})} \lesssim B^2(1+t)^{-1-\ell+\frac{j}{4}}, \quad \|\Gamma_{a,T}^j(t)\|_{H^4(\mathbb{T})} \lesssim B^2(1+t)^{-\frac{23}{24}+\frac{j}{4}}.$$

Let us now prove Lemmas 4.5 and 4.9.

*Proof of Lemma 4.5.* We have

$$\frac{\partial}{\partial t}\Delta^2\theta' = (1-G)\partial_x^2\theta' - 4\partial_z G \partial_z \partial_x^3\psi - 2\partial_z^2 G \partial_x^3\psi \\ - \sum_{k=1}^{4}\binom{4}{k}\partial_z^k G \partial_x \partial_z^{4-k}\psi + \Delta^2(\partial_z\psi\,\partial_x\theta')' - \Delta^2(\partial_x\psi\,\partial_z\theta')'. \quad (4\text{-}12)$$

We now take the trace of the above equation at $z = 0$, recalling that $G|_{z=0} = \partial_z G|_{z=0} = 0$ (see Lemma 2.1), $\psi|_{z=0} = \partial_z\psi|_{z=0} = 0$, and $\theta'|_{z=0} = \partial_z\theta'|_{z=0} = 0$. We obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\gamma_{\mathrm{bot}}^0 = -6\partial_z^2 G|_{z=0}\partial_x\partial_z^2\psi|_{z=0} + 6(\partial_z^3\psi|_{z=0}\partial_x\partial_z^2\theta'|_{z=0})' + 4(\partial_z^2\psi|_{z=0}\partial_x\partial_z^3\theta'|_{z=0})' \\ - 6(\partial_x\partial_z^2\psi|_{z=0}\partial_z^3\theta'|_{z=0})' - 4(\partial_x\partial_z^3\psi|_{z=0}\partial_z^2\theta'|_{z=0})'. \quad (4\text{-}13)$$

We then estimate each term in the right-hand side using Lemma 4.3. Note that $\partial_z^2 G|_{z=0}$ is bounded in $L^\infty(\mathbb{R}_+ \times (0, 1))$. We focus on the first term, which has the smallest decay. Using Theorem 3.1 in Chapter 1 of [Lions and Magenes 1968], we infer that, for any $s > 0$,

$$\|\partial_x\partial_z^2\psi|_{z=0}\|_{H^s} \le \|\partial_z^2\psi|_{z=0}\|_{H^{s+1}} \lesssim \|\psi\|_{H_x^\beta L_z^2}^{\frac{3}{8}}\|\partial_z^4\psi\|_{H_x^\gamma L_z^2}^{\frac{5}{8}},$$

where $\beta, \gamma$ are such that $\frac{3}{8}\beta + \frac{5}{8}\gamma = s+1$. Let us choose $\beta = \gamma = 10$, $s = 9$. According to Lemma 4.3, $\|\psi\|_{H_x^{10}L_z^2} \lesssim B(1+t)^{-2}$. As for the other term, using the short-hand notation from Section 2,

$$\|\partial_z^4\psi\|_{H_x^{10}L_z^2} \lesssim \|\theta\|_{H_x^{11}L_z^2} \lesssim \underbrace{\|\theta\|_{H_x^{10}L_z^2}^{\frac{1}{2}}\|\theta\|_{H_x^{12}L_z^2}^{\frac{1}{2}}}_{\frac{1}{2}\times 1 + \frac{1}{2}\times 0} \lesssim B(1+t)^{-\frac{1}{2}}.$$

Hence

$$\|\partial_x\partial_z^2\psi|_{z=0}\|_{H^9} \lesssim B(1+t)^{-\frac{17}{16}}.$$

The quadratic terms, involving traces of derivatives of $\psi$ and of $\theta'$, have a higher decay. Let us estimate for instance $\partial_z^3\psi\,\partial_x\partial_z^2\theta'$ at $z=0$. We have, for any $s>\frac{1}{2}$,

$$\|\partial_z^3\psi|_{z=0}\partial_x\partial_z^2\theta'|_{z=0}\|_{H^s(\mathbb{T})}\lesssim\|\partial_z^3\psi|_{z=0}\|_{L^\infty(\mathbb{T})}\|\partial_x\partial_z^2\theta'|_{z=0}\|_{H^s(\mathbb{T})}+\|\partial_z^3\psi|_{z=0}\|_{H^s(\mathbb{T})}\|\partial_x\partial_z^2\theta'|_{z=0}\|_{L^\infty(\mathbb{T})}.$$

Using once again Theorem 3.1 in Chapter 1 of [Lions and Magenes 1968], we find that

$$\|\partial_x\partial_z^2\theta'|_{z=0}\|_{L^\infty(\mathbb{T})}\lesssim\|\theta'\|_{H_x^4L_z^2}^{\frac{11}{16}}\|\partial_z^8\theta'\|_{L_x^2L_z^2}^{\frac{5}{16}}\lesssim B(1+t)^{-\frac{1}{2}},$$

and we recall that $\|\psi\|_{W^{3,\infty}}\lesssim B(1+t)^{-9/8}$ (see (4-4)). The same arguments together with Lemma 4.3 also imply

$$\|\partial_x\partial_z^2\theta'|_{z=0}\|_{H^{19/2}(\mathbb{T})}\lesssim\|\theta'\|_{H_x^{12}L_z^2}^{\frac{3}{8}}\|\partial_z^4\theta'\|_{H_x^8L_z^2}^{\frac{5}{8}}\lesssim\|\partial_x^8\Delta^2\theta'\|_{L^2}\lesssim B,$$

while

$$\|\partial_z^3\psi|_{z=0}\|_{H^s(\mathbb{T})}\lesssim\|\psi\|_{L_z^2H_x^\beta}^{\frac{1}{8}}\|\partial_z^4\psi\|_{L_z^2H_x^\gamma}^{\frac{7}{8}}\lesssim\|\psi\|_{L_z^2H_x^\beta}^{\frac{1}{8}}\|\theta'\|_{L_z^2H_x^{\gamma+1}}^{\frac{7}{8}},\tag{4-14}$$

with $\frac{1}{8}\beta+\frac{7}{8}\gamma=s$. Taking $\beta=10$ and $\gamma=9$, we obtain, for some $s>9$,

$$\|\partial_z^3\psi|_{z=0}\partial_x\partial_z^2\theta'|_{z=0}\|_{H^s(\mathbb{T})}\lesssim B^2(1+t)^{-\frac{9}{8}}.$$

The other terms are treated in a similar fashion. We infer that there exists $\delta>0$ such that

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\|\gamma_a^0(t)\|_{H^9(\mathbb{T})}\right|\lesssim\frac{B^2}{(1+t)^{1+\delta}}\quad\forall t\in(0,T).$$

This completes the proof of the estimate on $\gamma_a^0$.

The estimate for $\gamma_a^1$ follows from a similar argument. Taking the vertical derivative of (4-12), we have

$$\begin{aligned}\frac{\partial}{\partial t}\partial_z\Delta^2\theta'&=(1-G)\partial_z\partial_x^2\theta'\\&\quad-\sum_{k=1}^{3}\binom{3}{k}\partial_z^kG\partial_z^{3-k}\partial_x^3\psi-\sum_{k=1}^{5}\binom{5}{k}\partial_z^kG\partial_x\partial_z^{5-k}\psi+\partial_z\Delta^2(\partial_z\psi\partial_x\theta')'-\Delta^2(\partial_x\psi\partial_z\theta')'.\end{aligned}$$

Taking the trace of the above equation at $z=0$, we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\gamma_{\mathrm{bot}}^1&=-10\partial_z^2G|_{z=0}\partial_x\partial_z^3\psi|_{z=0}-10\partial_z^3G|_{z=0}\partial_x\partial_z^2\psi|_{z=0}\\&\quad+6\partial_x^2(\partial_z^2\psi|_{z=0}\partial_x\partial_z^2\theta'|_{z=0}-\partial_x\partial_z^2\psi|_{z=0}\partial_z^2\theta'|_{z=0})\\&\quad+10(\partial_z^3\psi|_{z=0}\partial_x\partial_z^3\theta'_{z=0}+\partial_z^4\psi_{z=0}\partial_x\partial_z^2\theta'|_{z=0})'+5(\partial_z^2\psi|_{z=0}\partial_x\partial_z^4\theta'_{z=0})'\\&\quad-10(\partial_x\partial_z^3\psi|_{z=0}\partial_z^3\theta'|_{z=0}+\partial_x\partial_z^4\psi_{z=0}\partial_z^2\theta'|_{z=0})'-5(\partial_x\partial_z^2\psi|_{z=0}\partial_z^4\theta'|_{z=0})'.\end{aligned}$$

The highest-order term is the first one. We recall that $\partial_z^2G$ and $\partial_z^3G$ are uniformly bounded by $CB$ in $L^\infty$, and that the trace of $\partial_z^3\psi$ in $H^9$ is evaluated thanks to (4-14). Once again, the quadratic terms have a higher decay and can be handled as perturbations. The trace of $\partial_z^4\theta'|_{z=0}$ can estimated thanks to $\gamma_{\mathrm{bot}}^0$. We then obtain, for some $\delta>0$,

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\|\gamma_a^1(t)\|_{H^8(\mathbb{T})}\right|\lesssim\frac{B^2}{(1+t)^{1+\delta}}\quad\forall t\in(0,T),$$

and the desired estimate for $\gamma_a^1$ follows.

Let us now address the convergence of $\partial_z^k G(t)|_{\partial\Omega}$ as $t \to \infty$. We recall that

$$\partial_t \partial_z^k G(t, z=0) = -\partial_z^{k+1} \overline{\boldsymbol{u} \cdot \nabla\theta'|_{z=0}} = \partial_z^{k+1} \overline{\partial_z \psi \partial_x \theta' - \partial_x \psi \partial_z \theta'|_{z=0}}.$$

Since $\psi|_{z=0} = \partial_{\boldsymbol{n}} \psi|_{z=0} = 0$ and $\theta'|_{z=0} = \partial_{\boldsymbol{n}}\theta'|_{z=0} = 0$, we have

$$\partial_t \partial_z^2 G(t, z=0) = \overline{3\partial_z^2 \psi|_{z=0} \partial_x \partial_z^2 \theta'|_{z=0} - \partial_x \partial_z^2 \psi|_{z=0} \partial_z^2 \theta'|_{z=0}}.$$

As above,

$$\begin{aligned}\|\partial_z^2 \psi|_{z=0} \partial_x \partial_z^2 \theta'|_{z=0}\|_{L^1(\mathbb{T})} &\le \|\partial_z^2 \psi|_{z=0}\|_{L^2(\mathbb{T})} \|\partial_x \partial_z^2 \theta'|_{z=0}\|_{L^2(\mathbb{T})} \\ &\lesssim \|\partial_z^2 \psi\|_{H^{1/2}(\Omega)} \|\partial_x \partial_z^2 \theta'\|_{H^{1/2}(\Omega)} \lesssim B^2 (1+t)^{-\frac{7}{4}}.\end{aligned}$$

The estimate on $\partial_t \partial_z^3 G(t, z=0)$ is similar and left to the reader. □

We now turn towards the decomposition of $\gamma_a^0$ and $\gamma_a^1$ for $a \in \{\text{top}, \text{bot}\}$:

*Proof of Lemma 4.9.* We focus on $a = \text{bot}$ by symmetry, and we start with the decomposition of $\gamma_{\text{bot}}^0$. Let us go back to (4-13). The main term in the right-hand side is $-6\partial_z^2 G|_{z=0} \partial_x \partial_z^2 \psi|_{z=0}$. Following Lemma 4.5 and using the decomposition $\theta' = \theta^{\text{BL}} + \theta^{\text{rem}}$, we write

$$\begin{aligned}\partial_z^2 G|_{z=0} \partial_x \partial_z^2 \psi|_{z=0} = (1+t)^{-\frac{3}{2}} g_{\text{bot},T}^2 \partial_x \partial_Z^2 \Psi_{\text{bot}|Z=0}^0 &+ (1+t)^{-\frac{7}{4}} g_{\text{bot},T}^2 \partial_x \partial_Z^2 \Psi_{\text{bot}|Z=0}^1 \\ &+ \sum_{j\ge 2} (1+t)^{-\frac{3}{2}-\frac{j}{4}} \partial_z^2 G|_{z=0} \partial_x \partial_Z^2 \Psi_{\text{bot}|Z=0}^j \\ &+ \sum_{j=0,1} (1+t)^{-\frac{3}{2}-\frac{j}{4}} (\partial_z^2 G|_{z=0} - g_{\text{bot},T}^2) \partial_x \partial_Z^2 \Psi_{\text{bot}|Z=0}^j \\ &+ \partial_z^2 G|_{z=0} \partial_x \partial_z^2 \psi^{\text{rem}}|_{z=0} + O(B^2 \exp(-c(1+t)^{\frac{1}{5}}),\end{aligned}$$

where the exponentially small term comes from the traces of derivatives of $\Psi_{\text{top}}^j$ evaluated at $Z = (1+t)^{1/4}$. The assumptions of the lemma and the bootstrap inequalities (4-2) ensure that, for all $t \in [0, T]$,

$$\Big\| \sum_{j\ge 2} (1+t)^{-\frac{3}{2}-\frac{j}{4}} \partial_z^2 G|_{z=0} \partial_x \partial_Z^2 \Psi_{\text{bot}|Z=0}^j \Big\|_{L^2(\mathbb{T})} \lesssim B^2 (1+t)^{-2},$$

$$\|\partial_z^2 G|_{z=0} \partial_x \partial_z^2 \psi^{\text{rem}}|_{z=0}\|_{L^2(\mathbb{T})} \lesssim \|\partial_z^2 G\|_\infty \|\partial_x \theta^{\text{rem}}\|_{L^2} \lesssim B^2 (1+t)^{-2}.$$

Furthermore, Lemma 4.5 ensures that for all $t \in (0, T)$

$$|\partial_z^2 G|_{z=0} - g_{\text{bot},T}^2| \lesssim B^2 (1+t)^{-\frac{3}{4}},$$

and therefore

$$\Big\| \sum_{j=0,1} (1+t)^{-\frac{3}{2}-\frac{j}{4}} (\partial_z^2 G|_{z=0} - g_{\text{bot},T}^2) \partial_x \partial_Z^2 \Psi_{\text{bot}|Z=0}^j \Big\|_{L^2(\mathbb{T})} \lesssim B^3 (1+t)^{-\frac{9}{4}}.$$

We now address the quadratic terms in (4-13), namely

$$\mathcal{B}(\psi, \theta') := 6\{\partial_z^2 \psi, \partial_z^2 \theta'\}'|_{z=0} + 4(\partial_z^2 \psi \partial_x \partial_z^3 \theta' - \partial_x \partial_z^3 \psi \partial_z^2 \theta')'|_{z=0}.$$

Decomposing $\psi$ and $\theta'$ into their boundary layer and their remainder part, we find that the main-order quadratic term is

$$\begin{aligned}(1+t)^{-\frac{7}{4}} \big[6\partial_Z^3 \Psi_{\text{bot}}^0 \partial_x \partial_Z^2 \Theta_{\text{bot}}^0 + 4\partial_Z^2 \Psi_{\text{bot}}^0 \partial_x \partial_Z^3 \Theta_{\text{bot}}^0 - 6\partial_x \partial_Z^2 \Psi_{\text{bot}}^0 \partial_Z^3 \Theta_{\text{bot}}^0 - 4\partial_x \partial_Z^3 \Psi_{\text{bot}}^0 \partial_Z^2 \Theta_{\text{bot}}^0\big]'\big|_{z=0} \\ =: (1+t)^{-\frac{7}{4}} \gamma_{\text{bot,NL}}^0,\end{aligned}$$

while all the other terms are bounded in $L^2(\mathbb{T})$ by $CB^2(1+t)^{-2}$.

Defining $\gamma_a^{0,j}$ by (4-11), we find that

$$\partial_t(\gamma_{\rm bot}^{0,2}(1+t)^{-\frac{1}{2}}) = -6(1+t)^{-\frac{3}{2}} g^2_{{\rm bot},T}\partial_x\partial_Z^2\Psi^0_{{\rm bot}|Z=0},$$
$$\partial_t(\gamma_{\rm bot}^{0,3}(1+t)^{-\frac{3}{4}}) = -6(1+t)^{-\frac{7}{4}} g^2_{{\rm bot},T}\partial_x\partial_Z^2\Psi^1_{{\rm bot}|Z=0} + (1+t)^{-\frac{7}{4}}\gamma^0_{\rm bot,NL},$$

where we recognize the main terms in $-6\partial_z^2 G|_{z=0}\partial_x\partial_z^2\psi|_{z=0}$. Now, define $\Gamma^0_{{\rm bot},T}$ by

$$\begin{aligned}\Gamma^0_{{\rm bot},T}(t) &= 6\int_t^T \sum_{j\ge 2}(1+s)^{-\frac{3}{2}-\frac{j}{4}}\partial_z^2 G(s)|_{z=0}\partial_x\partial_Z^2\Psi^j_{{\rm bot}|Z=0}\,{\rm d}s\\
&\quad + 6\int_t^T \sum_{j=0,1}(1+s)^{-\frac{3}{2}-\frac{j}{4}}(\partial_z^2 G(s)|_{z=0} - g^2_{{\rm bot},T})\partial_x\partial_Z^2\Psi^j_{{\rm bot}|Z=0}\,{\rm d}s\\
&\quad + \int_t^T \mathcal{B}\Big(\sum_{j=1}^4 (1+s)^{-2-\frac{j}{4}}\Psi^j_{\rm bot}(x,Z_{\rm bot}) + \psi^{\rm rem}, \theta'(s)\Big)\,{\rm d}s\\
&\quad + \int_t^T \mathcal{B}\Big((1+s)^{-2}\Psi^0_{\rm bot}(x,Z_{\rm bot}), \sum_{j=1}^4 (1+s)^{-1-\frac{j}{4}}\Theta^j_{\rm bot}(x,Z_{\rm bot}) + \theta^{\rm rem}\Big)\,{\rm d}s\\
&\quad + O(B^2\exp(-c(1+t)^{\frac{1}{4}})).\end{aligned}$$

The last — exponentially small — term comes once again from the trace of $\Psi^j_{\rm top}$ at the lower boundary, i.e., at $Z=(1+t)^{1/4}$. We do not write its full expression for the sake of readability.

Note that the assumptions (4-2) on $\theta^{\rm rem}$ ensure that

$$\|\partial_z^3\theta^{\rm rem}|_{z=0}\|_{L^2} \lesssim \|\theta^{\rm rem}\|_{L^2}^{\frac{9}{16}}\|\partial_z^8\theta^{\rm rem}\|_{L^2}^{\frac{7}{16}} \lesssim B(1+t)^{-\frac{9}{8}}.$$

Recalling Corollary 4.4 and using the assumption $\partial_Z^3\Theta_a^j|_{Z=0}\in H^{15/2}(\mathbb{T})$ (see (4-10)), we also infer that

$$\|\partial_z^3\theta^{\rm rem}|_{z=0}\|_{H^{15/2}} \lesssim B(1+t)^{-\frac{1}{8}}.$$

Interpolating between these two estimates, we find in particular that

$$\|\partial_z^3\theta^{\rm rem}|_{z=0}\|_{H^5} \lesssim B(1+t)^{-\frac{11}{24}}.$$

The above estimates ensure that for $k=0,1$

$$\|\partial_t^k\Gamma^0_{{\rm bot},T}(t)\|_{L^2(\mathbb{T})} \lesssim B^2(1+t)^{-k-1},$$
$$\|\Gamma^0_{{\rm bot},T}(t)\|_{H^4(\mathbb{T})} \lesssim B^2(1+t)^{-\frac{23}{24}}.$$

Therefore we obtain the decomposition announced in the lemma for $\gamma_a^0$.

The decomposition of $\gamma_a^1$ follows from similar arguments and is left to the reader. □

**4.4.** ***Iterative construction of the boundary layer profile.*** Let us now turn towards the construction of the boundary layer profile, and more generally, of an approximate solution. The purpose of this subsection is to prove the two following lemmas. Our first result, which is truly the core of the construction, is valid under the bootstrap assumption (4-3) on $\theta'$:

**Lemma 4.10.** *Let $\theta_0\in H^{14}(\Omega)$ such that $\|\theta_0\|_{H^{14}}\le B<1$, and $\theta_0|_{\partial\Omega}=\partial_{\boldsymbol{n}}\theta_0|_{\partial\Omega}=0$, $\partial_z^2\bar\theta_0|_{\partial\Omega}=0$.*

*Let $\theta=\bar\theta+\theta'$ be a solution of* (1-7), *and assume that the bounds* (4-3) *hold on* $(0,T)$. *Let $\gamma^0_{a,T}$, $\gamma^1_{a,T}$ be defined by Lemma 4.5.*

*Then there exist profiles* $\Theta_a^j \in H^8(\mathbb{T}\times\mathbb{R}_+)$, $\Psi_a^j \in H^9(\mathbb{T}\times\mathbb{R}_+)$, $j\in\{0,\dots,4\}$ *and a corrector* $\theta_c \in H^9(\Omega)$, *depending only on* $\gamma^0_{a,T}, \gamma^1_{a,T}, g^2_{a,T}$ *and* $g^3_{a,T}$, *such that, defining* $\theta^{\mathrm{BL}}$ *by* (4-8), $\gamma^j_{a,T}$ *by Lemma 4.5 and* $\gamma_a^{j,k}$ *by* (4-11), *the following properties hold*:

(1) *Bounds on the profiles*: $\Theta_a^j$, $\Psi_a^j$ *satisfy* (4-9).

(2) *Bound on the corrector*: *setting* $\psi_c = \Delta^{-2}\partial_x\theta_c$,

$$\sup_{t\in[0,T]}\big(\|\partial_x\Delta^4\theta_c(t)\|_{L^2}+(1+t)^2\|\partial_x^5\theta_c(t)\|_{L^2}\big)\lesssim\|\theta_0\|_{H^{14}}+B^2,$$
$$\sup_{t\in[0,T]}\big((1+t)^3(\|\partial_t\partial_x^4\theta_c(t)\|+\|\partial_x^5\psi_c(t)\|_{L^2})\big)\lesssim\|\theta_0\|_{H^{14}}+B^2.$$

(3) *Traces at the top and bottom*: *at* $z=0$,

$$\Delta^2(\theta^{\mathrm{BL}}+\theta_c)|_{z=0}=\gamma^0_{\mathrm{bot},T}+\gamma^{0,2}_{\mathrm{bot}}(1+t)^{-\frac12}+\gamma^{0,3}_{\mathrm{bot}}(1+t)^{-\frac34}-\gamma^{0,2}_{\mathrm{bot}}(1+T)^{-\frac12}-\gamma^{0,3}_{\mathrm{bot}}(1+T)^{-\frac34}, \tag{4-15}$$
$$\partial_z\Delta^2(\theta^{\mathrm{BL}}+\theta_c)|_{z=0}=\gamma^1_{\mathrm{bot},T}+\gamma^{1,1}_{\mathrm{bot},}(1+t)^{-\frac14}+\gamma^{1,2}_{\mathrm{bot}}(1+t)^{-\frac12}-\gamma^{1,1}_{\mathrm{bot}}(1+T)^{-\frac14}-\gamma^{1,2}_{\mathrm{bot}}(1+T)^{-\frac12}. \tag{4-16}$$

*Similar formulas hold at* $z=1$.

(4) *Evolution equation*: $\theta^{\mathrm{BL}}+\theta_c$ *satisfies*

$$\partial_t(\theta^{\mathrm{BL}}+\theta_c)=(1-G)\partial_x^2\Delta^{-2}(\theta^{\mathrm{BL}}+\theta_c)-(\nabla^\perp\Delta^{-2}\partial_x(\theta^{\mathrm{BL}}+\theta_c)\cdot\nabla(\theta^{\mathrm{BL}}+\theta_c))'+R^{\mathrm{BL}},$$

*and the remainder* $R^{\mathrm{BL}}$ *is such that*, *for* $\ell=0,1$, *for all* $t\in[0,T]$,

$$\|\partial_t^\ell\partial_x^4R^{\mathrm{BL}}\|_{L^2}\lesssim B^2(1+t)^{-3-\ell},\quad \|\partial_x^2\Delta^2R^{\mathrm{BL}}\|\lesssim B^2(1+t)^{-2},\quad \|\Delta^4R^{\mathrm{BL}}\|_{L^2}\lesssim B^2(1+t)^{-\frac98}.$$

**Remark 4.11.** The reader may compare formulas (4-15), (4-16) with the ones from Lemma 4.9. The terms $\Gamma^j_{a,T}$ are lifted neither by the boundary layer term $\theta^{\mathrm{BL}}$ nor by the corrector $\theta_c$, and an additional corrector will be built to handle them; see Lemma 4.13 below.

**Remark 4.12.** Actually, all profiles $\Theta_a^j$, $\Psi_a^j$, and therefore $\theta^{\mathrm{BL}}$, $\psi^{\mathrm{BL}}$, depend on $T$ through $\gamma^0_{a,T}, \gamma^1_{a,T}$. However, in order not to burden unnecessarily the notation, we will omit this dependency in the present section. The dependency will be restored in Section 4.5 when we perform the final bootstrap argument.

Once the boundary layer part is constructed, under an additional bootstrap assumption on the remainder, we can define a nonlinear corrector:

**Lemma 4.13.** *Let* $\theta_0\in H^{14}(\Omega)$ *such that* $\|\theta_0\|_{H^{14}}\le B<1$, *and* $\theta_0|_{\partial\Omega}=\partial_{\boldsymbol n}\theta_0|_{\partial\Omega}=0$, $\partial_z^2\bar\theta_0|_{\partial\Omega}=0$.

*Let* $\theta=\bar\theta+\theta'$ *be a solution of* (1-7), *and assume that the bounds* (4-3) *hold on* $(0,T)$.

*Let* $\theta^{\mathrm{BL}}$, $\psi^{\mathrm{BL}}$ *be given by Lemma 4.10, and let* $\theta^{\mathrm{rem}}=\theta'-\theta^{\mathrm{BL}}$. *Assume that* (4-2) *holds on* $(0,T)$, *and define* $\Gamma^j_{a,T}$ *as in Lemma 4.9.*

*Then there exists* $\sigma^{\mathrm{NL}}_{\mathrm{lift}}\in H^8(\Omega)$ *such that*

$$\Delta^2\sigma^{\mathrm{NL}}_{\mathrm{lift}}|_{z=0}=\Gamma^0_{\mathrm{bot},T},\quad \partial_z\Delta^2\sigma^{\mathrm{NL}}_{\mathrm{lift}}|_{z=0}=\Gamma^1_{\mathrm{bot},T},$$
$$\Delta^2\sigma^{\mathrm{NL}}_{\mathrm{lift}}|_{z=1}=\Gamma^0_{\mathrm{top},T},\quad \partial_z\Delta^2\sigma^{\mathrm{NL}}_{\mathrm{lift}}|_{z=1}=\Gamma^1_{\mathrm{top},T}$$

*and, for all* $t \in [0, T]$,

$$\|\Delta^4 \sigma_{\text{lift}}^{\text{NL}}\|_{L^2} \lesssim B^2(1+t)^{-\frac{1}{12}}, \quad \|\partial_x^2 \Delta^2 \sigma_{\text{lift}}^{\text{NL}}\|_{L^2} \lesssim B^2(1+t)^{-1+\frac{1}{64}}, \quad \|\partial_x^4 \sigma_{\text{lift}}^{\text{NL}}\|_{L^2} \lesssim B^2(1+t)^{-2}.$$

*As a consequence, setting* $\theta^{\text{app}} := \theta^{\text{BL}} + \theta_c + \sigma_{\text{lift}}^{\text{NL}}$, *we have*

$$\Delta^2 \theta^{\text{app}} = \Delta^2 \theta', \quad \partial_{\boldsymbol{n}} \Delta^2 \theta^{\text{app}} = \partial_{\boldsymbol{n}} \Delta^2 \theta' \quad \text{on } \partial\Omega.$$

*Furthermore* $\theta^{\text{app}}$ *is a solution of*

$$\partial_t \theta^{\text{app}} = (1-G)\partial_x^2 \Delta^{-2} \theta^{\text{app}} - (\nabla^\perp \Delta^{-2} \partial_x \theta^{\text{app}} \cdot \nabla \theta^{\text{app}})' + S_{\text{rem}},$$

*and the remainder* $S_{\text{rem}}$ *is such that, for all* $k, m \geq 0$ *with* $k + m \leq 8$,

$$\|\partial_t^\ell \partial_x^4 S_{\text{rem}}\|_{L^2} \lesssim B^2(1+t)^{-3-\ell}, \quad \|\partial_x^2 \Delta^2 S_{\text{rem}}\| \lesssim B^2(1+t)^{-2}, \quad \|\Delta^4 S_{\text{rem}}\|_{L^2} \lesssim B^2(1+t)^{-\frac{9}{8}}.$$

The main part of this section will be devoted to the proof of Lemma 4.10. The strategy will be very similar to the one of Section 3, and we will often refer the reader to the computations therein. We begin with the construction of the profiles $\Theta_a^j$, $\Psi_a^j$[1]. To that end, we plug the ansatz (4-8) into (1-7) and identify the powers of $1+t$ in the vicinity of $z=0$ or $z=1$. Note that for $z \ll 1$ and $t \in [0, T]$, setting $Z = (1+t)^{1/4} z$ and using Lemma 4.5,

$$\begin{aligned} G(t,z) &= \tfrac{1}{2}\partial_z^2 G(t,0) z^2 + \tfrac{1}{6}\partial_z^3 G(t,0) z^3 + O(z^4) \\ &= \tfrac{1}{2}(1+t)^{-\frac{1}{2}} g_{\text{bot},T}^2 Z^2 + \tfrac{1}{6}(1+t)^{-\frac{3}{4}} g_{\text{bot},T}^3 Z^3 + O((1+t)^{-1} Z^4 + (1+t)^{-\frac{5}{4}}(Z^2+Z^3)). \end{aligned} \tag{4-17}$$

A similar expansion holds in the vicinity of $z=1$. Furthermore, in the vicinity of $z=0$, setting $S = -(\nabla^\perp \psi \cdot \nabla \theta')'$ and assuming that (4-2) holds,

$$S = \sum_{0 \leq i,j \leq 4} (1+t)^{-3-\frac{i+j-1}{4}} (\partial_Z \Psi_{\text{bot}}^i \partial_x \Theta_{\text{bot}}^j - \partial_x \Psi_{\text{bot}}^i \partial_Z \Theta_{\text{bot}}^j)' + O((1+t)^{-\frac{15}{4}}) \quad \text{in } L^2.$$

Following the computations of the previous section and identifying the coefficient of $(1+t)^{-2-j/4}$, we obtain, for $j \in \{0, \dots, 3\}$ (compare with (3-8)),

$$-\big(1 + \tfrac{1}{4}j\big)\Theta_a^j + \tfrac{1}{4} Z \partial_Z \Theta_a^j = \partial_x \Psi_a^j + S_a^j, \tag{4-18}$$

where the source terms $S_a^j$ are defined by

$$\begin{aligned} S_a^0 &= S_a^1 = 0, \\ S_a^2 &= -\tfrac{1}{2} g_{a,T}^2 Z^2 \partial_x \Psi_a^0, \\ S_a^3 &= -\tfrac{1}{2} g_{a,T}^2 Z^2 \partial_x \Psi_a^1 - \eta_a \tfrac{1}{6} g_{a,T}^3 Z^3 \partial_x \Psi_a^0 + \eta_a (\partial_Z \Psi_a^0 \partial_x \Theta_a^0 - \partial_x \Psi_a^0 \partial_Z \Theta_a^0)', \end{aligned} \tag{4-19}$$

with $\eta_{\text{bot}} = 1$, $\eta_{\text{top}} = -1$.

Let us now proceed to define recursively the profiles $\Theta_a^j$, $\Psi_a^j$.

[1]We recall that these profiles are different from the ones constructed in Section 3, in spite of a similar notation.

***Main order boundary layer terms:* $\Theta_a^0$ *and* $\Theta_a^1$.** The role of the boundary layer profiles $\Theta_a^j$ for $j=0,1$ is to correct the traces of $\Delta^2\theta'$ and $\partial_z\Delta^2\theta'$ on $\partial\Omega$ *at main order*, i.e., $\gamma_{a,T}^j$ (see Lemma 4.9). Choosing $\Psi_a^j$ such that $\partial_Z^4\Psi_a^j=\partial_x\Theta_a^j$ for $j=0,1$ and recalling (4-18), we are led to

$$\begin{cases} Z\partial_Z^5\Theta_a^0=4\partial_x^2\Theta_a^0 \quad \text{in } \mathbb{T}\times(0,+\infty),\\ \partial_Z^4\Theta^0_{a|Z=0}=\gamma_{a,T}^0, \quad \partial_Z^5\Theta^0_{a|Z=0}=0,\\ \Theta^0_{a|Z=0}=0, \quad \partial_Z\Theta^0_{a|Z=0}=0, \quad \lim_{Z\to\infty}\Theta_a^0=0\end{cases}$$

and

$$\begin{cases} Z\partial_Z^6\Theta_a^1=4\partial_x^2\partial_Z\Theta_a^1 \quad \text{in } \mathbb{T}\times(0,+\infty),\\ \partial_Z^4\Theta^1_{a|Z=0}=0, \quad \partial_Z^5\Theta^1_{a|Z=0}=\eta_a\gamma_{a,T}^1,\\ \Theta^1_{a|Z=0}=0, \quad \partial_Z\Theta^1_{a|Z=0}=0, \quad \lim_{Z\to\infty}\Theta_a^1=0.\end{cases}$$

Note that these systems are identical to (3-2) and (3-3) respectively. As a consequence, as in the previous section (see (3-5)), we find that

$$\begin{aligned}\Theta_a^0(x,Z)&=\sum_{k\in\mathbb{Z}\setminus\{0\}}|k|^{-2}\hat{\gamma}_{a,T}^0(k)\chi_0(|k|^{\frac12}Z)e^{ikx},\\ \Psi_a^0(x,Z)&=\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{ik|k|^2}\hat{\gamma}_{a,T}^0(k)\big[\tfrac14|k|^{\frac12}Z\chi_0'(|k|^{\frac12}Z)-\chi_0(|k|^{\frac12}Z)\big]e^{ikx},\end{aligned} \tag{4-20}$$

where $\chi_0$ is defined in Corollary 3.3. Since $\|\gamma_{a,T}^0\|_{H^9}\lesssim\|\theta_0\|_{H^{14}}+B^2$ according to Lemma 4.5, it follows that

$$\begin{aligned}\|\Theta_a^0\|_{H_x^{11}L_Z^2}+\|\Theta_a^0\|_{L_x^2H_Z^{22}}&\lesssim\|\theta_0\|_{H^{14}}+B^2,\\ \|\Psi_a^0\|_{H_x^{12}L_Z^2}+\|\Psi_a^0\|_{L_x^2H_Z^{24}}&\lesssim\|\theta_0\|_{H^{14}}+B^2.\end{aligned}$$

In a similar fashion, recalling the definition of $\chi_1$ from Corollary 3.3 (see also (3-6)),

$$\begin{aligned}\Theta_a^1(x,Z)&=\eta_a\sum_{k\in\mathbb{Z}\setminus\{0\}}|k|^{-\frac52}\hat{\gamma}_{a,T}^1(k)\chi_1(|k|^{\frac12}Z)e^{ikx},\\ \Psi_a^1(x,Z)&=\eta_a\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{1}{ik|k|^{\frac52}}\hat{\gamma}_{a,T}^1(k)\big[\tfrac14|k|^{\frac12}Z\chi_1'(|k|^{\frac12}Z)-\tfrac54\chi_1(|k|^{\frac12}Z)\big]e^{ikx}.\end{aligned} \tag{4-21}$$

Since $\|\gamma_{a,T}^1\|_{H^8}\lesssim\|\theta_0\|_{H^{14}}+B^2$, we also have

$$\begin{aligned}\|\Theta_a^1\|_{H_x^{21/2}L_Z^2}+\|\Theta_a^1\|_{L_x^2H_Z^{21}}&\lesssim\|\theta_0\|_{H^{14}}+B^2,\\ \|\Psi_a^1\|_{H_x^{23/2}L_Z^2}+\|\Psi_a^1\|_{L_x^2H_Z^{23}}&\lesssim\|\theta_0\|_{H^{14}}+B^2.\end{aligned}$$

Let us now define the boundary terms $\gamma_a^{0,j}$ and $\gamma_a^{1,j}$ by (4-11). It follows from the above expressions for $\Psi_a^j$ and $\Theta_a^j$ and from the boundary conditions $\chi_0(0)=\chi_0'(0)=0$ that

$$\begin{aligned}\|\gamma_a^{0,2}\|_{H^{10}(\mathbb{T})}\lesssim B^2, \quad \|\gamma_a^{0,3}\|_{H^{17/2}(\mathbb{T})}\lesssim B^2,\\ \|\gamma_a^{1,1}\|_{H^{19/2}(T)}\lesssim B^2, \quad \|\gamma_a^{1,2}\|_{H^{8}(\mathbb{T})}\lesssim B^2.\end{aligned} \tag{4-22}$$

In a similar fashion, defining the source terms $S_a^2$, $S_a^3$ by (4-19), we have

$$\begin{aligned} \|S_a^2\|_{H_x^{12}L_Z^2} + \|S_a^2\|_{L_x^2 H_Z^{24}} &\lesssim B^2, \\ \|S_a^3\|_{H_x^{10}L_Z^2} + \|S_a^3\|_{L_x^2 H_Z^{20}} &\lesssim B^2. \end{aligned} \tag{4-23}$$

Note however that because of the quadratic term $\{\Psi_a^0, \Theta_a^0\}_{x,Z}$, $S_a^3$ does not have the same self-similar structure as $\Psi_a^j$, $\Theta_a^j$ for $j = 0, 1$, which is also shared by $S_a^2$.

***Correctors $\Theta_{c,a}^0$ and $\Theta_{c,a}^1$.*** We recall that the coefficients $\gamma_a^{j,k}$ are defined by (4-11), and are estimated in (4-22) above. The terms $\gamma_a^{0,2}(1+T)^{-1/2}$ and $\gamma_a^{0,3}(1+T)^{-3/4}$ in Lemma 4.9 are constant in time, but smaller (for $T \gg 1$) than $\gamma_{a,T}^0$. Hence they give rise to a profile $\Theta_{c,a}^0$ whose construction is very similar to the one of $\Theta_a^0$, but whose size is much smaller. More precisely, we set

$$\begin{aligned} \Theta_{c,a}^0(x, Z) &= \sum_{k\in\mathbb{Z}\setminus\{0\}} |k|^{-2}\big[\hat{\gamma}_a^{0,2}(k)(1+T)^{-\frac12} + \hat{\gamma}_a^{0,3}(k)(1+T)^{-\frac34}\big]\chi_0(|k|^{\frac12}Z)e^{ikx}, \\ \Theta_{c,a}^1(x, Z) &= \eta_a \sum_{k\in\mathbb{Z}\setminus\{0\}} |k|^{-\frac52}\big[\hat{\gamma}_a^{1,1}(k)(1+T)^{-\frac14} + \hat{\gamma}_a^{1,2}(1+T)^{-\frac12}\big]\chi_1(|k|^{\frac12}Z)e^{ikx}. \end{aligned}$$

Remembering (4-22), we have, for $j = 0, 1$,

$$\begin{aligned} \|\Theta_{c,a}^j\|_{H_x^{21/2}L_Z^2} + \|\Theta_{c,a}^j\|_{L_x^2 H_Z^{21}} &\lesssim B^2(1+T)^{-\frac12+\frac{j}{4}}, \\ \|\partial_Z^2\Theta_{c,a}^j|_{Z=0}\|_{H^{19/2}(\mathbb{T})} &\lesssim B^2(1+T)^{-\frac12+\frac{j}{4}}. \end{aligned}$$

Analogously to $\Psi_a^0$ and $\Psi_a^1$, we also define

$$\begin{aligned} \Psi_{c,a}^0 &= \sum_{k\in\mathbb{Z}^*} \frac{1}{ik|k|^2}\big[\hat{\gamma}_a^{0,2}(k)(1+T)^{-\frac12} + \hat{\gamma}_a^{0,3}(k)(1+T)^{-\frac34}\big]\big(\tfrac14\xi\chi_0'(\xi) - \chi_0(\xi)\big)|_{\xi=|k|^{1/2}Z}e^{ikx}, \\ \Psi_{c,a}^1 &= \eta_a \sum_{k\in\mathbb{Z}^*} \frac{1}{ik|k|^{\frac52}}\big[\hat{\gamma}_a^{1,1}(k)(1+T)^{-\frac14} + \hat{\gamma}_a^{1,2}(1+T)^{-\frac12}\big]\big(\tfrac14\xi\chi_1'(\xi) - \chi_1(\xi)\big)|_{\xi=|k|^{1/2}Z}e^{ikx}, \end{aligned}$$

so that $\partial_Z^4\Psi_{c,a}^j = \partial_x\Theta_{c,a}^j$, and we have

$$\|\Psi_{c,a}^j\|_{H_x^{23/2}L_Z^2} + \|\Psi_{c,a}^j\|_{L_x^2 H_Z^{23}} \lesssim B^2(1+T)^{-\frac12+\frac{j}{4}}.$$

***Lower-order boundary layer terms: $\Theta_a^2$, $\Theta_a^3$ and $\Theta_{c,a}^2$.*** We recall that $\Theta_a^j$, $\Psi_a^j$ must satisfy (4-18), where the source term $S_a^j$ is given by (4-19). Note that since $\Psi_a^0$, $\Psi_a^1$ and $\Theta_a^0$ have been constructed in the previous step, the source terms $S_a^2$ and $S_a^3$ are defined unequivocally and have exponential decay. Moreover, following Lemma 4.9 and noting that

$$\Delta^2\theta^{\mathrm{BL}}|_{z=0} = \sum_{j=0}^{3}(1+t)^{-\frac{j}{4}}\partial_Z^4\Theta_{\mathrm{bot}|Z=0}^j + 2\sum_{j=0}^{3}(1+t)^{-\frac12-\frac{j}{4}}\partial_x^2\partial_Z^2\Theta_{\mathrm{bot}|Z=0}^j + O((1+t)^{-1}),$$

we enforce the following boundary conditions:

$$\begin{aligned} \partial_Z^4\Theta_{a|Z=0}^2 &= \gamma_a^{0,2} - 2\partial_x^2\partial_Z^2\Theta_{a|Z=0}^0, \quad & \partial_Z^5\Theta_{a|Z=0}^2 &= \eta_a\gamma_a^{1,1} - 2\partial_x^2\partial_Z^3\Theta_{a|Z=0}^0, \\ \partial_Z^4\Theta_{a|Z=0}^3 &= \gamma_a^{0,3} - 2\partial_x^2\partial_Z^2\Theta_{a|Z=0}^1, \quad & \partial_Z^5\Theta_{a|Z=0}^3 &= \eta_a\gamma_a^{1,2} - 2\partial_x^2\partial_Z^3\Theta_{a|Z=0}^1, \end{aligned} \tag{4-24}$$

where the coefficients $\gamma_a^{j,k}$ are defined in (4-11) and estimated in (4-22). There remains to specify the relationship between $\Psi_a^j$ and $\Theta_a^j$. In order that $\Delta^2\psi^{\mathrm{BL}} = \partial_x\theta^{\mathrm{BL}}$ at main order, following the computations of the previous section (see in particular (3-9)), we take, for $j = 2, 3$,

$$\partial_Z^4\Psi_a^j + 2\partial_x^2\partial_Z^2\Psi_a^{j-2} = \partial_x\Theta_a^j.$$

Eliminating $\Psi_a^2$ from the equation on $\Theta_a^2$, we find that the system satisfied by $\Theta_a^2$ is

$$\begin{cases} Z\partial_Z^5\Theta_a^2 - 2\partial_Z^4\Theta_a^2 = 4\partial_x^2\Theta_a^2 - 2g_{\mathrm{bot},T}^2\partial_Z^4(Z^2\partial_x\Psi_a^0) - 8\partial_x^3\partial_Z^2\Psi_a^0, \\ \Theta^2_{a|Z=0} = \partial_Z\Theta^2_{a|Z=0} = 0, \\ \Theta_a^2(Z) \to 0 \quad \text{as } Z \to \infty, \end{cases} \tag{4-25}$$

together with (4-24). Note that $2\partial_x\partial_Z^2\Psi^0_{a|Z=0} = -\partial_Z^2\Theta^0_{a|Z=0}$ and $4\partial_x\partial_Z^3\Psi^0_{a|Z=0} = \partial_Z^3\Theta^0_{a|Z=0}$, so that the boundary conditions are (once again) redundant. In other words, taking the trace of (4-25) at $Z = 0$, we find $\partial_Z^4\Theta^2_{a|Z=0} = \gamma_a^{0,2} - 2\partial_x^2\partial_Z^2\Theta^0_{a|Z=0}$. Differentiating twice more with respect to $Z$, we find that the Fourier transform of $\partial_Z^2\Theta_a^2$, after a suitable lifting, satisfies an equation of the form (3-4) with boundary conditions of the type (iii) from Lemma 3.2. Using the explicit Fourier representation of $\Psi_a^0$ and $\Theta_a^0$ (4-20), we find that

$$\|\Theta_a^2\|_{H_x^{10}L_Z^2} + \|\Theta_a^2\|_{L_x^2H_Z^{20}} \lesssim \|\theta_0\|_{H^{14}} + B^2, \quad \|\Psi_a^2\|_{H_x^{11}L_Z^2} + \|\Psi_a^2\|_{L_x^2H_Z^{22}} \lesssim \|\theta_0\|_{H^{14}} + B^2.$$

In a similar fashion, $\Theta_a^3$ satisfies the system

$$\begin{cases} Z\partial_Z^5\Theta_a^3 - 3\partial_Z^4\Theta_a^3 = 4\partial_x^2\Theta_a^3 + 4\partial_Z^4 S_a^3 + 6\partial_x^2\partial_Z^2\Theta_a^1 - 2Z\partial_x^2\partial_Z^3\Theta_a^1, \\ \Theta^3_{a|Z=0} = \partial_Z\Theta^3_{a|Z=0} = 0, \\ \Theta_a^3(Z) \to 0 \quad \text{as } Z \to \infty, \end{cases}$$

together with (4-24). Once again, we find that the lifted Fourier transform of $\partial_Z^3\Theta_a^3$ satisfies an equation of the form (3-4) with boundary conditions of the type (iv) from Lemma 3.2. Using the explicit Fourier representation of $\Theta_a^1$ (4-21) together with the estimates on $S_a^3$ (4-23), we find that

$$\|\Theta_a^3\|_{H_x^{19/2}L_Z^2} + \|\Theta_a^3\|_{L_x^2H_Z^{19}} \lesssim \|\theta_0\|_{H^{14}} + B^2, \quad \|\Psi_a^3\|_{H_x^{21/2}L_Z^2} + \|\Psi_a^3\|_{L_x^2H_Z^{21}} \lesssim \|\theta_0\|_{H^{14}} + B^2.$$

Note that the Fourier representation of $\Theta_a^2$ and of the linear part of $\Theta_a^3$ also ensure that, for $j = 2, 3$,

$$\|\partial_Z^2\Theta_a^j|_{Z=0}\|_{H^8(\mathbb{T})} + \|\partial_Z^3\Theta_a^j|_{Z=0}\|_{H^{15/2}(\mathbb{T})} \lesssim \|\theta_0\|_{H^{14}} + B^2. \tag{4-26}$$

Eventually, we define $\Theta_{c,a}^2$ analogously to $\Theta_a^2$ so that

$$\begin{cases} Z\partial_Z^5\Theta_{c,a}^2 - 2\partial_Z^4\Theta_{c,a}^2 = 4\partial_x^2\Theta_{c,a}^2 - 8\partial_x^3\partial_Z^2\Psi_{c,a}^0, \\ \Theta_{c,a}^2|_{Z=0} = \partial_Z\Theta_{c,a}^2|_{Z=0} = 0, \\ \partial_Z^j\Theta_{c,a}^2|_{Z=0} = -2\partial_x^2\partial_Z^{j-2}\Theta_{c,a}^0|_{Z=0} \quad \forall j \in \{4, 5\}, \\ \Theta_{c,a}^2(x, Z) \to 0 \quad \text{as } Z \to \infty. \end{cases}$$

Once again, note that the boundary conditions are redundant. We also define $\Psi^2_{c,a}$ by $\partial_Z^4\Psi^2_{c,a} = \partial_x\Theta^2_{c,a} - 2\partial_x^2\partial_Z^2\Psi^2_{c,a}$, with homogeneous boundary conditions at $Z=0$. We obtain

$$\|\Theta^2_{c,a}\|_{H^9_x L^2_Z} + \|\Theta^2_{c,a}\|_{L^2_x H^{18}_Z} \lesssim B^2(1+T)^{-\frac{1}{2}},$$
$$\|\Psi^2_{c,a}\|_{H^{10}_x L^2_Z} + \|\Psi^2_{c,a}\|_{L^2_x H^{20}_Z} \lesssim B^2(1+T)^{-\frac{1}{2}}.$$

***Boundary layer corrector $\Theta^4_a$.*** As in the previous section, we need to define a higher-order boundary layer corrector $\Theta^4_a$, whose role is to ensure that

$$\|\partial_x^2\Delta^{-2}\theta^{\mathrm{BL}} - \partial_x\psi^{\mathrm{BL}}\|_{L^2} \lesssim B(1+t)^{-3}.$$

To that end, we choose $\Theta^4_a$, $\Psi^4_a$ so that

$$\partial_Z^4\Psi^4_a + 2\partial_x^2\partial_Z^2\Psi^2_a + \partial_x^4\Psi^0_a = \partial_x\Theta^4_a,$$
$$Z\partial_Z\Theta^4_a - 8\Theta^4_a = 4\partial_x\Psi^4_a.$$

Eliminating $\Psi^4_a$ from the equation, we find

$$Z\partial_Z^5\Theta^4_a - 4\partial_Z^4\Theta^4_a = 4\partial_x^2\Theta^4_a - 8\partial_x^3\partial_Z^2\Psi^2_a - 4\partial_x^5\Psi^0_a.$$

We enforce the boundary conditions (which are redundant)

$$\Theta^4_a|_{Z=0} = \partial_Z\Theta^4_a|_{Z=0} = 0, \quad \partial_Z^4\Theta^4_a|_{Z=0} = 2\partial_x^3\partial_Z^2\Psi^2_a|_{Z=0}, \quad \partial_Z^5\Theta^4_a|_{Z=0} = \tfrac{8}{3}\partial_x^3\partial_Z^3\Psi^2_a|_{Z=0},$$

together with a decay assumption at infinity. Looking at the equation satisfied by the Fourier transform and applying Lemma 3.2, we infer that there exists a (nonunique) solution $\Theta^4_a$ of this equation such that

$$\|\Theta^4_a\|_{H^9_x L^2_Z} + \|\Theta^4_a\|_{L^2_x H^{18}_Z} \lesssim \|\theta_0\|_{H^{14}} + B^2.$$

As in the previous section (see the discussion on page 1993), nonuniqueness comes from the fact that the Fourier transform of $\partial_Z^4\Theta^4_a$ satisfies an ODE of the form (3-4), with boundary conditions at $Z=0$ for $\partial_Z^4\Theta^4_a$ and $\partial_Z^5\Theta^4_a$. However, the boundary conditions above do not prescribe any condition on $\partial_Z^k\Theta^4_a$ for any $k \ge 6$. We lift this indetermination by requiring (somewhat arbitrarily) that $\partial_Z^8\Theta^4_a|_{Z=0} = 0$. The solution thus obtained satisfies the previous Sobolev estimates, and its trace satisfies

$$\|\partial_Z^2\Theta^4_a|_{Z=0}\|_{H^8(\mathbb{T})} + \|\partial_Z^3\Theta^4_a|_{Z=0}\|_{H^{15/2}(\mathbb{T})} \lesssim \|\theta_0\|_{H^{14}} + B^2. \tag{4-27}$$

***Lift of the remaining traces of order $B$.*** At this stage, we have defined $\Theta^j_a$, $\Psi^j_a$ for $0 \le j \le 4$ together with $\Theta^j_{c,a}$, $\Psi^j_{c,a}$ for $0 \le j \le 2$. Let $\chi \in C_c^\infty(\mathbb{R})$ be a cut-off function such that $\chi \equiv 1$ on $\left(-\frac{1}{4}, \frac{1}{4}\right)$ and $\operatorname{Supp}\chi \subset \left(-\frac{1}{2}, \frac{1}{2}\right)$. Setting $\Theta^j_{c,a} = \Psi^j_{c,a} = 0$ for $j \ge 3$ and $Z_{\mathrm{bot}} = (1+t)^{1/4}z$, $Z_{\mathrm{top}} = (1+t)^{1/4}(1-z)$, the main boundary layer term is given by

$$\theta^{\mathrm{BL}}_{\mathrm{main}} := \sum_{j=0}^{4}(1+t)^{-1-\frac{j}{4}}(\Theta^j_{\mathrm{bot}} + \Theta^j_{c,\mathrm{bot}})(x, Z_{\mathrm{bot}})\chi(z) + \sum_{j=0}^{4}(1+t)^{-1-\frac{j}{4}}(\Theta^j_{\mathrm{top}} + \Theta^j_{c,\mathrm{top}})(x, Z_{\mathrm{top}})\chi(1-z),$$
$$\psi^{\mathrm{BL}}_{\mathrm{main}} := \sum_{j=0}^{4}(1+t)^{-2-\frac{j}{4}}(\Psi^j_{\mathrm{bot}} + \Psi^j_{c,\mathrm{bot}})(x, Z_{\mathrm{bot}})\chi(z) + \sum_{j=0}^{4}(1+t)^{-2-\frac{j}{4}}(\Psi^j_{\mathrm{top}} + \Psi^j_{c,\mathrm{top}})(x, Z_{\mathrm{top}})\chi(1-z).$$

By construction, we have

$$\Delta^2\theta^{\mathrm{BL}}_{\mathrm{main}}|_{z=0} = \gamma^0_{\mathrm{bot},T} + \gamma^0_{\mathrm{bot},2}(1+t)^{-\frac12} + \gamma^0_{\mathrm{bot},3}(1+t)^{-\frac34} - \gamma^0_{\mathrm{bot},2}(1+T)^{-\frac12} - \gamma^0_{\mathrm{bot},3}(1+T)^{-\frac34} + 2(1+t)^{-\frac54}\partial_x^2\partial_Z^2\Theta^3_{\mathrm{bot}}|_{Z=0} + 2(1+t)^{-\frac32}\partial_x^2\partial_Z^2\Theta^4_{\mathrm{bot}}|_{Z=0},$$

$$\partial_z\Delta^2\theta^{\mathrm{BL}}_{\mathrm{main}}|_{z=0} = \gamma^1_{\mathrm{bot},T} + \gamma^1_{\mathrm{bot},1}(1+t)^{-\frac14} + \gamma^1_{\mathrm{bot},2}(1+t)^{-\frac12} - \gamma^1_{\mathrm{bot},1}(1+T)^{-\frac14} - \gamma^1_{\mathrm{bot},2}(1+t)^{-\frac12} + 2(1+t)^{-1}\partial_x^2\partial_Z^3\Theta^3_{\mathrm{bot}}|_{Z=0} + 2(1+t)^{-\frac54}\partial_x^2\partial_Z^3\Theta^4_{\mathrm{bot}}|_{Z=0}.$$

Similar formulas hold at $z = 1$. Comparing with Lemmas 4.9 and 4.10, we see that we need to lift the traces of $\partial_x^2\partial_Z^k\Theta^j_a$ for $k = 2, 3$ and $j \geq 3$. We lift these remaining traces thanks to a corrector $\sigma^{\mathrm{lin}}_{\mathrm{lift}}$ which we define in Fourier space in the following way. Let $\zeta_4, \zeta_5 \in C^\infty_c(\mathbb{R})$ such that $\zeta_j(Z) = Z^j/j!$ in a neighborhood of zero and such that $\operatorname{Supp}\zeta_j \subset \left(-\frac14, \frac14\right)$. In order to apply the last estimate of Lemma 3.6, we further choose $\zeta_j$ so that

$$\int_0^\infty Z^k\zeta_j(Z)\,\mathrm{d}Z = 0 \quad \forall k \in \{2, 3\}. \tag{4-28}$$

We then take

$$\widehat{\sigma^{\mathrm{lin}}_{\mathrm{lift}}}(t,k,z) = 2\sum_{l\geq3,\, j=0,1}(1+t)^{-\frac32-\frac{j+l}{4}}|k|^{-2-j}\,\widehat{\partial_Z^{2+j}\Theta^l_{\mathrm{bot}}}(k)|_{Z=0}\,\zeta_{4+j}(|k|z(1+t)^{\frac14}) + 2\sum_{l\geq3,\, j=0,1}(1+t)^{-\frac32-\frac{j+l}{4}}|k|^{-2-j}\,\widehat{\partial_Z^{2+j}\Theta^l_{\mathrm{top}}}(k)|_{Z=0}\,\zeta_{4+j}(|k|(1-z)(1+t)^{\frac14}),$$

so that

$$\Delta^2\sigma^{\mathrm{lin}}_{\mathrm{lift}}|_{z=0} = -2(1+t)^{-\frac54}\partial_x^2\partial_Z^2\Theta^3_{\mathrm{bot}}|_{Z=0} - 2(1+t)^{-\frac32}\partial_x^2\partial_Z^2\Theta^4_{\mathrm{bot}}|_{Z=0},$$
$$\partial_{\boldsymbol{n}}\Delta^2\sigma^{\mathrm{lin}}_{\mathrm{lift}}|_{z=0} = -2(1+t)^{-1}\partial_x^2\partial_Z^3\Theta^3_{\mathrm{bot}}|_{Z=0} - 2(1+t)^{-\frac54}\partial_x^2\partial_Z^3\Theta^4_{\mathrm{bot}}|_{Z=0}.$$

The estimates on the traces $\Theta^j_a$ for $j \geq 2$ (see (4-26), (4-27)) ensure that, for all $k, m \geq 0$ such that $k + m \leq 10$,

$$\begin{aligned}\|\sigma^{\mathrm{lin}}_{\mathrm{lift}}\|_{H^m_x H^k_z} &\lesssim (\|\theta_0\|_{H^{14}} + B^2)(1+t)^{-2-\frac18+\frac k4},\\ \|\partial_t\sigma^{\mathrm{lin}}_{\mathrm{lift}}\|_{H^m_x H^k_z} &\lesssim (\|\theta_0\|_{H^{14}} + B^2)(1+t)^{-3-\frac18+\frac k4}.\end{aligned} \tag{4-29}$$

We define an associated corrector $\phi^{\mathrm{lin}}_{\mathrm{lift}} = \Delta^{-2}\partial_x\sigma^{\mathrm{lin}}_{\mathrm{lift}}$. According to Lemma 3.6 and using (4-28), we have, for all $k, m \geq 0$ such that $k + m \leq 13$,

$$\begin{aligned}\|\phi^{\mathrm{lin}}_{\mathrm{lift}}\|_{H^m_x H^k_z} &\lesssim (\|\theta_0\|_{H^{14}} + B^2)(1+t)^{-3-\frac18+\frac k4},\\ \|\partial_t\phi^{\mathrm{lin}}_{\mathrm{lift}}\|_{H^m_x H^k_z} &\lesssim (\|\theta_0\|_{H^{14}} + B^2)(1+t)^{-4-\frac18+\frac k4}.\end{aligned} \tag{4-30}$$

***Evaluation of the remainder.*** Let us now focus on the different remainder terms in the equation satisfied by $\theta^{\mathrm{BL}}_{\mathrm{main}}$, in view of defining one last linear corrector.

• *Remainder stemming from the nonlinear term:* Using Lemma 3.5 together with the estimates on $\Theta^j_a$, we have

$$\begin{aligned}\nabla^\perp\psi^{\mathrm{BL}}_{\mathrm{main}}\cdot\nabla\theta^{\mathrm{BL}}_{\mathrm{main}} = &\sum_{0\leq j,k\leq4}(1+t)^{-3-\frac{k+j-1}{4}}\{\Psi^j_{\mathrm{bot}} + \Psi^j_{c,\mathrm{bot}}, \Theta^k_{\mathrm{bot}} + \Theta^k_{c,\mathrm{bot}}\}_{x,Z}(x, Z_{\mathrm{bot}})\chi(z)\\ &- \sum_{0\leq j,k\leq4}(1+t)^{-3-\frac{k+j-1}{4}}\{\Psi^j_{\mathrm{top}} + \Psi^j_{c,\mathrm{top}}, \Theta^k_{\mathrm{top}} + \Theta^k_{c,\mathrm{top}}\}_{x,Z}(x, Z_{\mathrm{top}})\chi(1-z)\\ &+ O(\exp(-c(1+t)^{\frac15})) \quad \text{in } H^9(\Omega).\end{aligned}$$

In the above expansion, we put aside the terms corresponding to $k=j=0$, which are part of $S_a^3$ and are lifted by $\Theta_a^3$. If $j+k\geq 1$, we have, when $0\leq s+r\leq 8$,

$$\|\{\Psi_{\text{bot}}^j+\Psi_{c,\text{bot}}^j,\,\Theta_{\text{bot}}^k+\Theta_{c,\text{top}}^k\}_{x,Z}(x,\,(1+t)^{\frac14}z)\chi(z)\|_{H_x^rH_z^s}\lesssim B^2(1+t)^{\frac{s}{4}-\frac18},$$

and the same estimate holds for the top boundary layer. We infer that

$$\begin{aligned}\nabla^\perp\psi_{\text{main}}^{\text{BL}}\cdot\nabla\theta_{\text{main}}^{\text{BL}}=(1+t)^{-\frac{11}{4}}\{\Psi_{\text{bot}}^0,\,\Theta_{\text{bot}}^0\}_{x,Z}(x,\,Z_{\text{bot}})\chi(z)&\\ -(1+t)^{-\frac{11}{4}}\{\Psi_{\text{top}}^0,\,\Theta_{\text{top}}^0\}_{x,Z}(x,\,Z_{\text{top}})\chi(1-z)+R_{\text{NL}},&\end{aligned}$$

where, for all $r,s\geq 0,\ r+s\leq 8$,

$$\|R_{\text{NL}}\|_{H_x^rH_z^s}\lesssim B^2(1+t)^{-3+\frac{s}{4}-\frac18}.$$

Note in particular that $\|R_{\text{NL}}\|_{H^8}\lesssim B^2(1+t)^{-1-\delta}$ with $\delta=\frac18$.

• *Remainder stemming from the Taylor expansion of $G$:* As explained in the construction of $\Theta_a^2$, $\Theta_a^3$, when defining the boundary layer term, we replaced $G$ by its Taylor expansion in the vicinity of $z=0$ and $z=1$. Recalling (4-17), we have, in the vicinity of $z=0$, setting $Z=(1+t)^{1/4}z$,

$$\begin{aligned}G\partial_x\psi_{\text{main}}^{\text{BL}}&=\frac{1}{2(1+t)^{\frac12}}g_{\text{bot},T}^2Z^2\partial_x\psi_{\text{main}}^{\text{BL}}+\frac{1}{6(1+t)^{\frac34}}g_{\text{bot},T}^3Z^3\partial_x\psi_{\text{main}}^{\text{BL}}+O((1+t)^{-1}(Z^2+Z^4)\partial_x\psi_{\text{main}}^{\text{BL}})\\&=\frac{1}{2(1+t)^{\frac52}}g_{\text{bot},T}^2Z^2\partial_x\Psi_{\text{bot}}^0(x,\,Z)\chi(z)\\&\qquad+(1+t)^{-\frac{11}{4}}\big(\tfrac12 g_{\text{bot},T}^2Z^2\partial_x\Psi_{\text{bot}}^1(x,\,Z)+\tfrac16 g_{\text{bot},T}^3Z^3\partial_x\Psi_{\text{bot}}^0(x,\,Z)\big)\chi(z)+R_G,\end{aligned}$$

where the first two terms enter the definition of $\Theta_{\text{bot}}^2$ and $\Theta_{\text{bot}}^3$ respectively, and the remainder term $R_G$ satisfies

$$\|R_G\|_{H_x^rH_z^s}\lesssim B^2(1+t)^{-3+\frac{s}{4}-\frac18}\quad\text{if } 0\leq r+s\leq 8.$$

• *Remainder stemming from $\psi^{\text{BL}}-\Delta^{-2}\partial_x\theta^{\text{BL}}$:* We now address the fact that $\Delta^2\psi_{\text{main}}^{\text{BL}}$ is not equal to $\partial_x\theta_{\text{main}}^{\text{BL}}$. More precisely, using the definition of $\Psi_a^j$, we have, in $\Omega\cap\{z\leq\frac12\}$,

$$\begin{aligned}\Delta^2\psi_{\text{main}}^{\text{BL}}-\partial_x\theta_{\text{main}}^{\text{BL}}=2\sum_{j=3,4}&(1+t)^{-2-\frac{j-2}{4}}\partial_x^2\partial_Z^2\Psi_{\text{bot}}^j(x,\,(1+t)^{\frac14}z)\chi(z)\\&+\sum_{j\geq1}(1+t)^{-2-\frac{j}{4}}\partial_x^4\Psi_{\text{bot}}^j(x,\,(1+t)^{\frac14}z)\chi(z)\\&+2\sum_{j=1,2}(1+t)^{-2-\frac{j-2}{4}}\partial_x^2\partial_Z^2\Psi_{c,\text{bot}}^j(x,\,(1+t)^{\frac14}z)\chi(z)\\&+\sum_{j\geq0}(1+t)^{-2-\frac{j}{4}}\partial_x^4\Psi_{c,\text{bot}}^j(x,\,(1+t)^{\frac14}z)\chi(z)\\&+O(\exp(-c(1+t)^{\frac15}))\quad\text{in } H^8(\Omega).\end{aligned}$$

A similar expression holds in $\Omega\cap\{z\geq\frac12\}$, replacing bot with top and $z$ with $1-z$. The exponentially small remainder comes from the commutator of the bilaplacian with multiplication by $\chi$ (see Lemma 3.5), and from the estimates on $\Psi_a^j$, $\Psi_{c,a}^j$. We now apply Lemma 3.6 and its variant Remark 3.9: more precisely, in order to avoid a high loss of horizontal derivatives, we apply the "self-similar version" from Remark 3.9 to the term involving $\partial_x^4 \Psi_\perp^1$, and the second statement from Lemma 3.6 to all other terms. We obtain

$$\partial_x \psi_{\text{main}}^{\text{BL}} - \Delta^{-2} \partial_x^2 \theta_{\text{main}}^{\text{BL}} =: R_{\Delta^2},$$

with

$$\sup_{t \in [0,T]} \big( (1+t)^3 \|\partial_x^5 R_{\Delta^2}\|_{L^2} + (1+t)^{2+\frac{3}{8}} \|\partial_x^3 \Delta^2 R_{\Delta^2}\|_{L^2} + (1+t)^{1+\frac{3}{8}} \|\partial_x \Delta^4 R_{\Delta^2}\|_{L^2} \big) \lesssim B.$$

Note that the decay of this remainder is similar to the one of $R_{\text{NL}}$ and $R_G$, but its order of magnitude is $B$. Hence we call it a "linear" remainder. In order to simplify the forthcoming bootstrap argument, we will lift it thanks to another (linear) corrector.

- *Remainder stemming from $\sigma_{\text{lift}}^{\text{lin}}$:* Recalling (4-29), (4-30) and using a variant of Remark 3.9, we have, setting $R_{c,\text{lin}} = \partial_t \sigma_{\text{lift}}^{\text{lin}} - \Delta^{-2} \partial_x^2 \sigma_{\text{lift}}^{\text{lin}}$, for $k + m \le 10$,

$$\sup_{t \in [0,T]} \big( (1+t)^3 \|\partial_x^5 R_{c,\text{lin}}\|_{L^2} + (1+t)^2 \|\partial_x^3 \Delta^2 R_{c,\text{lin}}\|_{L^2} + (1+t)^{\frac{9}{8}} \|\partial_x \Delta^4 R_{c,\text{lin}}\|_{L^2} \big) \lesssim B.$$

Once again, $R_{c,\text{lin}}$ is a linear remainder, and shall be lifted before the bootstrap argument of the next subsection. We also have

$$\|G \partial_x^7 \Delta^{-2} \sigma_{\text{lift}}^{\text{lin}}\|_{L^2} \lesssim B^2 (1+t)^{-3}.$$

In the remainders above, all terms of order $B^2(1+t)^{-3}$ in $L^2$ will be included in the remainder for the interior part (see Section 4.5), while the terms of order $B(1+t)^{-3}$ will be lifted thanks to a linear corrector $\sigma^R$, which we now construct.

***Definition of $\boldsymbol{\sigma^R}$.*** Let $\sigma^R$ be the solution of

$$\partial_t \sigma^R = \partial_x^2 \Delta^{-2} \sigma^R - R_{\Delta^2} - R_{c,\text{lin}}, \quad \sigma^R(t=0) = 0.$$

Note that $\partial_t \sigma^R|_{\partial\Omega} = \partial_t \partial_{\boldsymbol{n}} \sigma^R|_{\partial\Omega} = 0$, and therefore $\sigma^R|_{\partial\Omega} = \partial_{\boldsymbol{n}} \sigma^R|_{\partial\Omega} = 0$ for all $t > 0$. Applying $\Delta^2$ to the above equation and taking the trace at $z = 0$, we have, using the identity (4-18)

$$\begin{aligned}
\partial_t \Delta^2 \sigma^R|_{z=0} &= -\Delta^2 (R_{\Delta^2} + R_{c,\text{lin}})|_{z=0} \\
&= -\partial_x \Delta^2 \psi_{\text{main}}^{\text{BL}}|_{z=0} - \partial_t \partial_z^4 \sigma_{\text{lift}}^{\text{lin}}|_{z=0} \\
&= -2 \sum_{j \ge 3} (1+t)^{-2-\frac{j-2}{4}} \partial_x^3 \partial_Z^2 \Psi_{\text{bot}}^j|_{Z=0} + 2 \sum_{j \ge 3} \tfrac{1}{4}(2+j)(1+t)^{-2-\frac{j-2}{4}} \partial_x^2 \partial_Z^2 \Theta_{\text{bot}}^j|_{Z=0} = 0.
\end{aligned}$$

Hence $\Delta^2 \sigma^R|_{z=0} = 0$ for all $t \in (0, T)$. In a similar way, $\partial_z \Delta^2 \sigma^R|_{z=0} = 0$ for all $t \in (0, T)$, and the same properties hold at $z = 1$. Applying first Lemma 2.4 to $\partial_x \Delta^4 \sigma^R$, and then Proposition 2.6 to $\partial_x^3 \Delta^2 \sigma^R$, $\partial_x^5 \sigma^R$, and $\partial_t \partial_x^4 \sigma^R$, we infer

$$\begin{aligned}
\|\partial_x \Delta^4 \sigma^R\|_{L^2} &\lesssim \|\theta_0\|_{H^{14}} + B^2, \\
\|\partial_x^3 \Delta^2 \sigma^R\|_{L^2} &\lesssim (\|\theta_0\|_{H^{14}} + B^2)(1+t)^{-1}, \\
\|\partial_x^5 \sigma^R\|_{L^2} &\lesssim (\|\theta_0\|_{H^{14}} + B^2)(1+t)^{-2}, \\
\|\partial_t \partial_x^4 \sigma^R\|_{L^2} &\lesssim (\|\theta_0\|_{H^{14}} + B^2)(1+t)^{-3}, \\
\|\partial_x^6 \Delta^{-2} \sigma^R\|_{L^2} &\lesssim (\|\theta_0\|_{H^{14}} + B^2)(1+t)^{-3}.
\end{aligned}$$

Furthermore, looking at the expressions of $R_{\Delta^2}$ and $R_{c,\mathrm{lin}}$ and recalling the estimates on $\Psi_a^j$, we can perform similar estimates for $z\chi(z)\partial_z\sigma^R$ and $(1-z)\chi(1-z)\partial_z\sigma^R$. For instance, estimating the commutators, we find that

$$\begin{aligned}\partial_t\Delta^4(z\chi(z)\partial_z\sigma^R) &= \Delta^4(z\chi(z)\partial_z\partial_x^2\Delta^{-2}\sigma^R) - \Delta^4(z\chi(z)\partial_z(R_{\Delta^2}+R_{c,\mathrm{lin}}))\\ &= \partial_x^2\Delta^2(z\chi(z)\partial_z\sigma^R) + [\Delta^4, z\chi(z)\partial_z]\partial_x^2\Delta^{-2}\sigma^R + \partial_x^2[z\chi(z)\partial_z, \Delta^2]\sigma^R\\ &\qquad - \Delta^4(z\chi(z)\partial_z(R_{\Delta^2}+R_{c,\mathrm{lin}})),\end{aligned}$$

where

$$\|[\Delta^4, z\chi(z)\partial_z]\partial_x^2\Delta^{-2}\sigma^R\|_{L^2} + \|\partial_x^2[z\chi(z)\partial_z, \Delta^2]\sigma^R\|_{L^2} \lesssim \|\partial_x^2\Delta^2\sigma^R\|_{L^2} \lesssim (\|\theta_0\|_{H^{14}}+B^2)(1+t)^{-1},$$
$$\|\Delta^4(z\chi(z)\partial_z(R_{\Delta^2}+R_{c,\mathrm{lin}}))\|_{L^2} \lesssim (\|\theta_0\|_{H^{14}}+B^2)(1+t)^{-\frac{9}{8}}.$$

It follows that, for all $t\in[0,T]$,

$$\|\Delta^4(z\chi(z)\partial_z\sigma^R)\|_{L^2} \lesssim (\|\theta_0\|_{H^{14}}+B^2)\ln(2+t).$$

***Conclusion.*** Let

$$\theta_c := \sigma_{\mathrm{lift}}^{\mathrm{lin}} + \sigma^R + \theta_{\mathrm{bot}}^{\mathrm{BL}}(\chi(z)-1) + \theta_{\mathrm{top}}^{\mathrm{BL}}(\chi(1-z)-1), \quad \theta^{\mathrm{BL}} = \theta_{\mathrm{bot}}^{\mathrm{BL}} + \theta_{\mathrm{top}}^{\mathrm{BL}},$$

where

$$\theta_a^{\mathrm{BL}} = \sum_{j=0}^{4}(1+t)^{-1-\frac{j}{4}}(\Theta_a^j + \Theta_{c,a}^j)(x, Z_a), \quad a\in\{\mathrm{top},\mathrm{bot}\}.$$

Then (up to a redefinition of $\Theta_a^j+\Theta_{c,a}^j$ as $\Theta_a^j$), the bounds on the profiles $\Theta_a^j$ and the corrector $\theta_c$, together with the boundary conditions on $\theta^{\mathrm{BL}}+\theta_c$ announced in the statement of Lemma 4.10, are all satisfied.

Most of the remainder terms have already been evaluated. There only remains to evaluate the quadratic terms involving $\sigma_{\mathrm{lin}}^{\mathrm{lift}}$ and $\sigma^R$. We have for instance

$$\|\partial_x^2\Delta^2(\nabla^\perp\psi_{\mathrm{main}}^{\mathrm{BL}})\cdot\nabla\sigma^R\|_{H_x^rH_z^s} \lesssim B^2(1+t)^{-\frac{3}{4}}(1+t)^{-2+\frac{1}{4}} \lesssim B^2(1+t)^{-\frac{5}{2}}.$$

For the $H^8$ estimate, we write, for $z\le\frac{1}{2}$,

$$\partial_x\psi_{\mathrm{main}}^{\mathrm{BL}}\partial_z\sigma^R = \frac{\partial_x\psi_{\mathrm{main}}^{\mathrm{BL}}}{z}z\partial_z\sigma^R.$$

Both terms in the right-hand side belong to $H^8$, and we infer

$$\|\nabla^\perp\psi_{\mathrm{main}}^{\mathrm{BL}}\cdot\nabla\sigma^R\|_{H^8} \lesssim B^2(1+t)^{-\frac{5}{4}}\ln(2+t).$$

The statement of Lemma 4.10 follows. □

*Proof of Lemma 4.13.* Assume that $\theta^{\mathrm{rem}}=\theta'-\theta^{\mathrm{BL}}$ satisfies (4-2), and define $\Gamma_{a,T}^j$ as in Lemma 4.9. According to Lemma 4.9,

$$\|\Gamma_{a,T}^j(t)\|_{L^2(\mathbb{T})} \lesssim B^2(1+t)^{-1+\frac{j}{4}},$$
$$\|\partial_t\Gamma_{a,T}^j(t)\|_{L^2(\mathbb{T})} \lesssim B^2(1+t)^{-2+\frac{j}{4}},$$
$$\|\Gamma_{a,T}^j(t)\|_{H^4(\mathbb{T})} \lesssim B^2(1+t)^{-\frac{23}{24}+\frac{j}{4}}.$$

We now lift these traces thanks to a corrector $\sigma_{\text{lift}}^{\text{NL}}$, whose definition is similar to the one of $\sigma_{\text{lift}}^{\text{lin}}$, namely

$$\widehat{\sigma_{\text{lift}}^{\text{NL}}}(t,k,z) = \sum_{j=0,1} (1+t)^{-1-\frac{j}{4}} |k|^{-4-j} \widehat{\Gamma_{\text{bot},T}^{j}}(t,k) \zeta_{4+j}(|k| z (1+t)^{\frac{1}{4}}) + \sum_{j=0,1} (1+t)^{-1-\frac{j}{4}} |k|^{-4-j} (-1)^j \widehat{\Gamma_{\text{top},T}^{j}}(t,k) \zeta_{4+j}(|k|(1-z)(1+t)^{\frac{1}{4}}),$$

where we recall that $\zeta_j \in C_c^\infty(\mathbb{R})$, $\zeta(Z) = Z^j/j!$ in a neighborhood of zero, and $\zeta_j$ satisfies (4-28).

It follows from the estimates on $\Gamma_{a,T}^j$ and from the formula defining $\sigma_{\text{lift}}^{\text{NL}}$ that, for $\ell = 0, 1$,

$$\|\partial_t^\ell \sigma_{\text{lift}}^{\text{NL}}\|_{H_x^m H_z^k} \lesssim B^2 (1+t)^{-2-\ell+\frac{k}{4}-\frac{1}{8}} \quad \text{if } k+m \le \tfrac{9}{2},$$

$$\|\sigma_{\text{lift}}^{\text{NL}}\|_{H^8(\Omega)} \lesssim B^2(1+t)^{-\frac{1}{12}},$$

$$\|z\sigma_{\text{lift}}^{\text{NL}}\|_{H^9(\Omega\cap\{z\le\frac{1}{2}\})} + \|(1-z)\sigma_{\text{lift}}^{\text{NL}}\|_{H^9(\Omega\cap\{z\ge\frac{1}{2}\})} \lesssim B^2(1+t)^{-\frac{1}{12}}.$$

The function $\sigma_{\text{lift}}^{\text{NL}}$ has been designed so that

$$\Delta^2 \sigma_{\text{lift}}^{\text{NL}}|_{z=0} = \Gamma_{\text{bot}}^0(t), \quad \partial_{\boldsymbol{n}} \Delta^2 \sigma_{\text{lift}}^{\text{NL}}|_{z=0} = \Gamma_{\text{bot}}^1(t),$$

$$\Delta^2 \sigma_{\text{lift}}^{\text{NL}}|_{z=1} = \Gamma_{\text{top}}^0(t), \quad \partial_{\boldsymbol{n}} \Delta^2 \sigma_{\text{lift}}^{\text{NL}}|_{z=1} = \Gamma_{\text{top}}^1(t).$$

Furthermore, according to Remark 3.9 and using (4-28), we have, for all $k, m \ge 0$ such that $k+m \le 8$,

$$\|\Delta^{-2}\sigma_{\text{lift}}^{\text{NL}}\|_{H_x^m H_z^k} \lesssim B^2(1+t)^{-3-\frac{1}{8}+\frac{k}{4}},$$

$$\|\partial_t \Delta^{-2}\sigma_{\text{lift}}^{\text{NL}}\|_{H_x^m H_z^k} \lesssim B^2(1+t)^{-4-\frac{1}{8}+\frac{k}{4}}.$$

The statement of Lemma 4.13 follows immediately from these estimates and Lemmas 4.9 and 4.10. ☐

**4.5. *Bootstrap argument for $\theta^{\text{int}}$*.** In this subsection, we complete the proof of Theorem 1.3 thanks to a bootstrap argument (or rather, two nested bootstrap arguments). We start with an initial data $\theta_0 \in H^{14}(\Omega)$, with $\|\theta_0\|_{H^{14}(\Omega)} \le B$ and $\theta_0 = \partial_{\boldsymbol{n}}\theta_0 = 0$ on $\partial\Omega$, $\partial_z^2\bar{\theta}_0 = 0$ on $\partial\Omega$. We assume that $B \le B_0 < 1$, where $B_0$ is a small universal constant, so that Theorem 1.1 holds.

Let $\bar{C} \ge 2$ be a universal constant to be determined. We define

$$T_1 = \sup\{T > 0 : \text{(4-3) holds on } (0,T) \text{ with } B_1 = \bar{C}\|\theta_0\|_{H^{14}}\}.$$

By continuity, $T_1 > 0$. For any $T \in (0, T_1)$, we define an associated boundary layer profile $\theta_T^{\text{BL}}$ (see Lemma 4.10 and Remark 4.12) together with a corrector $\theta_c$. We recall that there exists a universal constant $C_1$ such that, for all $m, k \ge 0$ with $k+m \le 8$, for all $T \in (0, T_1)$, $t \in [0, T]$,

$$\|\theta_T^{\text{BL}}(t)\|_{H_x^m H_z^k} \le C_1(\|\theta_0\|_{H^{14}} + B_1^2)(1+t)^{-1+\frac{k}{4}-\frac{1}{8}} \le 2C_1\|\theta_0\|_{H^{14}}(1+t)^{-1+\frac{k}{4}-\frac{1}{8}},$$

provided $\bar{C}^2 B_0 \le 1$. Similarly, for all $t \in [0, T_1]$,

$$(1+t)^2\|\partial_x^5\theta_c\|_{L^2} + \|\partial_x\Delta^4\theta_c\|_{L^2} + (1+t)^3\|\partial_t\partial_x^4\theta_c\|_{L^2} + (1+t)^3\|\partial_x^5\psi_c\|_{L^2} \le 2C_1\|\theta_0\|_{H^{14}}.$$

We then introduce a new time

$$T_2 = \sup\{T \in (0, T_1) : \theta^{\text{rem}} = \theta' - \theta_T^{\text{BL}} \text{ satisfies (4-2) on } (0,T) \text{ with } B_2 = (2\bar{C} + 3C_1)\|\theta_0\|_{H^{14}}\}. \tag{4-31}$$

On $(0, T_2)$, relying on Lemma 4.13, we construct an approximate solution $\theta^{\text{app}}$. We now set $\theta^{\text{int}} = \theta' - \theta^{\text{app}} = \theta^{\text{rem}} - \theta_c - \sigma_{\text{lift}}^{\text{NL}}$. Note that we can always choose $\|\theta_0\|_{H^{14}}$ small enough so that for all $t \in (0, T_2)$, for $0 \le k+m \le 8$,

$$\|\sigma_{\text{lift}}^{\text{NL}}\|_{H_x^m H_z^k} \le \overline{C}\|\theta_0\|_{H^{14}}(1+t)^{-2+\frac{k}{4}-\frac{1}{8}}.$$

Consequently, $\theta^{\text{int}}$ satisfies (4-2) with $B_3 = (3\overline{C}+5C_1)\|\theta_0\|_{H^{14}}$ on $(0, T_2)$. Note that $B_j \lesssim B$ for $j = 1, 2, 3$.

Our goal is now to prove that $T_1 = T_2 = +\infty$ for a suitable choice of $\overline{C}$, provided $\|\theta_0\|_{H^{14}}$ is sufficiently small. To that end, we check that $\Delta^2\theta^{\text{int}}$ satisfies the assumptions of Proposition 2.6.

By construction (see Lemma 4.13),

$$\theta^{\text{int}} = \partial_{\boldsymbol{n}}\theta^{\text{int}} = \Delta^2\theta^{\text{int}} = \partial_{\boldsymbol{n}}\Delta^2\theta^{\text{int}} = 0 \quad \text{on } \partial\Omega.$$

Furthermore, defining the quadratic form

$$Q(f, g) = -(\nabla^{\perp}\Delta^{-2}\partial_x f \cdot \nabla g)',$$

we have, recalling Lemma 4.13,

$$\partial_t\theta^{\text{int}} = (1-G)\partial_x^2\Delta^{-2}\theta^{\text{int}} + S_{\text{rem}}^1, \tag{4-32}$$

where, recalling the definition of $S_{\text{rem}}$ from Lemma 4.13,

$$S_{\text{rem}}^1 = -S_{\text{rem}} + Q(\theta^{\text{app}} + \theta^{\text{int}}, \theta^{\text{int}}) + Q(\theta^{\text{int}}, \theta^{\text{app}}).$$

We claim that we have the following estimates on $S_{\text{rem}}^1$:

**Lemma 4.14** (estimates on $S_{\text{rem}}^1$). *Let $T_2$ be defined by* (4-31).

- *$L^2$ and $H^4$ estimates*: *for all $t \in [0, T_2)$,*

$$\|\partial_x^4 S_{\text{rem}}^1(t)\|_{L^2} \lesssim B^2\frac{1}{(1+t)^3}, \quad \|\partial_x^2\Delta^2 S_{\text{rem}}^2(t)\|_{L^2} \lesssim B^2\frac{1}{(1+t)^2}.$$

- *$H^8$ estimate*: *there exist $S_{\parallel}^2, S_{\perp}^2 \in L^\infty([0, T_2), L^2(\Omega))$ such that $\Delta^4 S_{\text{rem}}^1(t) = S_{\parallel}^2 + S_{\perp}^2$, with*

$$\|S_{\parallel}^2(t)\|_{L^2} \lesssim B^2\frac{1}{(1+t)^{\frac{9}{8}}} \quad \forall t \in [0, T_2) \quad \textit{and} \quad \int_\Omega S_{\perp}^2(t)\Delta^4\theta^{\text{int}}(t) = 0.$$

- *Estimates on the time derivative*: *for all $t \in [0, T_2)$,*

$$\|\partial_t\partial_x^4 S_{\text{rem}}^1(t)\|_{L^2} \lesssim B^2\frac{1}{(1+t)^4}.$$

*Proof.* We estimate each term separately. The estimates on $S_{\text{rem}}$ have already been proved in the previous subsection (see Lemma 4.13). Therefore we focus on the quadratic terms. It follows from the estimates of Lemmas 4.10, 4.13 and from the bootstrap estimates (4-2) on $\theta^{\text{rem}}$ that

$$\|\partial_x^4 Q(\theta^{\text{app}} + \theta^{\text{int}}, \theta^{\text{int}})\|_{L^2} \lesssim B^2(1+t)^{-3},$$
$$\|\partial_x^2\Delta^2 Q(\theta^{\text{app}} + \theta^{\text{int}}, \theta^{\text{int}})\|_{L^2} \lesssim B^2(1+t)^{-2}.$$

For the $H^8$ estimate, the situation is slightly different, because $\Delta^4 Q(\theta^{\mathrm{app}}+\theta^{\mathrm{int}}, \theta^{\mathrm{int}})$ involves derivatives of order 9 of $\theta^{\mathrm{int}}$, for which we have no estimate. Therefore, as in Section 2, we decompose $\Delta^4 Q(\theta^{\mathrm{app}}+\theta^{\mathrm{int}}, \theta^{\mathrm{int}})$ into two parts, writing

$$\Delta^4 Q(\theta^{\mathrm{app}}+\theta^{\mathrm{int}}, \theta^{\mathrm{int}}) = -(\nabla^\perp \Delta^{-2}\partial_x(\theta^{\mathrm{app}}+\theta^{\mathrm{int}}))\cdot\nabla\Delta^4\theta^{\mathrm{int}} - \partial_z^8\,\overline{\nabla^\perp\Delta^{-2}\partial_x(\theta^{\mathrm{app}}+\theta^{\mathrm{int}})\cdot\theta^{\mathrm{int}}} - [\Delta^4, \nabla^\perp\Delta^{-2}\partial_x(\theta^{\mathrm{app}}+\theta^{\mathrm{int}})\cdot\nabla]\theta^{\mathrm{int}}.$$

It can be easily checked that the term $[\Delta^4, \nabla^\perp\Delta^{-2}\partial_x(\theta^{\mathrm{app}}+\theta^{\mathrm{int}})\cdot\nabla]\theta^{\mathrm{int}}$ can be evaluated as above, and we have

$$\|[\Delta^4, \nabla^\perp\Delta^{-2}\partial_x(\theta^{\mathrm{app}}+\theta^{\mathrm{int}})\cdot\nabla]\theta^{\mathrm{int}}\|_{L^2} \lesssim B^2(1+t)^{-2+\frac{9}{4}}(1+t)^{-2+\frac{1}{4}} \lesssim B^2(1+t)^{-\frac{3}{2}}.$$

Furthermore, since $\langle \Delta^4\theta^{\mathrm{int}}(t,\cdot,z)\rangle = 0$ for all $t, z$,

$$\int_\Omega \partial_z^8\overline{\nabla^\perp\Delta^{-2}\partial_x(\theta^{\mathrm{app}}+\theta^{\mathrm{int}})\cdot\theta^{\mathrm{int}}}\,\Delta^4\theta^{\mathrm{int}} = 0.$$

Eventually, integrating by parts the remaining term,

$$-\int_\Omega\big((\nabla^\perp\Delta^{-2}\partial_x(\theta^{\mathrm{app}}+\theta^{\mathrm{int}}))\cdot\nabla\Delta^4\theta^{\mathrm{int}}\big)\Delta^4\theta^{\mathrm{int}} = \frac{1}{2}\int_\Omega \nabla\cdot\big(\nabla^\perp\Delta^{-2}\partial_x(\theta^{\mathrm{app}}+\theta^{\mathrm{int}})\big)|\Delta^4\theta^{\mathrm{int}}|^2 = 0.$$

Therefore, setting

$$\begin{aligned} S_\perp^2 &= -\nabla^\perp\Delta^{-2}\partial_x(\theta^{\mathrm{app}}+\theta^{\mathrm{int}})\cdot\nabla\Delta^4\theta^{\mathrm{int}} - \partial_z^8\overline{\nabla^\perp\Delta^{-2}\partial_x(\theta^{\mathrm{app}}+\theta^{\mathrm{int}})\cdot\theta^{\mathrm{int}}},\\ S_\parallel^2 &= \Delta^4 Q(\theta^{\mathrm{int}}, \theta^{\mathrm{app}}) - [\Delta^4, \nabla^\perp\Delta^{-2}\partial_x(\theta^{\mathrm{app}}+\theta^{\mathrm{int}})\cdot\nabla]\theta^{\mathrm{int}}, \end{aligned}$$

we obtain the desired $H^8$ estimates.

We now need to estimate the time derivative of $\partial_x^4 S^1_{\mathrm{rem}}$ in $L^2$. Note that the definition of time $T_2$ (see (4-31)) ensures that

$$\|\partial_t\partial_x^4\theta^{\mathrm{int}}(t)\|_{L^2} \lesssim B(1+t)^{-3} \quad \forall t\in[0, T_2].$$

Setting $\psi^{\mathrm{int}} = \Delta^{-2}\partial_x\theta^{\mathrm{int}}$, it follows that

$$\|\partial_t\partial_x^3\psi^{\mathrm{int}}\|_{H^4} \lesssim B(1+t)^{-3} \quad \forall t\in[0, T_2].$$

From there, differentiating with respect to time $\partial_x^4 S^1_{\mathrm{rem}}$, we obtain the desired estimate in $L^2$. The only problematic term is $\partial_t\partial_x^5\psi^{\mathrm{int}}\partial_z\theta^{\mathrm{app}}$, which we decompose as

$$\partial_t\partial_x^5\psi^{\mathrm{int}}\partial_z\theta^{\mathrm{app}}\chi(z) + \partial_t\partial_x^5\psi^{\mathrm{int}}\partial_z\theta^{\mathrm{app}}(1-\chi(z)),$$

with $\chi\in C_c^\infty(\mathbb{R})$ such that $\chi\equiv 1$ in a neighborhood of zero and $\chi(z)=0$ for $|z|\geq\frac{1}{2}$. Let us consider the first term. Recalling that $\psi^{\mathrm{int}}(z{=}0)=0$, we write, using the Hardy inequality,

$$\begin{aligned} \|\partial_t\partial_x^5\psi^{\mathrm{int}}\partial_z\theta^{\mathrm{app}}\chi(z)\|_{L^2} &\leq \left\|\frac{1}{z}\partial_t\partial_x^5\psi^{\mathrm{int}}\right\|_{L^2}\|z\partial_z\theta^{\mathrm{app}}\chi(z)\|_{L^\infty}\\ &\lesssim \|\partial_t\partial_x^5\partial_z\psi^{\mathrm{int}}\|_{L^2}\|z\partial_z\theta^{\mathrm{app}}\chi(z)\|_{L^\infty}\\ &\lesssim B(1+t)^{-3}\times B(1+t)^{-1} \lesssim B^2(1+t)^{-4}. \end{aligned}$$

The term involving $(1-\chi(z))$ is treated similarly, exchanging the roles of $z=0$ and $z=1$. □

***Conclusion.*** We apply the operator $\Delta^4$ to (4-32). We recall that by construction, $\Delta^2\theta^{\text{int}} = \partial_{\boldsymbol{n}}\Delta^2\theta^{\text{int}} = 0$ on $\partial\Omega$. We obtain

$$\partial_t \Delta^4\theta^{\text{int}} = (1-G)\partial_x\Delta^2\theta^{\text{int}} + \Delta^4 S^1_{\text{rem}} - [\Delta^4, G]\partial_x\psi^{\text{int}}.$$

Let us now check that the assumptions of Lemma 2.4 are satisfied. The decay assumptions on $\Delta^4 S^1_{\text{rem}}$ follow from Lemma 4.14. Therefore it suffices to check that the decay of the commutator term satisfies the desired bounds. Using (2-16) together with the bounds on $G$ (see Lemma 4.3), we have, for all $t \in (0, T_2)$,

$$\begin{aligned}\|[\Delta^4, G]\partial_x\psi^{\text{int}}\|_{L^2} &\lesssim \|G\|_{W^{1,\infty}}\|\partial_x\psi^{\text{int}}\|_{H^7} + \|G\|_{H^8}\|\partial_x\psi^{\text{int}}\|_\infty \\ &\lesssim B^2(1+t)^{-\frac{5}{4}} + B^2(1+t)^{\frac{1}{2}}(1+t)^{-\frac{11}{4}} \lesssim B^2(1+t)^{-\frac{5}{4}}.\end{aligned}$$

Therefore, according to Lemma 2.4, there exists a universal constant $C_2$ such that, for all $t \in (0, T_2)$, setting $B = (3+2C_1)\overline{C}\|\theta_0\|_{H^{14}}$ (see (4-31)),

$$\|\Delta^4\theta^{\text{int}}(t)\|_{L^2} \le C_2(\|\theta^{\text{int}}(t=0)\|_{H^8} + B^2).$$

From there, we apply Proposition 2.6 twice (first to $\partial_x^2\Delta^2\theta^{\text{int}}$ and then to $\partial_x^4\theta^{\text{int}}$), and we obtain, up to a change in the constant $C_2$, for all $t \in [0, T_2]$

$$\begin{aligned}\|\partial_x^2\Delta^2\theta^{\text{int}}(t)\|_{L^2} &\le C_2(\|\theta^{\text{int}}(t=0)\|_{H^8} + B^2)(1+t)^{-1}, \\ \|\partial_x^4\theta^{\text{int}}(t)\|_{L^2} &\le C_2(\|\theta^{\text{int}}(t=0)\|_{H^8} + B^2)(1+t)^{-2}.\end{aligned}$$

There remains to bound $\partial_t\partial_x^4\theta^{\text{int}}$ and $\partial_x^5\psi^{\text{int}}$ in $L^2$. To that end, we differentiate (4-32) with respect to time, and we obtain

$$\partial_t\partial_t\partial_x^4\theta^{\text{int}} = (1-G)\partial_x^5\partial_t\psi^{\text{int}} + \partial_t\partial_x^4 S^1_{\text{rem}} - \partial_t G\partial_x^5\psi^{\text{int}}.$$

The source term $\partial_t\partial_x^4 S^1_{\text{rem}}$ is evaluated in Lemma 4.14. As for the commutator term, we have

$$\|\partial_t G\partial_x^5\psi^{\text{int}}\|_{L^2} \le \|\partial_t G\|_{L^\infty}\|\partial_x^5\psi^{\text{int}}\|_{L^2} \lesssim B^3(1+t)^{-3+\frac{1}{2}-3} \lesssim B^3(1+t)^{-4}.$$

Using Proposition 2.6, we find that, for any $t \in (0, T_2)$,

$$\|\partial_t\partial_x^4\theta^{\text{int}}\|_{L^2} \le C_2(\|\theta^{\text{int}}(t=0)\|_{H^8} + B^2)\frac{1}{(1+t)^3}.$$

Using (4-32),

$$\|\partial_x^5\psi^{\text{int}}(t)\|_{L^2} \le C_2(\|\theta^{\text{int}}(t=0)\|_{H^8} + B^2)\frac{1}{(1+t)^3} \quad \forall t \in (0, T_2).$$

Grouping these estimates with the ones on $\sigma^{\text{NL}}_{\text{lift}}$ from Lemma 4.13, we infer that, up to a change of the constant $C_2$, for any $t \in (0, T_2)$

$$\begin{aligned}(1+t)^2\|\partial_x^4\theta^{\text{rem}}(t)\|_{L^2} + \|\Delta^4\theta^{\text{rem}}(t)\|_{L^2} &\le C_2(\|\theta_0\|_{H^8} + B^2), \\ (1+t)^3\big(\|\partial_t\partial_x^4\theta^{\text{rem}}(t)\|_{L^2} + \|\partial_x^5\psi^{\text{rem}}(t)\|_{L^2}\big) &\le C_2(\|\theta_0\|_{H^8} + B^2).\end{aligned}$$

We now recall that $B_3 = (3\overline{C} + 5C_1)\|\theta_0\|_{H^{14}}$ for some constant $\overline{C}$ that remains to be chosen. We want to pick $\overline{C}$ so that

$$C_2\big(\|\theta_0\|_{H^8} + (3\overline{C}+5C_1)^2\|\theta_0\|^2_{H^{14}}\big) \le (\overline{C} + C_1)\|\theta_0\|_{H^{14}}.$$

It is sufficient to take $\overline{C}$ such that $2C_2 \le (\overline{C} + C_1)$, and $\|\theta_0\|_{H^{14}}$ sufficiently small. We then infer that the bounds within (4-31) are satisfied with $B_2$ replaced by $B_2/2$. It follows that $T_2 = T_1$. From there, recalling the estimates on $\theta^{\mathrm{BL}}$, we deduce that there exists a universal constant $C_3$ such that, for all $t \in (0, T_1)$, for all $k \in \{4, \dots, 8\}$,

$$\begin{aligned}\|\partial_x^k \theta'\|_{L^2} &\le \|\partial_x^k \theta^{\mathrm{BL}}\|_{L^2} + \|\partial_x^k \theta^{\mathrm{rem}}\|_{L^2} \\ &\le C_3(\|\theta_0\|_{H^{14}} + B^2)((1+t)^{-\frac{9}{8}} + (1+t)^{\frac{k-8}{2}}).\end{aligned}$$

Similar estimates hold for $\partial_z^k \theta'$ and $\partial_x^5 \psi$ in $L^2$. Hence we further choose the constant $\overline{C}$ so that

$$2C_3(\|\theta_0\|_{H^{14}} + B^2) \le \overline{C} \|\theta_0\|_{H^{14}}$$

provided $\|\theta_0\|_{H^{14}}$ is sufficiently small. We conclude that $T_1 = +\infty$. Theorem 1.3 follows.

## Appendix A: Well-posedness of the Stokes-transport equation in Sobolev spaces

The aim of this section is to prove the well-posedness of the Stokes-transport on the domain of interest of the present paper, namely $\Omega = \mathbb{T} \times (0, 1)$. The proof is also valid on any regular enough bounded domain of $\mathbb{R}^d$ with $d = 2$ or 3.

**Theorem A.1.** *Let $\Omega$ satisfy either*

(1) $\Omega = \mathbb{T} \times (0, 1)$ *or*

(2) $\Omega$ *is a simply connected compact subdomain of* $\mathbb{R}^d$, $d = 2, 3$, *regular enough.*

*Let* $m \ge 3$, $\rho_0 \in H^m(\Omega)$ (*and* $\Omega$ *of regularity* $C^{m+2}$). *The system*

$$\begin{cases} \partial_t \rho + \boldsymbol{u} \cdot \nabla \rho = 0, \\ -\Delta \boldsymbol{u} + \nabla p = -\rho \boldsymbol{e}_z, \\ \operatorname{div} \boldsymbol{u} = 0, \\ \boldsymbol{u}|_{\partial\Omega} = \boldsymbol{0}, \\ \rho|_{t=0} = 0 \end{cases} \tag{A-1}$$

*has a unique global solution for the present regularity*

$$(\rho, \boldsymbol{u}) \in C(\mathbb{R}_+; H^m(\Omega)) \times C(\mathbb{R}_+; H^{m+2}(\Omega)).$$

*Moreover*, *the solution obeys the energy estimate*

$$\|\rho(t)\|_{H^m} \le \|\rho_0\|_{H^m} \exp\Big(C \int_0^t \|\nabla \boldsymbol{u}(s)\|_{L^\infty} + \|\nabla \rho(s)\|_{L^\infty} \, \mathrm{d}s\Big). \tag{A-2}$$

The proof of this result follows rather classical techniques. It also relies on a previous work of one of the authors [Leblond 2022], including in particular the well-posedness in a weak sense of the system (A-1).

**Remark A.2.** The well-posedness in the weak sense of (A-1) in $\mathbb{T} \times (0, 1)$ is a direct consequence its the well-posedness in $\mathbb{R} \times (0, 1)$ stated in [Leblond 2022, Theorem 1.2]. In this latter unbounded domain, the Poiseuille flows are avoided thanks to a zero flux condition on the velocity field. In the periodic case, this condition is no longer required as the periodicity of the solution prevents the existence of Poiseuille flows.

***A priori estimate.*** Formally, the energy estimate for any derivative of order $m$ can be written as

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\partial^m \rho\|_{L^2}^2 = -\int_\Omega [\partial^m, \boldsymbol{u}\cdot\nabla]\rho\partial^m\rho,$$

due to the divergence-free condition satisfied by $\boldsymbol{u}$. We apply the tame estimate (2-16), together with the continuous Sobolev embedding of $H^m(\Omega)$ in $L^\infty(\Omega)$ and the Stokes equation regularization estimate $\|\boldsymbol{u}\|_{H^m} \lesssim \|\rho\|_{H^{m-2}}$, to get

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\partial^m \rho\|_{L^2}^2 \lesssim (\|\nabla\boldsymbol{u}\|_{L^\infty} + \|\nabla\rho\|_{L^\infty})\|\rho\|_{H^m}^2.$$

One therefore obtains the same inequality with the complete $H^m$ norm on the left-hand side, and the estimate (A-2) follows. This energy estimate tells us that $\rho$ remains in $H^m(\Omega)$ as long as $\|\nabla\boldsymbol{u}\|_{L^\infty}$ and $\|\nabla\rho\|_{L^\infty}$ are integrable in time. Regarding the properties we know from [Leblond 2022] about the solutions of this equation it is enough to prove that the solution exists globally and is unique. Let us recall from [Galdi 2011, Theorem IV.6.1] and [Leblond 2022, Section 2.1] that the source term and the solution of the Stokes equation satisfy for all times

$$\|\boldsymbol{u}\|_{H^m} \lesssim \|\rho\|_{H^{m-2}}, \qquad \|\boldsymbol{u}\|_{W^{1,\infty}} \lesssim \|\rho\|_{L^\infty}. \tag{A-3}$$

Also, the uniform norm of $\rho$ is constant since $\rho$ is transported by an incompressible vector field. We also observe

$$\|\nabla\rho\|_{L^\infty} \le \|\nabla\rho_0\|_{L^\infty}\exp\Big(C\int_0^t \|\nabla\boldsymbol{u}(s)\|_{L^\infty}\,\mathrm{d}s\Big) \le \|\nabla\rho_0\|_{L^\infty}\exp(C\|\rho_0\|_{L^\infty}t). \tag{A-4}$$

Putting these considerations together leads to

$$\|\rho\|_{H^m} \le \|\rho_0\|_{H^m}\exp\bigg(C\|\rho_0\|_{L^\infty}t + \frac{\|\nabla\rho_0\|_{L^\infty}}{\|\rho_0\|_{L^\infty}}(\exp(C\|\rho_0\|_{L^\infty}t) - 1)\bigg).$$

This suggests that if $\rho_0 \in H^m \cap W^{1,\infty}$, the solution exists globally in time in $H^m$. In particular, if $m$ is large enough so that $H^m(\Omega) \hookrightarrow W^{1,\infty}(\Omega)$, the Stokes-transport system is well-posed in $H^m$.

***Proof.*** An iterative scheme allows us to formalize the previous considerations. Let $\rho^0 : t \mapsto \rho_0$, which belongs to $C(\mathbb{R}_+, H^m(\Omega))$. Now if $\rho^N$ belongs to $C(\mathbb{R}_+, H^m(\Omega))$, which is true for $N = 0$, we know that the Stokes system

$$\begin{cases} -\Delta\boldsymbol{u}^N + \nabla p^N = -\rho^N\boldsymbol{e}_z, \\ \operatorname{div}\boldsymbol{u}^N = 0, \\ \boldsymbol{u}^N|_{\partial\Omega} = \boldsymbol{0} \end{cases}$$

admits for any time a unique solution $\boldsymbol{u}^N(t) \in H^{m+2}(\Omega)$ obeying inequalities (A-3). By linearity of the problem, $\boldsymbol{u}^N$ in $H^{m+2}(\Omega)$ inherits the continuity of $\rho^N$ in $H^m(\Omega)$. Then since $\boldsymbol{u}^N$ belongs in particular to $C(\mathbb{R}_+, H^{m+2}(\Omega))$, the transport equation

$$\begin{cases} \partial_t\rho^{N+1} + \boldsymbol{u}^N\cdot\nabla\rho^{N+1} = 0, \\ \rho^{N+1}|_{t=0} = \rho_0 \end{cases}$$

has a unique strong solution $\rho^{N+1} \in C(\mathbb{R}_+, H^m(\Omega))$. This concludes the definition of the sequences $(\rho^N)_N$ and $(\boldsymbol{u}^N)_N$. We thereafter show that for any $T > 0$ the sequence $(\rho^N)_N$ is bounded in $L^\infty((0,T), H^m(\Omega))$

and equicontinuous in $C((0,T),H^{m-1}(\Omega))$, so it converges in $C((0,T),H^{m-1}(\Omega))$ to a solution of the original system up to an extraction. Since this is true for any $T>0$ and by uniqueness of the weak solution ensured by [Leblond 2022, Theorems 1.1 and 1.2], we get the well-posedness of the system and the proposition is proven.

***Boundedness.*** Let us show that we have, for any $N\in\mathbb{N}$,

$$\|\rho^N\|_{H^m}\le\|\rho_0\|_{H^m}\exp\biggl(C\|\rho_0\|_{L^\infty}t+\frac{\|\nabla\rho_0\|_{L^\infty}}{\|\rho_0\|_{L^\infty}}(\exp(C\|\rho_0\|_{L^\infty}t)-1)\biggr)=:B_{\rho_0}(t). \tag{A-5}$$

This inequality is immediately satisfied for $N=0$ since $\rho^0$ is constant in time and equal to $\rho_0$. Let $N\in\mathbb{N}$ such that (A-5) is satisfied. Then the tame estimate (2-15) provides here

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\rho^{N+1}(t)\|_{H^m}^2\lesssim\|\nabla\boldsymbol{u}^N\|_{L^\infty}\|\rho^{N+1}\|_{H^m}^2+\|\nabla\rho^{N+1}\|_{L^\infty}\|\boldsymbol{u}^N\|_{H^m}\|\rho^{N+1}\|_{H^m}.$$

The considerations (A-3) and (A-4) applied to $\rho^N$, $\rho^{N+1}$ and $\boldsymbol{u}^N$ lead here to

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\rho^{N+1}\|_{H^m}\lesssim\|\rho_0\|_{L^\infty}\|\rho^{N+1}\|_{H^m}+\|\nabla\rho_0\|_{L^\infty}\exp(C\|\rho_0\|_{L^\infty}t)\|\rho^{N+1}\|_{H^m}.$$

From here, we use the Grönwall lemma to estimate $\|\rho^{N+1}\|_{H^m}$,

$$\|\rho^{N+1}\|_{H^m}\le\exp(C\|\rho_0\|_{L^\infty}t)\biggl(\|\rho_0\|_{H^m}+C\|\nabla\rho_0\|_{L^\infty}\int_0^t\|\rho^N(s)\|_{H^m}\,\mathrm{d}s\biggr). \tag{A-6}$$

Then, according to the assumption on $\rho^N$, we observe that

$$\begin{aligned}
C\|\nabla\rho_0\|_{L^\infty}&\int_0^t\|\rho^N(s)\|_{H^m}\,\mathrm{d}s\\
&\le C\|\rho_0\|_{H^m}\|\nabla\rho_0\|_{L^\infty}\int_0^t\exp\biggl(C\|\rho_0\|_{L^\infty}s+\frac{\|\nabla\rho_0\|_{L^\infty}}{\|\rho_0\|_{L^\infty}}(\exp(C\|\rho_0\|_{L^\infty}s)-1)\biggr)\mathrm{d}s\\
&\le C\|\rho_0\|_{H^m}\|\nabla\rho_0\|_{L^\infty}\int_0^{\exp(C\|\rho_0\|_{L^\infty}t)-1}\exp\biggl(\frac{\|\nabla\rho_0\|_{L^\infty}}{\|\rho_0\|_{L^\infty}}r\biggr)\frac{\mathrm{d}r}{C\|\rho_0\|_{L^\infty}}\\
&=\|\rho_0\|_{H^m}\biggl(\exp\biggl(\frac{\|\nabla\rho_0\|_{L^\infty}}{\|\rho_0\|_{L^\infty}}(\exp(C\|\rho_0\|_{L^\infty}t)-1)\biggr)-1\biggr).
\end{aligned}$$

The latter bound substituted in (A-6) yields exactly the result (A-5). Therefore, for any $T>0$ the sequence $(\rho^N)_N$ is uniformly bounded in $L^\infty(0,T,H^m(\Omega))$.

***Equicontinuity.*** We find a uniform bound on $(\partial_t\rho^N)_N$ in $H^{m-1}(\Omega)$ to show the equicontinuity of the sequence in $C((0,T),H^{m-1}(\Omega))$. This bound, uniform in $N\in\mathbb{N}$ and $t\in[0,T]$, is obtained thanks to the tame estimate, the bounds (A-3) and the uniform bound (A-5) on $\rho^N$,

$$\begin{aligned}
\|\partial_t\rho^N\|_{H^{m-1}}&=\|\boldsymbol{u}^{N-1}\cdot\nabla\rho^N\|_{H^{m-1}}\\
&\lesssim\|\boldsymbol{u}^{N-1}\|_{L^\infty}\|\nabla\rho^N\|_{H^{m-1}}+\|\nabla\rho^N\|_{L^\infty}\|\boldsymbol{u}^{N-1}\|_{H^{m-1}}\\
&\lesssim\|\rho_0\|_{L^\infty}\|\rho^N\|_{H^m}+\|\rho^{N-1}\|_{H^m}\|\nabla\rho_0\|_\infty\exp(C\|\rho_0\|_\infty t)\\
&\lesssim\|\rho_0\|_{W^{1,\infty}}\|\rho_0\|_{H^m}\exp\biggl(C\|\rho_0\|_{L^\infty}t+\frac{\|\nabla\rho_0\|_{L^\infty}}{\|\rho_0\|_{L^\infty}}(\exp(C\|\rho_0\|_{L^\infty}t)-1)\biggr).
\end{aligned}$$

***Regularity.*** Let us show that the limit $\rho$ belongs to $L^\infty((0,T), H^m(\Omega))$. For any $t \in [0,T]$, $(\rho^N(t))_N$ is uniformly bounded in $H^m(\Omega)$ with respect to $N$ and $t$. Hence according to Banach–Alaoglu theorem, for any $t$ the sequence is weakly compact in $H^m(\Omega)$. Thus up to an extraction, $\rho^N(t)$ converges weakly toward a $\bar\rho(t) \in H^m(\Omega)$, and this limit satisfies

$$\|\bar\rho(t)\|_{H^m} \le \liminf_N \|\rho^N(t)\|_{H^m},$$

where the right-hand side is uniformly bounded thanks to (A-5). As $\rho^N$ already converges weakly in $H^m(\Omega)$, we can identify $\bar\rho$ and $\rho$, which then belongs to $L^\infty((0,T), H^m(\Omega))$. Finally, to reach the regularity $C((0,T), H^m(\Omega))$, Lemma II.5.6 in [Boyer and Fabrie 2013] tells us that since in particular $\rho \in L^\infty((0,T), H^m(\Omega)) \cap C^0_{\mathrm{w}}((0,T), H^{m-1}(\Omega))$ then $\rho \in C^0_{\mathrm{w}}((0,T), H^m(\Omega))$. Hence it is enough to show that $t \mapsto \|\rho(t)\|_{H^m}$ is continuous to prove the strong continuity of $\rho$ in $H^m(\Omega)$. By weak continuity,

$$\|\rho_0\|_{H^m} \le \liminf_{t\searrow 0} \|\rho(t)\|_{H^m}.$$

Also by weak convergence

$$\|\rho(t)\|_{H^m} \le \liminf_{N\to\infty} \|\rho^N(t)\|_{H^m} \le B_{\rho_0}(t),$$

which proves by clamping that $t \mapsto \|\rho(t)\|_{H^m}$ is continuous at $t=0$. This can be performed for any $t \in [0,T]$, hence the continuity. Finally, $\boldsymbol{u} \in C^0((0,T), H^{m+2}(\Omega))$ by (A-3) and linearity of the Stokes equation.

## Appendix B: About the bilaplacian equation

We use throughout the paper the following classical regularity result:

**Lemma B.1** (regularity). *Let $f \in H^m(\Omega)$, $m \ge -2$. The problem*

$$\Delta^2\psi = f, \quad \psi|_{\partial\Omega} = \partial_{\boldsymbol{n}}\psi|_{\partial\Omega} = 0,$$

*admits a unique strong solution $\psi \in H^2_0 \cap H^{m+4}(\Omega)$ such that*

$$\|\psi\|_{H^{m+4}} \lesssim \|f\|_{H^m}.$$

The eigenvalues and eigenfunctions of the bilaplacian in a channel can be semiexplicitly computed (see [Leblond 2023] for the details):

**Lemma B.2** (spectrum of the bilaplacian). *The eigenvalues of the operator $\Delta^2$ on $H^2_0$ in $\mathbb{T}\times(-1,1)$ are the union, for all $k \in \mathbb{Z}$, of strictly increasing sequences $(\lambda_{n,k})_{n\in\mathbb{N}}$ such that*

$$\lambda_{n,k} \simeq (n^2+k^2)^2,$$

*with associated (unnormalized) eigenfunctions*

$$b_{n,k} = e^{ikx}\begin{cases} \cos(\omega_{n,k}z) - \dfrac{\cos(\omega_{n,k})}{\cosh(r_{n,k})}\cosh(r_{n,k}z), & n \in 2\mathbb{N}, \\ \sin(\omega_{n,k}z) - \dfrac{\sin(\omega_{n,k})}{\sinh(r_{n,k})}\sinh(r_{n,k}z), & n \in 2\mathbb{N}+1, \end{cases}$$

*with $\omega_{n,k} = (k^2 - \lambda_{n,k}^{1/2})^{1/2}$ and $r_{n,k} = (k^2 + \lambda_{n,k}^{1/2})^{1/2}$. Note that to simplify the calculations, the domain was chosen to be $\mathbb{T}\times(-1,1)$ and not $\Omega = \mathbb{T}\times(0,1)$.*

## Appendix C: Proof of Lemma 3.2

The proof of the lemma relies on energy estimates in weighted Sobolev spaces, with weights that grow like $\exp(cZ^{4/5})$ for $Z \gg 1$. Unfortunately, we have not been able to treat all four cases for the boundary conditions simultaneously, but we will treat (i) and (iii) (resp. (ii) and (iv)) together. Note that when (3-4) is multiplied (formally) by $\Psi w$ or by $-\partial_Z \Psi w$, where $w \in C^\infty([0, +\infty))$ is an arbitrary weight function, there are many commutator terms when we integrate by parts the fifth-order derivative. The main idea is that if the weight is adequately chosen, all these commutators can be absorbed in the main-order terms, which will be designed to have a positive sign. Hence we start with the following result, which will allow us to control the commutators:

**Lemma C.1.** *Let $\Psi \in C_c^\infty([0, +\infty))$ such that $\Psi(0) = 0$, and let $r \in (0, 1)$.*

(1) *Let $W \in C^\infty([0, +\infty))$ such that $W(Z) = \exp(Z^{4/5})$ for $Z \geq 1$, and $W \geq 1$, $\partial_z^2 W \geq 0$, $W \equiv 1$ in a neighborhood of zero.*

*Then, for $k \in \{1, 2\}$, there exists a constant $C_k$, independent of $r$, such that, for all $r \in (0, 1)$,*

$$\left| \int_0^\infty |\partial_Z^k \Psi(Z)|^2 \frac{|\partial_Z^{3-k} W(rZ)|^2}{W(rZ)} \, \mathrm{d}Z \right| \leq C_k r^{-\frac{2}{3}(3-k)} \left[ \int_0^\infty |\partial_Z^3 \Psi(Z)|^2 W(rZ) \, \mathrm{d}Z + \int_0^\infty \Psi^2(Z) \left( \frac{r \partial_Z W(rZ)}{Z} + \frac{W(rZ)}{Z^2} \right) \mathrm{d}Z \right].$$

(2) *Let $\Phi : Z \mapsto \Psi(Z)/Z$. Then, for $k \in \{1, 2, 3, 4\}$, for all $c > 0$, for $r$ sufficiently small,*

$$\int_0^\infty \mathbf{1}_{rZ>c} (rZ)^{\frac{2+2k}{5}} \exp((rZ)^{\frac{4}{5}}) \partial_Z^k \Phi(Z)^2 \, \mathrm{d}Z \lesssim_c r^{\frac{2(k-1)}{3}} \left[ \int_0^\infty \partial_Z^4 \Psi(Z)^2 \exp((rZ)^{\frac{4}{5}}) \, \mathrm{d}Z + \int_0^\infty \partial_Z \Phi(Z)^2 (rZ)^{\frac{4}{5}} \exp((rZ)^{\frac{4}{5}}) \, \mathrm{d}Z \right].$$

*Proof.* • For $k = 0, \dots, 3$, let us consider weights $\omega_k \in W^{1,\infty}_{\mathrm{loc}}((0, +\infty))$ such that

$$\forall k \in \{0, \dots, 3\}, \quad \forall Z \geq 1, \quad \omega_k(Z) = e^{-1} Z^{-\frac{2}{5}(3-k)} \exp(Z^{\frac{4}{5}}),$$
$$\forall Z \in (0, 1), \quad \omega_1(Z) = \omega_3(Z) = 1, \quad \omega_0(Z) = Z^{-2}, \quad \omega_2(Z) = Z^2.$$

Note that the weights $\omega_k$ satisfy the following assumptions:

- For $k \in \{1, 2\}$, $\omega_k \leq \sqrt{\omega_{k-1}\omega_{k+1}}$.
- For $k \in \{1, 2\}$, $|\partial_Z \omega_k| \leq C_k \sqrt{\omega_k \omega_{k-1}}$ for some constant $C_k$.
- $\omega_2(0) = 0$.
- $\omega_3 \leq CW$, $\omega_0(Z) \leq C(Z^{-2} W(Z) + Z^{-1} \partial_Z W(Z))$.

Let us now introduce, for $k = 0, \dots, 3$,

$$I_k := \int_0^\infty |\partial_Z^k \Psi(Z)|^2 \omega_k(rZ) \, \mathrm{d}Z.$$

Then by the definition of $W$, $\omega_0$, $\omega_3$, there exists a constant $C$ (independent of $r > 0$) such that

$$r^2 I_0 + I_3 \le C \int_0^\infty |\partial_Z^3 \Psi(Z)|^2 W(rZ)\,\mathrm{d}Z + \int_0^\infty \Psi^2(Z)\left(\frac{r\partial_Z W(rZ)}{Z} + \frac{W(rZ)}{Z^2}\right)\mathrm{d}Z.$$

Let us set $E = r^2 I_0 + I_3$. For $k = 1, 2$, integrating by parts and using the conditions $\Psi(0) = \omega_2(0) = 0$, we have

$$I_k = -\int_0^\infty \partial_Z^{k-1}\Psi(Z)\partial_Z^{k+1}\Psi(Z)\omega_k(rZ)\,\mathrm{d}Z - r\int_0^\infty \partial_Z^{k-1}\Psi(Z)\partial_Z^k\Psi(Z)\partial_Z\omega_k(rZ)\,\mathrm{d}Z.$$

Using the properties of $\omega_k$, we deduce that there exist constants $C_k$ such that

$$I_k \le C_k\big(\sqrt{I_{k-1}I_{k+1}} + r\sqrt{I_k I_{k-1}}\big).$$

Since $r^2 I_0 \le E$, we deduce first that $I_1 \lesssim E + r^{-1}\sqrt{I_2 E}$, and plugging this inequality into the bound on $I_2$, we find, since $r \in (0, 1)$,

$$I_1 \lesssim r^{-\frac{4}{3}}E, \quad I_2 \lesssim r^{-\frac{2}{3}}E.$$

The first inequality from Lemma C.1 then follows easily by noticing that

$$\frac{(\partial_Z W)^2}{W} \lesssim \omega_2, \quad \frac{(\partial_Z^2 W)^2}{W} \lesssim \omega_1.$$

• Let us now set, for $k \in \{1, \dots, 4\}$,

$$J_k := \int_0^\infty \partial_Z^k \Phi(Z)^2 \zeta_k(rZ)\,\mathrm{d}Z,$$

where the weights $\zeta_k \in W^{1,\infty}_{\mathrm{loc}}(\mathbb{R})$ satisfy $\zeta_k \equiv 0$ in a neighborhood of zero, $\zeta_k(Z) = Z^{(2+2k)/5}\exp(Z^{4/5})$ for $Z$ large enough, and $\zeta_k \lesssim \sqrt{\zeta_{k-1}\zeta_{k+1}}$, $\partial_Z\zeta_k \lesssim \sqrt{\zeta_k\zeta_{k-1}}$. Let $F := r^{-2}J_4 + J_1$. As above, for $k \in \{2, 3\}$, we have

$$J_k \le C_k\Big(\sqrt{J_{k-1}J_k} + r\sqrt{J_k J_{k-1}}\Big).$$

From there, we infer that, for $k \in \{1, \dots 4\}$, $J_k \lesssim r^{2(k-1)/3}F$.

Now, since $\Psi = Z\Phi$, we have $\partial_Z^4\Psi = Z\partial_Z^4\Phi + 4\partial_Z^3\Phi$. It follows that

$$\begin{aligned}
&\int_0^\infty (\partial_Z^4\Psi(Z))^2\exp((rZ)^{4/5})\,\mathrm{d}Z \\
&\quad= \int_0^\infty (Z\partial_Z^4\Phi + 4\partial_Z^3\Phi)^2\exp((rZ)^{4/5})\,\mathrm{d}Z \\
&\quad= \int_0^\infty Z^2(\partial_Z^4\Phi)^2\exp((rZ)^{4/5})\,\mathrm{d}Z + \int_0^\infty 16(\partial_Z^3\Phi)^2\exp((rZ)^{4/5})\,\mathrm{d}Z \\
&\qquad\qquad - 4\int_0^\infty (\partial_Z^3\Phi)^2\frac{\partial}{\partial Z}(Z\exp((rZ)^{4/5}))\,\mathrm{d}Z \\
&\quad= \int_0^\infty Z^2(\partial_Z^4\Phi)^2\exp((rZ)^{4/5})\,\mathrm{d}Z + 16\int_0^\infty (\partial_Z^3\Phi)^2\exp((rZ)^{4/5})\,\mathrm{d}Z \\
&\qquad\qquad - 4\int_0^\infty \big(1 + \tfrac{4}{5}(rZ)^{\frac{4}{5}}\big)(\partial_Z^3\Phi)^2\exp((rZ)^{4/5})\,\mathrm{d}Z.
\end{aligned}$$

We split the last integral into two parts, for $rZ \le 1$ and $rZ \ge 1$. When $rZ \le 1$,

$$4\int_0^{r^{-1}} \left(1+\tfrac{4}{5}(rZ)^{\frac{4}{5}}\right)(\partial_Z^3\Phi)^2 \exp((rZ)^{4/5})\,\mathrm{d}Z \le 8\int_0^{r^{-1}} (\partial_Z^3\Phi)^2 \exp((rZ)^{4/5})\,\mathrm{d}Z.$$

And for $rZ \ge 1$, for a suitable choice of $\zeta_3$

$$\int_{r^{-1}}^{\infty} \left(1+\tfrac{4}{5}(rZ)^{\frac{4}{5}}\right)(\partial_Z^3\Phi)^2 \exp((rZ)^{4/5})\,\mathrm{d}Z \lesssim J_3 \lesssim r^{\frac{4}{3}} F.$$

We infer that

$$\int_0^{\infty} (\partial_Z^4\Psi(Z))^2 \exp((rZ)^{4/5})\,\mathrm{d}Z \ge C^{-1}r^{-2}J_4 - Cr^{\frac{4}{3}}F,$$

and therefore, for $r$ sufficiently small,

$$F \lesssim \int_0^{\infty} (\partial_Z^4\Psi(Z))^2 \exp((rZ)^{4/5})\,\mathrm{d}Z + J_1.$$

The result follows. □

We now turn towards the proof of Lemma 3.2. In both cases, we start with a formal a priori estimate, from which we deduce an appropriate notion of variational solution in a suitable Hilbert space. Existence and uniqueness then follow in a straightforward manner from the Lax–Milgram lemma.

**First case:** conditions (ii) and (iv). As explained above, we start with a formal a priori estimate. Let $w \in C^\infty(\mathbb{R}_+)$ be an arbitrary weight function, and multiply (3-4) by $\partial_Z(\Psi(Z)w(Z))/Z$. On the one hand,

$$\int_0^{\infty} \partial_Z^5\Psi(Z)\partial_Z(\Psi w)(Z)\,\mathrm{d}Z = \int_0^{\infty} \partial_Z^3\Psi(Z)\partial_Z^3(\Psi w)(Z)\,\mathrm{d}Z - \partial_Z^4\Psi(0)\partial_Z(\Psi w)(0) + \partial_Z^3\Psi(0)\partial_Z^2(\Psi w)(0).$$

Note that the two boundary terms vanish in cases (ii) and (iv). We obtain

$$\int_0^{\infty} \partial_Z^5\Psi(Z)\partial_Z(\Psi w)(Z)\,\mathrm{d}Z = \int_0^{\infty} (\partial_Z^3\Psi(Z))^2 w(Z) + \sum_{k=1}^{3}\binom{3}{k}\int_0^{\infty} \partial_Z^3\Psi(Z)\partial_Z^{3-k}\Psi(Z)\partial_Z^k w(Z)\,\mathrm{d}Z.$$

On the other hand, since $\Psi(0)=0$,

$$\begin{aligned}\int_0^{\infty} \Psi(Z)\partial_Z(\Psi w)(Z)\frac{\mathrm{d}Z}{Z} &= \int_0^{\infty} (\Psi w)(Z)\partial_Z(\Psi w)(Z)\frac{\mathrm{d}Z}{Zw(Z)}\\ &= -\frac{1}{2}\int_0^{\infty} (\Psi w)^2(Z)\frac{\mathrm{d}}{\mathrm{d}Z}\left(\frac{1}{Zw(Z)}\right)\mathrm{d}Z.\end{aligned}$$

Choosing $w$ such that $\partial_Z w \ge 0$, the right-hand side has a positive sign. We then choose $w(Z) = W(rZ)$ for some $W \in C^\infty(\mathbb{R}_+)$ such that $W(\xi) = \exp(\xi^{4/5})$ for $\xi \ge 1$, $W(\xi) = 1$ for $\xi$ in a neighborhood of zero, $\partial_\xi W \ge 0$, and $r > 0$ small enough. With this choice, the positive terms in the energy are bounded from below by

$$\int_0^{\infty} (\partial_Z^3\Psi(Z)^2 W(rZ)\,\mathrm{d}Z + \int_0^{\infty} \Psi^2(Z)\left(\frac{W(rZ)}{Z^2} + r\frac{\partial_Z W(rZ)}{Z}\right)\mathrm{d}Z.$$

Lemma C.1 then implies that there exists an explicit constant $\delta > 0$ such that, for $k = 1, 2, 3$,

$$\left|\int_0^\infty \partial_Z^3\Psi(Z)\partial_Z^{3-k}\Psi(Z)\partial_Z^k w(Z)\,\mathrm{d}Z\right| \\ \leq r^\delta\left[\int_0^\infty (\partial_Z^3\Psi(Z))^2 W(rZ)\,\mathrm{d}Z + \int_0^\infty \Psi^2(Z)\left(\frac{W(rZ)}{Z^2} + r\frac{\partial_Z W(rZ)}{Z}\right)\mathrm{d}Z\right].$$

Therefore, for $r > 0$ sufficiently small, we obtain

$$\int_0^\infty (\partial_Z^3\Psi(Z))^2 W(rZ)\,\mathrm{d}Z + \int_0^\infty \Psi^2(Z)\left(\frac{W(rZ)}{Z^2} + r\frac{\partial_Z W(rZ)}{Z}\right)\mathrm{d}Z \\ \lesssim \int_0^1 \frac{S(Z)^2}{Z^2}\,\mathrm{d}Z + \int_0^\infty S(Z)^2 W(rZ)\,\mathrm{d}Z.$$

This leads us to the following formulation: let

$$\mathcal{H} := \left\{\Psi \in H^3(\mathbb{R}_+) : \Psi(0) = 0,\ \int_0^\infty (\partial_Z^3\Psi(Z))^2 \exp((rZ)^{4/5})\,\mathrm{d}Z < +\infty, \right. \\ \left. \int_0^\infty \Psi(Z)^2(Z^{-2} + Z^{-\frac{1}{5}})\exp((rZ)^{4/5})\,\mathrm{d}Z < +\infty\right\},$$

and let

$$\mathcal{H}_0 := \{\Psi \in \mathcal{H} : \partial_Z\Psi(0) = \partial_Z^2\Psi(0) = 0\}.$$

We endow $\mathcal{H}$ and $\mathcal{H}_0$ with the norm

$$\|\Psi\|_{\mathcal{H}}^2 = \int_0^\infty (\partial_Z^3\Psi(Z))^2 W(rZ)\,\mathrm{d}Z + \int_0^\infty \Psi^2(Z)\left(\frac{W(rZ)}{Z^2} + r\frac{\partial_Z W(rZ)}{Z}\right)\mathrm{d}Z,$$

where $W$ is the previous weight. We say that $\Psi \in \mathcal{H}$ is a solution of (3-4)-(ii) (resp. $\Psi \in \mathcal{H}_0$ is a solution of (3-4)-(iv)) if and only if, for all $\Phi \in \mathcal{H}$ (resp. $\Phi \in \mathcal{H}_0$),

$$\int_0^\infty \partial_Z^3\Psi\partial_Z^3(\Phi W(r\,\cdot)) + \int_0^\infty \Psi(Z)\frac{\partial_Z(\Phi(Z)W(rZ))}{Z}\,\mathrm{d}Z = \int_0^\infty \frac{S(Z)}{Z}\partial_Z(\Phi(Z)W(rZ))\,\mathrm{d}Z.$$

Existence and uniqueness of solutions of (3-4)-(ii) (resp. of (3-4)-(iv)) in $\mathcal{H}$ (resp. $\mathcal{H}_0$) follow easily from the Lax–Milgram lemma. Using the equation, we then infer that

$$\int_0^\infty (\partial_Z^5\Psi(Z))^2 \exp((rZ)^{4/5})\,\mathrm{d}Z < +\infty.$$

The result follows.

**Second case:** conditions (i) and (iii). The estimates in the case of conditions (i) and (iii) are similar, but slightly less straightforward, since we shall need to combine two estimates.

We first multiply (3-4) by $-\partial_Z^3\Psi(Z)w_1(Z)/Z$, with a weight $w_1$ to be chosen later. We obtain on the one hand

$$-\int_0^\infty \partial_Z^5\Psi(Z)\partial_Z^3\Psi(Z)w_1(Z)\,\mathrm{d}Z = \int_0^\infty (\partial_Z^4\Psi(Z))^2 w_1(Z)\,\mathrm{d}Z + \int_0^\infty \partial_Z^4\Psi(Z)\partial_Z^3\Psi(Z)\partial_Z w_1(Z)\,\mathrm{d}Z.$$

Note that the boundary term $-\partial_Z^4\Psi(0)\partial_Z^3\Psi(0)w_1(0)$ vanishes in cases (i) and (iii). The first term gives a positive contribution to the energy, and the second one will be treated below with the help of Lemma C.1. On the other hand, we obtain for the zeroth order term, noticing that either $\partial_Z^2\Psi(0)=0$ (in case (iii)) or $(Z^{-1}\Psi(Z))|_{z=0}=\partial_Z\Psi(0)=0$ (in case (i)),

$$-\int_0^\infty \frac{\Psi(Z)}{Z}\partial_Z^3\Psi(Z)w_1(Z)\,\mathrm{d}Z=\int_0^\infty \partial_Z^2\Psi(Z)\frac{\mathrm{d}}{\mathrm{d}Z}\left(\frac{\Psi(Z)}{Z}w_1(Z)\right)\mathrm{d}Z.$$

As in Lemma C.1, we set $\Phi(Z)=\Psi(Z)/Z$. Let us write $\partial_Z^2\Psi$ as

$$\partial_Z^2\Psi(Z)=\partial_Z^2(Z\Phi(Z))=2\partial_Z\Phi(Z)+Z\partial_Z^2\Phi(Z).$$

Performing integrations by parts and assuming that $\partial_Z w_1(0)=0$, we obtain

$$\begin{aligned}\int_0^\infty &\partial_Z^2\Psi(Z)\frac{\mathrm{d}}{\mathrm{d}Z}(\Phi(Z)w_1(Z))\,\mathrm{d}Z\\ &=\frac{3}{2}\int_0^\infty(\partial_Z\Phi)^2(w_1(Z)-Z\partial_Z w_1(Z))\,\mathrm{d}Z+\frac{1}{2}\int_0^\infty \Psi(Z)^2Z^{-1}\partial_Z^3 w_1(Z)\,\mathrm{d}Z.\end{aligned}$$

We shall choose $w_1$ so that $\partial_Z^3 w_1\geq 0$, so that the last term has a positive sign. However, for $Z\gg 1$, $w_1-Z\partial_Z w_1<0$, and therefore we need to add another term to the energy. More precisely, we now multiply (3-4) by $-Z^{-1}\partial_Z(\partial_Z\Phi w_2)$, with a weight $w_2$ which vanishes identically in a neighborhood of zero. We obtain

$$-\int_0^\infty \frac{\Psi(Z)}{Z}\partial_Z(\partial_Z\Phi w_2)\,\mathrm{d}Z=\int_0^\infty(\partial_Z\Phi(Z))^2w_2(Z)\,\mathrm{d}Z.$$

We then take $w_i(Z)=W_i(rZ)$, with $0<r\ll 1$ and $W_1$, $W_2$ satisfying the following properties:

- $W_1\equiv 1$, $W_2\equiv 0$ in a neighborhood of zero.
- $\partial_Z^3 W_1\geq 0$.
- $W_1(Z)=C\exp(Z^{4/5})$ for $Z$ large enough.
- $W_2+\frac{3}{2}(W_1-Z\partial_Z W_1)\gtrsim(1+Z^{4/5})\exp(Z^{4/5})$.

Our energy is then

$$\begin{aligned}\int_0^\infty(\partial_Z^4\Psi(Z))^2W_1(rZ)\,\mathrm{d}Z&+\int_0^\infty(\partial_Z\Phi(Z))^2\big[W_2+\tfrac{3}{2}(W_1-ZW_1')\big](rZ)\,\mathrm{d}Z\\ &\qquad+\tfrac{1}{2}r^3\int_0^\infty\Psi(Z)^2Z^{-1}\partial_Z^3W_1(rZ)\,\mathrm{d}Z\\ \gtrsim\int_0^\infty(\partial_Z^4\Psi(Z))^2\exp((rZ)^{\frac{4}{5}})\,\mathrm{d}Z&+\int_0^\infty(\partial_Z\Phi(Z))^2(1+(rZ)^{\frac{4}{5}})\exp((rZ)^{\frac{4}{5}})\,\mathrm{d}Z.\end{aligned}$$

Let us now consider the two commutator terms, namely

$$r\int_0^\infty\partial_Z^4\Psi(Z)\partial_Z^3\Psi(Z)\partial_Z W_1(rZ)\,\mathrm{d}Z\quad\text{and}\quad\int_0^\infty\partial_Z^4\Psi(Z)\partial_Z^2(\partial_Z\Phi w_2)\,\mathrm{d}Z.$$

For the first one, we write $\partial_Z^3\Psi = Z\partial_Z^3\Phi + 3\partial_Z^2\Phi$. We note that $\partial_Z W_1(rZ) \lesssim \mathbf{1}_{rZ>c}(rZ)^{-1/5}\exp((rZ)^{4/5})$. Using the second part of Lemma C.1, we infer that there exists $\delta > 0$ such that

$$\begin{aligned}
r\int_0^\infty & \partial_Z^4\Psi(Z)\partial_Z^3\Psi(Z)W_1'(rZ)\,\mathrm{d}Z \\
&\lesssim \left(\int_0^\infty (\partial_Z^4\Psi(Z))^2 W_1(rZ)\,\mathrm{d}Z\right)^{\frac12}\left(\int_0^\infty \mathbf{1}_{rZ>c}(rZ)^{\frac85}(\partial_Z^3\Phi(Z))^2\exp((rZ)^{\frac45})\,\mathrm{d}Z\right)^{\frac12} \\
&\qquad + r\left(\int_0^\infty (\partial_Z^4\Psi(Z))^2 W_1(rZ)\,\mathrm{d}Z\right)^{\frac12}\left(\int_0^\infty \mathbf{1}_{rZ>c}(rZ)^{-\frac25}(\partial_Z^2\Phi(Z))^2\exp((rZ)^{\frac45})\,\mathrm{d}Z\right)^{\frac12} \\
&\lesssim r^\delta\left[\int_0^\infty (\partial_Z^4\Psi(Z))^2\exp((rZ)^{\frac45})\,\mathrm{d}Z + \int_0^\infty (\partial_Z\Phi(Z))^2(rZ)^{\frac45}\exp((rZ)^{\frac45}\,\mathrm{d}Z\right].
\end{aligned}$$

Let us now address the second commutator term. We have for instance, using once again Lemma C.1,

$$\begin{aligned}
\int_0^\infty & \partial_Z^4\Psi(Z)\partial_Z^3\Phi w_2(rZ)\,\mathrm{d}Z \\
&\lesssim \left(\int_0^\infty (\partial_Z^4\Psi(Z))^2 W_1(rZ)\,\mathrm{d}Z\right)^{\frac12}\left(\int_0^\infty \mathbf{1}_{rZ>c}(rZ)^{\frac85}(\partial_Z^3\Phi)^2\exp((rZ)^{\frac45})\,\mathrm{d}Z\right)^{\frac12} \\
&\lesssim r^{\frac23}\left[\int_0^\infty (\partial_Z^4\Psi(Z))^2\exp((rZ)^{\frac45})\,\mathrm{d}Z + \int_0^\infty (\partial_Z\Phi(Z))^2(rZ)^{\frac45})\exp((rZ)^{\frac45}\,\mathrm{d}Z\right].
\end{aligned}$$

The two other terms are treated in a similar fashion. As in the first case, we find that for $r$ small enough, the energy is controlled by

$$\int_0^1 \frac{S(Z)^2}{Z^2}\,\mathrm{d}Z + \int_0^\infty S(Z)^2(1+(rZ)^{\frac25})\exp{(rZ)^{\frac45}}\,\mathrm{d}Z.$$

We conclude by a Lax–Milgram type argument. □

## Acknowledgments

The authors thank Miguel Rodrigues for fruitful discussions especially for the identification of the limit profile, and the referees for the quality and thoroughness of their comments. This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 637653, project BLOC) and by the French National Research Agency (grants ANR-18-CE40-0027, project SingFlows and ANR-23-CE40-0014-01, project BOURGEONS. A.-L. Dalibard acknowledges the support of the Institut Universitaire de France. J. Guillod acknowledges the support of the Initiative d'Excellence (Idex) of Sorbonne University through the Emergence program.

## References

[Antontsev et al. 2000] S. Antontsev, A. Meirmanov, and B. V. Yurinsky, "A free-boundary problem for Stokes equations: classical solutions", *Interfaces Free Bound.* **2**:4 (2000), 413–424. MR

[Boyer and Fabrie 2013] F. Boyer and P. Fabrie, *Mathematical tools for the study of the incompressible Navier–Stokes equations and related models*, Applied Mathematical Sciences **183**, Springer, 2013. MR Zbl

[Castro et al. 2019a] A. Castro, D. Córdoba, and D. Lear, "Global existence of quasi-stratified solutions for the confined IPM equation", *Arch. Ration. Mech. Anal.* **232**:1 (2019), 437–471. MR Zbl

[Castro et al. 2019b] A. Castro, D. Córdoba, and D. Lear, "On the asymptotic stability of stratified solutions for the 2D Boussinesq equations with a velocity damping term", *Math. Models Methods Appl. Sci.* **29**:7 (2019), 1227–1277. MR Zbl

[Constantin et al. 2015] P. Constantin, V. Vicol, and J. Wu, "Analyticity of Lagrangian trajectories for well posed inviscid incompressible fluid models", *Adv. Math.* **285** (2015), 352–393. MR Zbl

[Córdoba et al. 2007] D. Córdoba, F. Gancedo, and R. Orive, "Analytical behavior of two-dimensional incompressible flow in porous media", *J. Math. Phys.* **48**:6 (2007), art. id. 065206. MR Zbl

[Córdoba et al. 2011] D. Cordoba, D. Faraco, and F. Gancedo, "Lack of uniqueness for weak solutions of the incompressible porous media equation", *Arch. Ration. Mech. Anal.* **200**:3 (2011), 725–746. MR Zbl

[Drazin and Reid 2004] P. G. Drazin and W. H. Reid, *Hydrodynamic stability*, 2nd ed., Cambridge University Press, Cambridge, 2004. MR Zbl

[Elgindi 2017] T. M. Elgindi, "On the asymptotic stability of stationary solutions of the inviscid incompressible porous medium equation", *Arch. Ration. Mech. Anal.* **225**:2 (2017), 573–599. MR

[Galdi 2011] G. P. Galdi, *An introduction to the mathematical theory of the Navier–Stokes equations*: *steady-state problems*, 2nd ed., Springer, 2011. MR Zbl

[Gancedo et al. 2025] F. Gancedo, R. Granero-Belinchón, and E. Salguero, "Long time interface dynamics for gravity Stokes flow", *SIAM J. Math. Anal.* **57**:2 (2025), 1680–1724. MR Zbl

[Grayer 2023] H. Grayer, II, "Dynamics of density patches in infinite Prandtl number convection", *Arch. Ration. Mech. Anal.* **247**:4 (2023), art. id. 69. MR Zbl

[Höfer 2018] R. M. Höfer, "Sedimentation of inertialess particles in Stokes flows", *Comm. Math. Phys.* **360**:1 (2018), 55–101. MR Zbl

[Inversi 2023] M. Inversi, "Lagrangian solutions to the transport-Stokes system", *Nonlinear Anal.* **235** (2023), art. id. 113333. MR Zbl

[Isett and Vicol 2015] P. Isett and V. Vicol, "Hölder continuous solutions of active scalar equations", *Ann. PDE* **1**:1 (2015), art. id. 2. MR Zbl

[Kiselev and Yao 2023] A. Kiselev and Y. Yao, "Small scale formations in the incompressible porous media equation", *Arch. Ration. Mech. Anal.* **247**:1 (2023), art. id. 1. MR Zbl

[Leblond 2022] A. Leblond, "Well-posedness of the Stokes-transport system in bounded domains and in the infinite strip", *J. Math. Pures Appl.* (9) **158** (2022), 120–143. MR Zbl

[Leblond 2023] A. Leblond, *Well-posedness and long-time behaviour of the Stokes-transport equation*, Ph.D. thesis, Sorbonne Université, 2023, available at https://theses.hal.science/tel-04356556.

[Lieb and Loss 2001] E. H. Lieb and M. Loss, *Analysis*, 2nd ed., Graduate Studies in Mathematics **14**, Amer. Math. Soc., Providence, RI, 2001. MR Zbl

[Lions and Magenes 1968] J.-L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications*, *Vol.* 1, Travaux et Recherches Mathématiques **17**, Dunod, Paris, 1968. MR

[Mecherbet 2021] A. Mecherbet, "On the sedimentation of a droplet in Stokes flow", *Commun. Math. Sci.* **19**:6 (2021), 1627–1654. MR

[Mecherbet and Sueur 2024] A. Mecherbet and F. Sueur, "A few remarks on the transport-Stokes system", *Ann. H. Lebesgue* **7** (2024), 1367–1408. MR

[Park 2025] J. Park, "Stability analysis of the incompressible porous media equation and the Stokes transport system via energy structure", *Calc. Var. Partial Differential Equations* **64**:5 (2025), art. id. 169. MR Zbl

[Shvydkoy 2011] R. Shvydkoy, "Convex integration for a class of active scalar equations", *J. Amer. Math. Soc.* **24**:4 (2011), 1159–1174. MR Zbl

[Xue 2009] L. Xue, "On the well-posedness of incompressible flow in porous media with supercritical diffusion", *Appl. Anal.* **88**:4 (2009), 547–561. MR Zbl

[Yu and He 2014] W. Yu and Y. He, "On the well-posedness of the incompressible porous media equation in Triebel–Lizorkin spaces", *Bound. Value Probl.* (2014), art. id. 95. MR Zbl

Received 24 Nov 2023. Accepted 23 Sep 2024.

ANNE-LAURE DALIBARD: anne-laure.dalibard@sorbonne-universite.fr
*Laboratoire Jacques-Louis Lions, UMR CNRS 7598, Sorbonne Université, Université Paris Cité, Inria, Paris, France*

and

*Département de Mathématiques et applications, École Normale Supérieure, Université PSL, Paris, France*

JULIEN GUILLOD: julien.guillod@sorbonne-universite.fr
*Laboratoire Jacques-Louis Lions, UMR CNRS 7598, Sorbonne Université, Université Paris Cité, Inria, Paris, France*

and

*Département de Mathématiques et applications, École Normale Supérieure, Université PSL, Paris, France*

ANTOINE LEBLOND: antoine.leblond@mpimet.mpg.de
*Laboratoire Jacques-Louis Lions, UMR CNRS 7598, Sorbonne Université, Université Paris Cité, Inria, Paris, France*

**mathematical sciences publishers** msp

ANALYSIS AND PDE
Vol. 18 (2025), No. 8, pp. 2033–2060
DOI: 10.2140/apde.2025.18.2033

# RECONSTRUCTION FOR THE CALDERÓN PROBLEM WITH LIPSCHITZ CONDUCTIVITIES

PEDRO CARO, MARÍA ÁNGELES GARCÍA-FERRERO AND KEITH M. ROGERS

We determine the conductivity of the interior of a body using electrical measurements on its surface. We assume only that the conductivity is bounded below by a positive constant and that the conductivity and surface are Lipschitz continuous. To determine the conductivity we first solve an associated integral equation in a ball $B$ that properly contains the body, finding solutions in $H^1(B)$. A key ingredient is to equip this Sobolev space with an equivalent norm which depends on two auxiliary parameters that can be chosen to yield a contraction.

## 1. Introduction

We consider the conductivity equation in a bounded domain $\Omega \subset \mathbb{R}^n$ and place electric potentials $\phi \in H^{1/2}(\partial\Omega)$ on the Lipschitz boundary $\partial\Omega$:

$$\begin{cases} \nabla \cdot (\sigma \nabla u) = 0 & \text{in } \Omega, \\ u|_{\partial\Omega} = \phi. \end{cases} \tag{1}$$

Throughout the article, the conductivity $\sigma$ is assumed to be bounded above and below by positive constants, so that (1) has a unique weak solution $u$ in the $L^2$-Sobolev space $H^1(\Omega)$. The Dirichlet-to-Neumann map $\Lambda_\sigma$ can then be formally defined by

$$\Lambda_\sigma : \phi \mapsto \sigma\, \partial_\nu u|_{\partial\Omega}, \tag{2}$$

where $\nu$ denotes the outward unit normal vector to $\partial\Omega$. This provides us with the steady-state perpendicular currents induced by the electric potentials $\phi$.

Motivated by the possibility of creating an image of the interior of a body from these noninvasive voltage-to-current measurements on its surface, Calderón [2006] asked whether the conductivity $\sigma$ is uniquely determined by $\Lambda_\sigma$ and, if so, whether $\sigma$ can be calculated from $\Lambda_\sigma$. In two dimensions, Astala and Päivärinta answered the uniqueness part [2006b] and also provided a reconstruction algorithm [2006a]. The two-dimensional problem has distinct mathematical characteristics, so from now on we consider only $n \geqslant 3$.

With $n \geqslant 3$, it has so far been necessary to make additional regularity assumptions. Kohn and Vogelius [1984] proved uniqueness for real-analytic conductivities, and Sylvester and Uhlmann [1987] improved this to smooth conductivities. Nachman, Sylvester and Uhlmann [Nachman et al. 1988] then proved uniqueness for twice continuously differentiable conductivities, and Nachman [1988] and Novikov [1988]

*MSC2020:* 35R30.
*Keywords:* Calderón inverse problem, conductivity, reconstruction, low regularity.

© 2025 The Authors, under license to MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

provided reconstruction algorithms. These pioneering articles provoked a great deal of interesting work, including that of Brown [1996], Päivärinta, Panchenko and Uhlmann [Päivärinta et al. 2003] and Brown and Torres [2003] for conductivities with $\frac{3}{2}$ derivatives. In the past decade, a breakthrough was made by Haberman and Tataru [2013], who proved uniqueness for continuously differentiable conductivities or Lipschitz conductivities with $\|\nabla \log \sigma\|_\infty$ sufficiently small. García and Zhang [2016] then provided a reconstruction algorithm under the same assumptions. Two of the authors removed the smallness condition from the uniqueness result in [Caro and Rogers 2016], and the purpose of this article will be to extend this work to a reconstruction algorithm that holds for all Lipschitz conductivities. We will not assume that the conductivity is constant near the boundary, nor will we extend the conductivity in order to achieve this, leading to simpler formulas than those of [García and Zhang 2016]; see Section 3.

Before we outline the proof, we remark that there are also uniqueness results for conductivities in Sobolev spaces; see [Haberman 2015; Ham et al. 2021; Ponce-Vanegas 2021]. In particular, [Haberman 2015] proved that uniqueness holds for bounded conductivities in $W^{1,n}(\overline{\Omega})$, with $n = 3$ or 4. Note that this is a strictly larger class than Lipschitz, however there are obstacles to reconstruction via their methods; see Remark 11.2 for more details. It has been conjectured that Lipschitz continuity is the sharp threshold within the scale of Hölder continuity; see for example [Brown 1996] or [Uhlmann 1998, Open Problem 1].

When $\sigma$ is Lipschitz, weak solutions to (1) are in fact strong solutions; see for example [Zhang and Bao 2012, Theorem 1.3]. Defining the Dirichlet-to-Neumann map as in (2) by identifying $\sigma\, \partial_\nu u|_{\partial\Omega}$ with the normal trace of $\sigma\nabla u$, we have the divergence identity

$$\int_{\partial\Omega} \Lambda_\sigma[\phi]\psi = \int_\Omega \sigma\nabla u \cdot \nabla\psi$$

whenever $(\phi, \psi) \in H^{1/2}(\partial\Omega) \times H^1(\Omega)$; see for example [Kim and Kwon 2022, Proposition 2.4]. Given this identity, it is possible to describe the heuristic which underlies the reconstruction: For each $\xi \in \mathbb{R}^n$, one hopes to choose an oscillating pair $(\phi, \psi)$ so that the right-hand side becomes a nonlinear Fourier transform of $\sigma$ evaluated at $\xi$. As the left-hand side can be calculated from the measurements, the conductivity might then be recoverable by Fourier inversion. Indeed, much of the literature, including the original work of Calderón [2006], has involved pairs $(e^{\rho\cdot x}, e^{\rho'\cdot x})$, with $\rho, \rho' \in \mathbb{C}^n$ chosen carefully, so that $\rho + \rho'$ is equal to a real constant multiple of $-i\xi$, where $i := \sqrt{-1}$. The hope is that the essentially harmonic $u$ is not so different from $e_\rho := e^{\rho\cdot x}$, and so the complex vector $\rho$ is chosen in such a way that $\rho \cdot \rho = 0$, so that $e_\rho$ is harmonic.

In fact we begin by noting that $u$ is a solution to the conductivity equation if and only if $v = \sigma^{1/2}u$ is a solution to the Schrödinger equation

$$\Delta v = qv \quad \text{in } \Omega, \tag{3}$$

where formally $q = \sigma^{-1/2}\Delta\sigma^{1/2}$. Kohn and Vogelius [1985] observed that if $\sigma|_{\partial\Omega}$ and $\nu \cdot \nabla\sigma|_{\partial\Omega}$ are known, then the Dirichlet-to-Neumann map $\Lambda_q$ for the Schrödinger equation (3) can be written in terms of $\Lambda_\gamma$, and so the literature has mainly considered the essentially equivalent problem of recovering $q$ from $\Lambda_q$ (which is intimately connected to inverse scattering at fixed energy). We will only partially use the equivalence however: we will recover $q$ directly from $\Lambda_\gamma$, circumventing the need to calculate

$\nu \cdot \nabla\sigma|_{\partial\Omega}$. This is connected to the fact that our conductivities are not regular enough to define $q$ in a pointwise fashion. However, as noted by Brown [1996], it suffices to define $\langle qv, \psi\rangle := \langle q, v\psi\rangle$ for suitable test functions $\psi$, with

$$\langle q, \bullet\rangle := -\int_\Omega \nabla\sigma^{1/2} \cdot \nabla(\sigma^{-1/2}\bullet). \tag{4}$$

By the product rule and the Cauchy–Schwarz inequality, $\langle q, \bullet\rangle$ and $\langle qv, \bullet\rangle$ are bounded linear functionals on $H^1(B)$, where $B$ is a ball that properly contains $\Omega$, so in particular we can make sense of $q$ and $qv$ as distributions.

Rather than solving (3) directly, we consider solutions to the Lippmann–Schwinger-type equation

$$v = \Delta^{-1} \circ \mathrm{M}_q[v] + e_\rho, \tag{5}$$

where $\mathrm{M}_q : f \mapsto qf$ and the inverse of the Laplacian is defined using the Faddeev fundamental solution; see Section 2.1. Integral equations like this are usually solved globally, however we will find a $v \in H^1(B)$ which is a solution of (5) in the ball $B$. Writing $v = e_\rho(1 + w)$ and additionally requiring that the remainders $w$ vanish in some sense as $|\rho| \to \infty$ gives hope that the nonlinear Fourier transform will converge to the linear Fourier transform in the limit. Solutions of this type were introduced to the problem by Sylvester and Uhlmann [1987] and have since become known as CGO solutions, where CGO stands for complex geometrical optics. Substituting into (3) and multiplying by $e_{-\rho}$, we find that

$$\Delta_\rho w = \mathrm{M}_q[1 + w] \quad \text{in } \Omega, \tag{6}$$

where $\Delta_\rho := \Delta + 2\rho \cdot \nabla$. In much of the literature $\Delta_\rho$ is inverted using the Fourier transform and the resulting integral equation is solved globally via a contraction for $\Delta_\rho^{-1} \circ \mathrm{M}_q$ and Neumann series. In order to reconstruct $\sigma$ from $\Lambda_\sigma$ (as opposed to just proving uniqueness), we must additionally determine which electric potentials should be placed on the boundary in order to generate the CGO solutions. A contraction for $\Delta_\rho^{-1} \circ \mathrm{M}_q$ can also be helpful in this step; however, the need for such a contraction was circumvented in the uniqueness result of [Caro and Rogers 2016], instead solving the differential equation (6) via the method of a priori estimates.

Nachman and Street [2010] were able to recover the boundary values of CGO solutions that had been constructed via a priori estimates, however, we were unable to take advantage of their ideas; see Remark 11.1 for more details. Instead we will reprove the existence of CGO solutions, this time via Neumann series; however, we will adopt the previously mentioned intermediate approach of solving the integral equation in the ball $B$. That is to say, we find a $w \in H^1(B)$ such that

$$(\mathrm{I} - \Delta_\rho^{-1} \circ \mathrm{M}_q)w = \Delta_\rho^{-1} \circ \mathrm{M}_q[1], \tag{7}$$

where the identity holds as elements of $H^1(B)$. This is equivalent to (5) when $\Delta^{-1} \circ \mathrm{M}_q$ is defined appropriately; see Remark 9.3.

Most of the article will be occupied by the proof of the contraction for $\Delta_\rho^{-1} \circ \mathrm{M}_q$ in Sections 4–9. In Section 4 we give a sketch of its proof before proving the key Carleman estimate in Section 5. In Section 6 we incorporate the associated convex weights into our localised versions of the Haberman–Tataru norms, so that they not only depend on $\rho$ but also on an auxiliary parameter $\lambda > 1$. The final estimate for $\Delta_\rho^{-1}$, proved in Section 7, is somewhat weaker and easier to prove than the main estimate of [Caro and Rogers 2016], so the present article also simplifies the uniqueness result of that work. In Section 8 we bound $\mathrm{M}_q$ with respect to the new norms, and in Section 9 we choose the parameters in order to yield the contraction.

In Section 2 we list some of the main definitions before presenting the reconstruction algorithm in Section 3. The reconstruction formulas will not make mention of the new norms, which are only used in Section 11 to prove the validity of the formulas. In the final Section 12 we suggest some simplifications that could make the algorithm easier to implement.

## 2. Preliminary notation

We invert our main operator $\Delta_\rho$ initially on the space of Schwartz functions $\mathcal{S}(\mathbb{R}^n)$ using the Fourier transform defined by

$$\hat{f}(\xi) := \int_{\mathbb{R}^n} e^{-i\xi\cdot x} f(x)\,\mathrm{d}x$$

for all $\xi \in \mathbb{R}^n$ and $f \in \mathcal{S}(\mathbb{R}^n)$. By integration by parts, one can calculate that

$$\widehat{\Delta_\rho f}(\xi) = m_\rho(\xi)\hat{f}(\xi), \quad \text{where } m_\rho(\xi) := -|\xi|^2 + 2i\rho\cdot\xi, \tag{8}$$

for all $\xi \in \mathbb{R}^n$. The reciprocal of this Fourier multiplier is integrable on compact sets, so we can define an inverse by

$$\Delta_\rho^{-1} g(x) := \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\frac{1}{m_\rho(\xi)}\hat{g}(\xi)\,\mathrm{d}\xi$$

for all $x \in \mathbb{R}^n$ and $g \in \mathcal{S}(\mathbb{R}^n)$.

**2.1.** ***The Faddeev fundamental solutions.*** Writing the inverse Fourier transform of the product as a convolution, we find

$$\Delta_\rho^{-1} g(x) = \int_{\mathbb{R}^n} F_\rho(x-y)g(y)\,\mathrm{d}y \tag{9}$$

for all $x \in \mathbb{R}^n$ and $g \in \mathcal{S}(\mathbb{R}^n)$, where the fundamental solution $F_\rho$ for $\Delta_\rho$ is defined by

$$F_\rho(x) := \lim_{r\to\infty}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\frac{1}{m_\rho(\xi)}\hat{\chi}(\xi/r)\,\mathrm{d}\xi.$$

Here $\chi \in \mathcal{S}(\mathbb{R}^n)$ must be positive and satisfy $\hat{\chi}(0) = 1$, but the limit is insensitive to the precise choice of $\chi$ and so the integral is often written formally, taking $\hat{\chi} = 1$. This fundamental solution was first considered by Faddeev [1965] in the context of quantum inverse scattering.

We also consider the associated fundamental solution $G_\rho := e_\rho F_\rho$ for the Laplacian, and we will often write $G_\rho(x, y) := G_\rho(x-y)$. This is not so different from the usual potential-theoretic fundamental solution. Indeed, by subtracting one from the other, one obtains a harmonic function which is thus smooth by Weyl's lemma:

$$H_\rho(x) := G_\rho(x) - \frac{c_n}{(2-n)}\frac{1}{|x|^{n-2}}, \tag{10}$$

where $c_n$ denotes the reciprocal of the measure of the unit sphere. For more details regarding the properties of Faddeev's fundamental solutions, see [Newton 1989, Section 6.1].

**2.2.** ***The boundary integral.*** For notational compactness we write the reconstruction formulas in terms of the bilinear functional $BI_{\Lambda_\sigma} : H^{1/2}(\partial\Omega) \times H^1(\Omega) \to \mathbb{C}$ defined by

$$BI_{\Lambda_\sigma}(\phi, \psi) := \int_{\partial\Omega} (\sigma^{-1/2}\Lambda_\sigma[\sigma^{-1/2}\phi] - \nu \cdot \nabla P_0[\phi])\psi, \tag{11}$$

where $P_0[\phi]$ denotes the harmonic extension of $\phi$. Brown [2001] calculated $\sigma|_{\partial\Omega}$ from $\Lambda_\sigma$, so the boundary integral $BI_{\Lambda_\sigma}$ can be recovered from $\Lambda_\sigma$. In Lemma 10.1 we will prove that

$$BI_{\Lambda_\sigma}(\phi, G_\rho(\mathrm{x}, \bullet)) \in H^1(B \setminus \overline{\Omega}),$$

where $B$ properly contains $\Omega$ and $f(\mathrm{x})$ denotes a function that takes the values $f(x)$ for all $x$ in the domain. This allows us to define $\Gamma_{\Lambda_\sigma} : H^{1/2}(\partial\Omega) \to H^{1/2}(\partial\Omega)$ by taking the outer trace on $\partial\Omega$:

$$\Gamma_{\Lambda_\sigma}[\phi] := BI_{\Lambda_\sigma}(\phi, G_\rho(\mathrm{x}, \bullet))|_{\partial\Omega}. \tag{12}$$

As $H_\rho$ is smooth, the singularity of $G_\rho$ is the same as that of the usual potential-theoretic fundamental solution, so $\Gamma_{\Lambda_\sigma}$ shares many properties with the single layer potential; see for example [Mitrea and Taylor 1999, Propositions 3.8 and 7.9]. However, we will not need these types of estimates going forward.

## 3. The reconstruction algorithm

Recall our a priori assumptions, that the boundary and conductivity are Lipschitz continuous and that the conductivity is bounded below by a positive constant.

The first step of the reconstruction algorithm is to determine the electric potentials that we place on the boundary in order to generate the CGO solutions. As in the previous reconstruction formulas of [García and Zhang 2016; Nachman 1988; Novikov 1988], we resort to the Fredholm alternative; however, once we have obtained the contraction, the argument will be direct, avoiding the use of generalised double layer potentials. The proof is postponed until Section 11.

**Theorem 3.1.** *Consider $\rho \in \mathbb{C}^n$ such that $\rho \cdot \rho = 0$ and $|\rho|^2 = \rho \cdot \bar{\rho}$ is sufficiently large. Let $\Gamma_{\Lambda_\sigma}$ be defined by* (12). *Then*

(i) $\Gamma_{\Lambda_\sigma} : H^{1/2}(\partial\Omega) \to H^{1/2}(\partial\Omega)$ *is bounded compactly,*

(ii) *if* $\Gamma_{\Lambda_\sigma}[\phi] = \phi$, *then* $\phi = 0$,

(iii) $\mathrm{I} - \Gamma_{\Lambda_\sigma}$ *has a bounded inverse on* $H^{1/2}(\partial\Omega)$,

*and if* $v = e_\rho(1 + w)$, *where* $w \in H^1(B)$ *is a solution to* (7), *then*

(iv) $v|_{\partial\Omega} = (\mathrm{I} - \Gamma_{\Lambda_\sigma})^{-1}[e_\rho|_{\partial\Omega}]$.

Next we provide a formula for the Fourier transform $\hat{q}(\xi) := \langle q, e^{-i\xi\cdot\mathrm{x}}\rangle$, where $q$ is defined in (4). Again we postpone the proof until the penultimate section.

**Theorem 3.2.** *Let $\Pi$ be a two-dimensional linear subspace orthogonal to $\xi \in \mathbb{R}^n$, and define*

$$S^1 := \Pi \cap \{\theta \in \mathbb{R}^n : |\theta| = 1\}.$$

*For $\theta \in S^1$, let $\vartheta \in S^1$ be such that $\{\theta, \vartheta\}$ is an orthonormal basis of $\Pi$, and define*

$$\rho := \tau\theta + i\left(-\frac{\xi}{2} + \left(\tau^2 - \frac{|\xi|^2}{4}\right)^{1/2}\vartheta\right), \quad \rho' := -\tau\theta + i\left(-\frac{\xi}{2} - \left(\tau^2 - \frac{|\xi|^2}{4}\right)^{1/2}\vartheta\right),$$

*where $\tau > 1$. Let $BI_{\Lambda_\sigma}$ and $\Gamma_{\Lambda_\sigma}$ be defined by* (11) *and* (12), *respectively. Then*

$$\hat{q}(\xi) = \lim_{T\to\infty} \frac{1}{2\pi T}\int_T^{2T}\int_{S^1} BI_{\Lambda_\sigma}((\mathrm{I} - \Gamma_{\Lambda_\sigma})^{-1}[e_\rho|_{\partial\Omega}], e_{\rho'})\,\mathrm{d}\theta\,\mathrm{d}\tau.$$

Finally, we recover $\sigma$ from $q$ using the approach of [García and Zhang 2016]. By [Brown 2001] and Plancherel's identity, we can now calculate the right-hand side of

$$\begin{cases} \Delta w + |\nabla w|^2 = q & \text{in } \Omega, \\ w|_{\partial\Omega} = \frac{1}{2}\log\sigma|_{\partial\Omega}. \end{cases} \tag{13}$$

If $w \in H^1(\Omega)$ is the unique bounded solution to (13), we then have

$$\sigma = e^{2w} \quad \text{in } \Omega.$$

This completes the reconstruction algorithm.

That $w = \log\sigma^{1/2}$ solves (13) follows directly by inspection of the definition (4) of $q$. For uniqueness, note that if $\tilde{w}$ also solved (13), then $u = w - \tilde{w}$ would solve

$$\begin{cases} \nabla\cdot(\gamma\nabla u) = 0 & \text{in } \Omega, \\ u|_{\partial\Omega} = 0, \end{cases}$$

where $\gamma := e^{w+\tilde{w}}$. Then $u = 0$ by uniqueness of solutions for elliptic equations; see for example [Gilbarg and Trudinger 1983, Corollary 8.2].

## 4. Sketch of the proof of the contraction for $\Delta_\rho^{-1} \circ \mathrm{M}_q$

One of the main ideas of [Haberman and Tataru 2013] was to extend the domain of $\Delta_\rho^{-1}$ using Bourgain-type spaces that are adapted to the problem, instead of the usual Sobolev spaces. With $s = \frac{1}{2}$ or $-\frac{1}{2}$, their norms are defined by

$$\|\bullet\|_{\dot{X}^s_\rho} : f \in \mathcal{S}(\mathbb{R}^n) \mapsto \||m_\rho|^s \hat{f}\|_{L^2(\mathbb{R}^n)},$$

where $m_\rho$ is the multiplier defined in (8). Then $\dot{X}^s_\rho$ is defined to be the Banach completion of $\mathcal{S}(\mathbb{R}^n)$ with respect to this norm. It is immediate from the definitions that

$$\|\Delta_\rho^{-1} g\|_{\dot{X}^{1/2}_\rho} \leqslant \|g\|_{\dot{X}^{-1/2}_\rho} \tag{14}$$

whenever $g \in \mathcal{S}(\mathbb{R}^n)$, which can be used to continuously extend the operator. For ease of reference we will call (14) *the trivial inequality*.

On the other hand, Haberman and Tataru also proved that $\mathrm{M}_q : f \mapsto qf$ satisfies

$$\|\mathrm{M}_q f\|_{\dot{X}_\rho^{-1/2}} \leqslant C\|\nabla \log\sigma\|_\infty(1+|\rho|^{-1}\|\nabla\log\sigma\|_\infty)\|f\|_{\dot{X}_\rho^{1/2}} \tag{15}$$

whenever $f \in \dot{X}_\rho^{1/2}$; see [Haberman and Tataru 2013, Theorem 2.1]. Together these inequalities yield a contraction for $\Delta_\rho^{-1} \circ \mathrm{M}_q$ whenever $|\rho| > 1$ and $\|\nabla \log\sigma\|_\infty$ is sufficiently small. In order to remove this smallness condition, we will alter the norms in such a way that the constant of (15) can be taken small for any Lipschitz conductivity, while maintaining a version of (14).

There is a natural gain for the higher frequencies in (15) whereas a gain for the lower frequencies can be engineered in (14) by introducing convex weights. This was the key observation of [Caro and Rogers 2016]. In order to have a gain for all frequencies, in at least one of the inequalities, we dampen the higher frequencies relative to the lower frequencies in our main norm (with the lower frequencies dampened relative to the higher frequencies in the dual norm), so that the gain for the lower frequencies in our version of (14) is passed through to our version of (15).

We prove the Carleman estimate in Section 5, we define new Banach spaces in Section 6, and then we extend the domain of $\Delta_\rho^{-1}$ via density in Section 7. We prove our version of (15) in Section 8 and then combine the estimates to obtain the contraction in Section 9.

## 5. Bounds for $\Delta_\rho^{-1}$ with convex weights

Let $B$ be an open ball centred at the origin, with radius

$$R := 2\sup_{x\in\Omega}|x|,$$

so that we comfortably have $\Omega \subset B$. The forthcoming constants will invariably depend on this $R$, but never on the auxiliary parameters $\rho \in \mathbb{C}^n$ or $\lambda > 1$.

**5.1.** ***The Carleman estimate.*** Here we will deduce our estimate for $\Delta_\rho^{-1}$ from a Carleman estimate for $\Delta_\rho$ before defining the main spaces and their duals in the following section. We improve upon the estimate

$$|\rho|\|\Delta_\rho^{-1} f\|_{L^2(B)} \leqslant C\|f\|_{L^2(\mathbb{R}^n)} \tag{16}$$

whenever $f \in C_c^\infty(B)$, which does not seem strong enough to construct CGO solutions for Lipschitz conductivities. The inequality (16) follows by combining

$$|\rho|^{1/2}\|g\|_{L^2(B)} \leqslant C\|g\|_{\dot{X}_\rho^{1/2}} \tag{17}$$

whenever $g \in \dot{X}_\rho^{1/2}$ with the trivial inequality (14), and then

$$|\rho|^{1/2}\|f\|_{\dot{X}_\rho^{-1/2}} \leqslant C\|f\|_{L^2(\mathbb{R}^n)} \tag{18}$$

whenever $f \in C_c^\infty(B)$. The constants $C > 1$ depend only on $R$. Away from the zero set of the Fourier multiplier $m_\rho$, these inequalities are obvious, and the localisation serves to blur out the effect of the zero set; see Lemma 2.2 of [Haberman and Tataru 2013] for the proof.

In the following lemma we improve the constant in (16) by introducing exponential weights that depend on the auxiliary parameter $\lambda > 1$. The extra gain in terms of $\lambda$ will be key to constructing our CGO solutions for Lipschitz conductivities.

**Lemma 5.1.** *Consider $\rho \in \mathbb{C}^n$ such that $\rho \cdot \rho = 0$, and write $\theta := \operatorname{Re}\rho/|\operatorname{Re}\rho|$. Then*

$$\int_B |\Delta_\rho^{-1} f(x)|^2 e^{\lambda(\theta\cdot x)^2}\,\mathrm{d}x \leqslant \frac{2}{\lambda|\rho|^2}\int_{\mathbb{R}^n} |f(x)|^2 e^{\lambda(\theta\cdot x)^2}\,\mathrm{d}x$$

*whenever $f \in C_c^\infty(\mathbb{R}^n)$ and $\lambda > 1$ satisfies $|\rho| \geqslant 4\lambda R$.*

*Proof.* If $m_\rho$ had been defined slightly differently at the beginning, including a superfluous $\rho\cdot\rho$ term, we could have proved a version of this lemma without the hypothesis that $\rho\cdot\rho = 0$. In fact, we begin by reducing to a purely real vector case. Indeed, letting $\operatorname{Re}\rho, \operatorname{Im}\rho \in \mathbb{R}^n$ denote the real and imaginary parts of $\rho$, respectively, we define $\Delta_{\operatorname{Re}\rho}^{-1}$ as in Section 2, but with $m_\rho$ replaced by

$$m_{\operatorname{Re}\rho}(\xi) := -|\xi|^2 + 2i\operatorname{Re}\rho\cdot\xi + \operatorname{Re}\rho\cdot\operatorname{Re}\rho$$

for all $\xi \in \mathbb{R}^n$. Then, observing that

$$m_\rho(\xi) = -|\xi|^2 + 2i\rho\cdot\xi + \rho\cdot\rho = m_{\operatorname{Re}\rho}(\xi + \operatorname{Im}\rho)$$

and defining the modulation operator by $\operatorname{Mod}_{\operatorname{Im}\rho} f(\mathrm{x}) := e^{i\operatorname{Im}\rho\cdot\mathrm{x}} f(\mathrm{x})$, we find that

$$\operatorname{Mod}_{\operatorname{Im}\rho}[\Delta_\rho^{-1} f] = \Delta_{\operatorname{Re}\rho}^{-1}[\operatorname{Mod}_{\operatorname{Im}\rho} f]$$

whenever $f \in C_c^\infty(\mathbb{R}^n)$. Recalling that $|\rho|^2 = 2|\operatorname{Re}\rho|^2$ if $\rho\cdot\rho = 0$, it will therefore suffice to prove

$$\int_B |\Delta_{\operatorname{Re}\rho}^{-1} f(x)|^2 e^{\lambda(\theta\cdot x)^2}\,\mathrm{d}x \leqslant \frac{1}{\lambda|\operatorname{Re}\rho|^2}\int_{\mathbb{R}^n} |f(x)|^2 e^{\lambda(\theta\cdot x)^2}\,\mathrm{d}x$$

whenever $|\operatorname{Re}\rho| \geqslant 2\lambda R$. Recalling that $\theta := \operatorname{Re}\rho/|\operatorname{Re}\rho|$, by rotating to $\mathrm{e}_n$, this would follow from

$$\int_B |\Delta_{\tau\mathrm{e}_n}^{-1} f(x)|^2 e^{\lambda x_n^2}\,\mathrm{d}x \leqslant \frac{1}{\lambda\tau^2}\int_{\mathbb{R}^n} |f(x)|^2 e^{\lambda x_n^2}\,\mathrm{d}x \tag{19}$$

whenever $f \in C_c^\infty(\mathbb{R}^n)$ and $\tau \geqslant 2\lambda R$.

In order to prove (19), we will first prove the closely related Carleman estimate

$$\|g\|_{L^2(B)}^2 \leqslant \frac{1}{\lambda\tau^2}\|e^{\lambda \mathrm{x}_n^2/2}(\Delta + 2\tau\mathrm{e}_n\cdot\nabla + \tau^2)(e^{-\lambda \mathrm{x}_n^2/2} g)\|_{L^2(\mathbb{R}^n)}^2 \tag{20}$$

whenever $g \in \mathcal{S}(\mathbb{R}^n)$. Defining $\varphi(x) = \tau x_n + \frac{1}{2}\lambda x_n^2$, the integrand of the right-hand side can be rewritten as

$$e^{\varphi}\Delta(e^{-\varphi} g) = \Delta g - \nabla\varphi\cdot\nabla g - \nabla\cdot(\nabla\varphi g) + |\nabla\varphi|^2 g.$$

Defining the formally self-adjoint A and skew-adjoint B by

$$\mathrm{A}g = \Delta g + |\nabla\varphi|^2 g \quad\text{and}\quad \mathrm{B}g = -\nabla\varphi\cdot\nabla g - \nabla\cdot(\nabla\varphi g)$$

and integrating by parts, we have that

$$\|(\mathrm{A}+\mathrm{B})g\|^2_{L^2(\mathbb{R}^n)} = \|\mathrm{A}g\|^2_{L^2(\mathbb{R}^n)} + \|\mathrm{B}g\|^2_{L^2(\mathbb{R}^n)} + \int_{\mathbb{R}^n} [\mathrm{A},\mathrm{B}]g\bar{g}, \tag{21}$$

where $[\mathrm{A},\mathrm{B}] = \mathrm{AB} - \mathrm{BA}$ denotes the commutator. By the definition of $\varphi$, we have

$$\mathrm{A}g(x) = \Delta g(x) + (\tau + \lambda x_n)^2 g(x),$$
$$\mathrm{B}g(x) = -2(\tau + \lambda x_n)\,\partial_{x_n} g(x) - \lambda g(x),$$

which yields

$$[\mathrm{A},\mathrm{B}]g(x) = -4\lambda \partial^2_{x_n} g(x) + 4\lambda(\tau + \lambda x_n)^2 g(x).$$

After another integration by parts, we find

$$\int_{\mathbb{R}^n} [\mathrm{A},\mathrm{B}]g\bar{g} = 4\lambda \int_{\mathbb{R}^n} |\partial_{x_n} g|^2 + 4\lambda \int_{\mathbb{R}^n} |\nabla\varphi|^2 |g|^2,$$

so that, substituting this into (21) and throwing three of the terms away, we find

$$\|e^{\varphi}\Delta(e^{-\varphi}g)\|^2_{L^2(\mathbb{R}^n)} \geqslant 4\lambda \int_{\mathbb{R}^n} |\nabla\varphi|^2 |g|^2.$$

As $|\nabla\varphi(x)| \geqslant \tau - \lambda R$ whenever $|x_n| \leqslant R$, this yields

$$\|e^{\varphi}\Delta(e^{-\varphi}g)\|^2_{L^2(\mathbb{R}^n)} \geqslant 4\lambda(\tau - \lambda R)^2 \|g\|^2_{L^2(B)},$$

which implies (20) whenever $\tau \geqslant 2\lambda R$ and $g \in \mathcal{S}(\mathbb{R}^n)$.

Finally, by density, the inequality (20) also holds for every $g \in L^2_{\mathrm{loc}}(\mathbb{R}^n)$ such that

$$e^{\lambda x_n^2/2}(\Delta + 2\tau \mathrm{e}_n \cdot \nabla + \tau^2)(e^{-\lambda x_n^2/2} g) \in L^2(\mathbb{R}^n).$$

Choosing $g = e^{\lambda x_n^2/2} \Delta^{-1}_{\tau \mathrm{e}_n} f$ with $f \in C^\infty_{\mathrm{c}}(\mathbb{R}^n)$, we find that (20) implies (19). □

**Remark 5.2.** The proof of Lemma 5.1 yields the following strengthened estimate: if $\rho \in \mathbb{C}^n$ and $\theta := \operatorname{Re}\rho / |\operatorname{Re}\rho|$, then

$$\int_{|\theta\cdot x| < |\operatorname{Re}\rho|/(2\lambda)} |\Delta^{-1}_{\rho} f(x)|^2 e^{\lambda(\theta\cdot x)^2}\,\mathrm{d}x \leqslant \frac{1}{\lambda |\operatorname{Re}\rho|^2} \int_{\mathbb{R}^n} |f(x)|^2 e^{\lambda(\theta\cdot x)^2}\,\mathrm{d}x$$

whenever $f \in \mathcal{S}(\mathbb{R}^n)$ is such that the right-hand side is finite and $\lambda > 1$.

**5.2.** ***Estimates for derivatives.*** The inequality of Lemma 5.1 has a gain in the sense of $L^2$, however, this is not enough to construct CGO solutions for Lipschitz conductivities since we need to control the first-order partial derivatives present in the operator $\mathrm{M}_q$. For this we consider

$$\|\bullet\|_{X^{1/2}_{\lambda,\rho}} := \lambda^{1/4} |\rho|^{1/2} \|\bullet\|_{L^2(B, e^{\lambda(\theta\cdot x)^2})} + \frac{1}{\lambda^{1/4}} \|\bullet\|_{\dot{X}^{1/2}_{\rho}} \tag{22}$$

and combine Lemma 5.1 with the trivial inequality (14).

**Lemma 5.3.** *Consider $\rho \in \mathbb{C}^n$ such that $\rho \cdot \rho = 0$, and write $\theta := \operatorname{Re}\rho/|\operatorname{Re}\rho|$. Then there is a constant $C > 1$, depending only on the radius $R$ of $B$, such that*

$$\|\Delta_\rho^{-1} f\|_{X^{1/2}_{\lambda,\rho}} \leqslant \frac{C}{\lambda^{1/4}|\rho|^{1/2}} \|f\|_{L^2(\mathbb{R}^n, e^{\lambda(\theta\cdot x)^2})}$$

*whenever $f \in C_c^\infty(B)$ and $\lambda > 1$ satisfies $|\rho| \geqslant 4\lambda R$.*

*Proof.* The first term in the definition (22) is bounded using Lemma 5.1, so it remains to bound the second term. Combining the trivial inequality (14) with (18), we see that

$$\|\Delta_\rho^{-1} f\|_{\dot{X}^{1/2}_\rho} \leqslant \|f\|_{\dot{X}^{-1/2}_\rho} \leqslant \frac{C}{|\rho|^{1/2}} \|f\|_{L^2(\mathbb{R}^n)} \leqslant \frac{C}{|\rho|^{1/2}} \|f\|_{L^2(\mathbb{R}^n, e^{\lambda(\theta\cdot x)^2})}$$

whenever $f \in C_c^\infty(B)$, where the constant $C > 1$ depends only on $R$. Dividing by $\lambda^{1/4}$ yields the desired estimate for the second term. □

**Lemma 5.4.** *Consider $\rho \in \mathbb{C}^n$ such that $\rho \cdot \rho = 0$, and write $\theta := \operatorname{Re}\rho/|\operatorname{Re}\rho|$. Then there is a constant $C > 1$, depending only on the radius $R$ of $B$, such that*

$$\|\Delta_\rho^{-1} f\|_{X^{1/2}_{\lambda,\rho}} \leqslant C\lambda^{1/4} e^{\lambda R^2/2} \|f\|_{\dot{X}^{-1/2}_\rho}$$

*whenever $f \in C_c^\infty(B)$ and $\lambda > 1$.*

*Proof.* The second term in the definition (22) can be bounded easily using the trivial inequality (14), so it remains to bound the first term. By (17), we have

$$|\rho|^{1/2} \|g\|_{L^2(B, e^{\lambda(\theta\cdot x)^2})} \leqslant e^{\lambda R^2/2} |\rho|^{1/2} \|g\|_{L^2(B)} \leqslant C e^{\lambda R^2/2} \|g\|_{\dot{X}^{1/2}_\rho}$$

whenever $g \in \dot{X}^{1/2}_\rho$, where the constant $C > 1$ depends only on $R$. Taking $g = \Delta_\rho^{-1} f$ and multiplying the inequality by $\lambda^{1/4}$ yields

$$\lambda^{1/4} |\rho|^{1/2} \|\Delta_\rho^{-1} f\|_{L^2(B, e^{\lambda(\theta\cdot x)^2})} \leqslant C\lambda^{1/4} e^{\lambda R^2/2} \|\Delta_\rho^{-1} f\|_{\dot{X}^{1/2}_\rho}.$$

A final application of the trivial inequality (14) yields the desired estimate. □

## 6. The new spaces

We must extend the domain of $\Delta_\rho^{-1}$ by taking limits, so we carefully define Banach spaces using equivalence classes. We define

$$\dot{X}^{1/2}_\rho(B) := \{[f]_B : f \in \dot{X}^{1/2}_\rho\},$$

where the equivalence class $[f]_B$ is given by

$$[f]_B := \{g \in \dot{X}^{1/2}_\rho : \operatorname{ess\,supp}(f - g) \subset \mathbb{R}^n \setminus B\}.$$

The space can be endowed with the norm

$$\|[f]_B\|_{\dot{X}^{1/2}_\rho(B)} := \inf\{\|g\|_{\dot{X}^{1/2}_\rho} : g \in [f]_B\},$$

so that

$$(\dot{X}^{1/2}_\rho(B), \|\bullet\|_{\dot{X}^{1/2}_\rho(B)}) \text{ is a Banach space.}$$

We can rephrase the inequality (17) in terms of this norm. Indeed, as

$$|\rho|^{1/2}\|g\|_{L^2(B)} \leqslant C\|g\|_{\dot{X}^{1/2}_\rho}$$

whenever $g \in [f]_B$, where $C > 1$ is a constant depending only on $R$, we can take the infimum to find

$$|\rho|^{1/2}\|f\|_{L^2(B)} \leqslant C\|f\|_{\dot{X}^{1/2}_\rho(B)}. \tag{23}$$

Identifying the elements $[f]_B$ of $\dot{X}^{1/2}_\rho(B)$ with $f|_B$, the restriction of $f$ to $B$, this yields the embedding

$$\dot{X}^{1/2}_\rho(B) \hookrightarrow L^2(B). \tag{24}$$

Moreover, we have the following equivalence of norms.

**6.1.** ***Equivalence with the Sobolev norm.*** There are constants $c, C > 0$, depending only on $R$, such that

$$c|\rho|^{-1/2}\|f\|_{H^1(B)} \leqslant \|f\|_{\dot{X}^{1/2}_\rho(B)} \leqslant C|\rho|^{1/2}\|f\|_{H^1(B)} \tag{25}$$

whenever $f \in H^1(B)$ and $|\rho| > 1$. To see this, note that

$$|m_\rho(\xi)| \leqslant 2(1+|\rho|)(1+|\xi|^2)$$

for all $\xi \in \mathbb{R}^n$, so that

$$\|g\|_{\dot{X}^{1/2}_\rho} \leqslant 2^{1/2}(1+|\rho|)^{1/2}\|g\|_{H^1(\mathbb{R}^n)}$$

whenever $g \in H^1(\mathbb{R}^n)$. Thus $H^1$-extensions are also $\dot{X}^{1/2}_\rho$-extensions, so the right-hand inequality of (25) follows by taking the infimum over $H^1$-extensions $g$ of $f \in H^1(B)$.

For the left-hand inequality, consider $g_B := \chi_B g$, where $\chi_B$ is a smooth function equal to one on $B$ and supported on $2B$. Then, separating the low and high frequencies,

$$\begin{aligned}\|g_B\|^2_{H^1(\mathbb{R}^n)} &\leqslant \|g_B\|^2_{L^2(\mathbb{R}^n)} + 16|\rho|^2 \int_{|\xi|\leqslant 4|\rho|} |\widehat{g_B}(\xi)|^2\,\mathrm{d}\xi + 2\int_{|\xi|>4|\rho|} |m_\rho(\xi)|\,|\widehat{g_B}(\xi)|^2\,\mathrm{d}\xi \\ &\leqslant C|\rho|\|g\|^2_{\dot{X}^{1/2}_\rho}\end{aligned}$$

whenever $|\rho| > 1$, where the second inequality follows from Lemma 2.2 of [Haberman and Tataru 2013]. Restricting the left-hand side to $B$, we find that

$$\|g\|_{H^1(B)} \leqslant C|\rho|^{1/2}\|g\|_{\dot{X}^{1/2}_\rho}.$$

Now if $g$ is an $\dot{X}^{1/2}_\rho$-extension of $f \in \dot{X}^{1/2}_\rho(B)$, then $f = g$ almost everywhere in $B$, so we can replace $g$ on the left-hand side by $f$ and take the infimum over $g$ to obtain the left-hand inequality of (25).

**6.2.** ***The main space.*** We define our main norm by

$$\|\bullet\|_{X^{1/2}_{\lambda,\rho}(B)} : f \in \dot{X}^{1/2}_{\rho}(B) \mapsto \lambda^{1/4}|\rho|^{1/2}\|f\|_{L^2(B,e^{\lambda(\theta\cdot x)^2})} + \frac{1}{\lambda^{1/4}}\|f\|_{\dot{X}^{1/2}_{\rho}(B)},$$

and note that by (23) it is equivalent to the homogeneous norm:

$$\lambda^{-1/4}\|f\|_{\dot{X}^{1/2}_{\rho}(B)} \leqslant \|f\|_{X^{1/2}_{\lambda,\rho}(B)} \leqslant C\lambda^{1/4}e^{\lambda R^2/2}\|f\|_{\dot{X}^{1/2}_{\rho}(B)}, \tag{26}$$

where $C > 1$ depends only on $R$. Thus we can conclude that

$$(\dot{X}^{1/2}_{\rho}(B), \|\bullet\|_{X^{1/2}_{\lambda,\rho}(B)}) \text{ is a Banach space.} \tag{27}$$

Later we will use that the constants in this norm equivalence are independent of $|\rho|$.

**6.3.** ***A minor variant of the main space.*** We also consider the norm $\|\bullet\|_{Y^{1/2}_{\lambda,-\rho}(B)}$ defined by

$$f \in \dot{X}^{1/2}_{-\rho}(B) \mapsto \max\left\{\lambda^{1/4}|\rho|^{1/2}\|f\|_{L^2(B,e^{-\lambda(\theta\cdot x)^2})}, \frac{1}{\lambda^{1/4}e^{\lambda R^2/2}}\|f\|_{\dot{X}^{1/2}_{-\rho}(B)}\right\}.$$

Notice that little more than some signs have changed. As before, this norm is equivalent to the homogeneous norm:

$$\frac{1}{\lambda^{1/4}e^{\lambda R^2/2}}\|f\|_{\dot{X}^{1/2}_{-\rho}(B)} \leqslant \|f\|_{Y^{1/2}_{\lambda,-\rho}(B)} \leqslant C\lambda^{1/4}\|f\|_{\dot{X}^{1/2}_{-\rho}(B)}, \tag{28}$$

where $C > 1$ depends only on $R$, and so

$$(\dot{X}^{1/2}_{-\rho}(B), \|\bullet\|_{Y^{1/2}_{\lambda,-\rho}(B)}) \text{ is a Banach space.} \tag{29}$$

Recalling the embedding (24), this can be identified with the intersection of the spaces

$$(L^2(B), \lambda^{1/4}|\rho|^{1/2}\|\bullet\|_{L^2(B,e^{-\lambda(\theta\cdot x)^2})}) \quad \text{and} \quad \left(\dot{X}^{1/2}_{-\rho}(B), \frac{1}{\lambda^{1/4}e^{\lambda R^2/2}}\|\bullet\|_{\dot{X}^{1/2}_{-\rho}(B)}\right).$$

As (29) is dense in both of these spaces, we can identify the dual of their intersection with the sum of their duals; see for example [Bennett 1974, Theorem 3.1]. This provides an alternative identification of the dual of (29) which we describe now.

**6.4.** ***The dual space.*** Let $\dot{X}^{-1/2}_{\rho,c}(B)$ denote the Banach completion of $C^\infty_c(B)$ with respect to the norm

$$\|\bullet\|_{\dot{X}^{-1/2}_{\rho}} : f \in C^\infty_c(B) \mapsto \||m_\rho|^{-1/2}\hat{f}\|_{L^2(\mathbb{R}^n)}.$$

We endow $L^2(B) + \dot{X}^{-1/2}_{\rho,c}(B)$ with the norm

$$\|f\|_{Y^{-1/2}_{\lambda,\rho,c}(B)} := \inf_{f=f^\flat+f^\sharp}\left(\frac{1}{\lambda^{1/4}|\rho|^{1/2}}\|f^\flat\|_{L^2(B,e^{\lambda(\theta\cdot x)^2})} + \lambda^{1/4}e^{\lambda R^2/2}\|f^\sharp\|_{\dot{X}^{-1/2}_{\rho}}\right),$$

with the infimum taken over all $f^\flat \in L^2(B)$ and $f^\sharp \in \dot{X}^{-1/2}_{\rho,c}(B)$. Then

$$(L^2(B) + \dot{X}^{-1/2}_{\rho,c}(B), \|\bullet\|_{Y^{-1/2}_{\lambda,\rho,c}(B)}) \text{ is a Banach space.} \tag{30}$$

With real-bracket pairings, Plancherel's identity takes the form

$$\langle f, g\rangle = \int_{\mathbb{R}^n} \hat{f}(\xi)\check{g}(\xi)\,\mathrm{d}\xi = \int_{\mathbb{R}^n} \hat{f}(\xi)\hat{g}(-\xi)\,\mathrm{d}\xi, \tag{31}$$

so that, by similar arguments to those used for Sobolev spaces, we find

$$\Big(\dot{X}^{1/2}_{-\rho}(B),\, \frac{1}{\lambda^{1/4}e^{\lambda R^2/2}}\|\bullet\|_{\dot{X}^{1/2}_{-\rho}(B)}\Big)^* \cong (\dot{X}^{-1/2}_{\rho,\mathrm{c}}(B),\, \lambda^{1/4}e^{\lambda R^2/2}\|\bullet\|_{\dot{X}^{-1/2}_{\rho}});$$

see for example [Jerison and Kenig 1995, Proposition 2.9]. On the other hand, it is easy to see that

$$(L^2(B),\, \lambda^{1/4}|\rho|^{1/2}\|\bullet\|_{L^2(B,e^{-\lambda(\theta\cdot x)^2})})^* \cong \Big(L^2(B),\, \frac{1}{\lambda^{1/4}|\rho|^{1/2}}\|\bullet\|_{L^2(B,e^{\lambda(\theta\cdot x)^2})}\Big).$$

Thus the dual of (29) can be identified with the sum of the two dual spaces as described in (30); see for example [Bennett 1974, Theorem 3.1].

## 7. The locally defined extension of $\Delta_\rho^{-1}$

We are now ready to extend the domain of $\Delta_\rho^{-1}$ by combining Lemmas 5.3 and 5.4. This extension will make no sense outside of $B$ in contrast with the globally defined extension of $f \in C^\infty_\mathrm{c}(B) \mapsto \Delta_\rho^{-1} f$ given by the trivial inequality (14). We denote the globally defined extension by $\Delta_\rho^{-1}$ and the locally defined extension by $\mathrm{T}^B_\rho$.

**Corollary 7.1.** *Consider $\rho \in \mathbb{C}^n$ such that $\rho\cdot\rho = 0$ and $\lambda > 1$. Then there is a continuous linear extension $\mathrm{T}^B_\rho$ of*

$$f \in C^\infty_\mathrm{c}(B) \mapsto \Delta_\rho^{-1} f|_B$$

*and a constant $C > 1$, depending only on the radius $R$ of $B$, such that*

$$\|\mathrm{T}^B_\rho f\|_{X^{1/2}_{\lambda,\rho}(B)} \leqslant C\|f\|_{Y^{-1/2}_{\lambda,\rho,\mathrm{c}}(B)}$$

*whenever $f \in L^2(B) + \dot{X}^{-1/2}_{\rho,\mathrm{c}}(B)$ and $|\rho| \geqslant 4\lambda R$.*

*Proof.* By Lemma 5.3 and the density of $C^\infty_\mathrm{c}(B)$ in

$$\Big(L^2(B),\, \frac{1}{\lambda^{1/4}|\rho|^{1/2}}\|\bullet\|_{L^2(B,e^{\lambda(\theta\cdot x)^2})}\Big),$$

we can extend $f \in C^\infty_\mathrm{c}(B) \mapsto \Delta_\rho^{-1} f|_B$ to a bounded linear operator $\mathrm{T}^B_\rho$ that satisfies

$$\|\mathrm{T}^B_\rho f\|_{X^{1/2}_{\lambda,\rho}(B)} \leqslant \frac{C}{\lambda^{1/4}|\rho|^{1/2}}\|f\|_{L^2(B,e^{\lambda(\theta\cdot x)^2})}$$

whenever $f \in L^2(B)$. The constant $C > 1$ depends only on $R$. On the other hand, by Lemma 5.4 and the density of $C^\infty_\mathrm{c}(B)$ in

$$(\dot{X}^{-1/2}_{\rho,\mathrm{c}}(B),\, \lambda^{1/4}e^{\lambda R^2/2}\|\bullet\|_{\dot{X}^{-1/2}_{\rho}}),$$

we can extend $f \in C_c^\infty(B) \mapsto \Delta_\rho^{-1} f|_B$ to a bounded linear operator $\mathrm{T}_\rho^B$ that satisfies

$$\|\mathrm{T}_\rho^B f\|_{X_{\lambda,\rho}^{1/2}(B)} \leqslant C\lambda^{1/4} e^{\lambda R^2/2} \|f\|_{\dot{X}_\rho^{-1/2}}$$

whenever $f \in \dot{X}_{\rho,\mathrm{c}}^{-1/2}(B)$. Again, the constant $C > 1$ depends only on $R$.

Considering now $f = f^\flat + f^\sharp$ with $f^\flat \in L^2(B)$ and $f^\sharp \in \dot{X}_{\rho,\mathrm{c}}^{-1/2}(B)$, we define

$$\mathrm{T}_\rho^B f := \mathrm{T}_\rho^B f^\flat + \mathrm{T}_\rho^B f^\sharp.$$

One can show that this is well defined using the linearity of the previous extensions and the density of $C_c^\infty(B)$. Then, by the triangle inequality and the previous bounds,

$$\|\mathrm{T}_\rho^B f\|_{X_{\lambda,\rho}^{1/2}(B)} \leqslant C\left(\frac{1}{\lambda^{1/4}|\rho|^{1/2}} \|f^\flat\|_{L^2(e^{\lambda(\theta\cdot x)^2})} + \lambda^{1/4} e^{\lambda R^2/2} \|f^\sharp\|_{\dot{X}_\rho^{-1/2}}\right),$$

where the constant $C$ depends only on $R$. Since the left-hand side is independent of the representation $f = f^\flat + f^\sharp$, we can take the infimum over such representations, and the desired inequality follows. □

## 8. The bound for $\mathrm{M}_q$

With a view to further applications, we write part of this section in greater generality. Consider bounded functions $a_0, a_1, \ldots, a_n \subset L^\infty(\mathbb{R}^n)$ with compact support:

$$\operatorname{supp} a_j \subset \Omega \subset B = \{x \in \mathbb{R}^n : |x| < R\},$$

where $R := 2\sup_{x\in\Omega} |x|$. Define the bilinear form $\mathcal{B} : H^1(B) \times H^1(B) \to \mathbb{C}$ by

$$\mathcal{B}(f, g) := \int_\Omega a_0 f g + \int_\Omega A \cdot \nabla(f g),$$

where $A$ is the vector field with components $(a_1, \ldots, a_n)$. This is well defined by an application of the product rule, followed by the Cauchy–Schwarz inequality.

**Proposition 8.1.** *Consider $\rho \in \mathbb{C}^n$ such that $\rho\cdot\rho = 0$ and $\lambda > 1$. Then there is a constant $C > 1$, depending only on the radius $R$ of $B$, such that*

$$|\mathcal{B}(f, g)| \leqslant C\left(\frac{1}{\lambda^{1/2}|\rho|} + \frac{1}{\lambda^{1/2}} + \frac{e^{\lambda R^2/2}}{|\rho|^{1/2}}\right) \sum_{j=0}^n \|a_j\|_{L^\infty(\Omega)} \|f\|_{X_{\lambda,\rho}^{1/2}(B)} \|g\|_{Y_{\lambda,-\rho}^{1/2}(B)}$$

*whenever $(f, g) \in \dot{X}_\rho^{1/2}(B) \times \dot{X}_{-\rho}^{1/2}(B)$.*

*Proof.* For the first term, we note that, by the Cauchy–Schwarz inequality,

$$\left|\int_\Omega a_0 f g\right| \leqslant \|a_0\|_\infty \|e^{\lambda(\theta\cdot x)^2/2} f\|_{L^2(B)} \|e^{-\lambda(\theta\cdot x)^2/2} g\|_{L^2(B)} \leqslant \frac{1}{\lambda^{1/2}} \frac{1}{|\rho|} \|a_0\|_\infty \|f\|_{X_{\lambda,\rho}^{1/2}(B)} \|g\|_{Y_{\lambda,-\rho}^{1/2}(B)}$$

whenever $(f, g) \in \dot{X}_\rho^{1/2}(B) \times \dot{X}_{-\rho}^{1/2}(B)$. The second inequality follows directly from the weightings in the definition of the norms.

For the more difficult first-order term, we consider a positive and smooth function $\chi$, equal to 1 on the ball of radius $\frac{1}{2}$, supported in the unit ball, and bounded above by 1. Then we work with $f_B := \chi_B f$ and $g_B := \chi_B g$, where $\chi_B := \chi(\bullet/R)$ is equal to 1 on $\Omega$ and supported on $B$. Letting $A^\flat$ denote the vector field with components

$$a_j^\flat(x) := \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\chi\Big(\frac{\xi}{16|\rho|}\Big)\hat{a}_j(\xi)\,\mathrm{d}\xi$$

for all $x \in \mathbb{R}^n$ and $j = 1, \ldots, n$, and letting $A^\sharp := A - A^\flat$, by integration by parts,

$$\int_\Omega A\cdot\nabla(fg) = -\int_{\mathbb{R}^n}\nabla\cdot A^\flat f_B g_B + \int_{\mathbb{R}^n} A^\sharp\cdot\nabla(f_B g_B).$$

Noting that $\|\nabla\cdot A^\flat\|_\infty \leqslant C|\rho|\|A\|_\infty$, the first term can be bounded as before:

$$\begin{aligned}\Big|\int_{\mathbb{R}^n}\nabla\cdot A^\flat f_B g_B\Big| &\leqslant C\|\nabla\cdot A^\flat\|_\infty\|e^{\lambda(\theta\cdot x)^2/2}f\|_{L^2(B)}\|e^{-\lambda(\theta\cdot x)^2/2}g\|_{L^2(B)}\\ &\leqslant C\|A\|_\infty|\rho|^{1/2}\|f\|_{L^2(B,e^{\lambda(\theta\cdot x)^2})}|\rho|^{1/2}\|g\|_{L^2(B,e^{-\lambda(\theta\cdot x)^2})}.\end{aligned}$$

Again by the weightings in the definitions of the norms, this implies that

$$\Big|\int_{\mathbb{R}^n}\nabla\cdot A^\flat f_B g_B\Big| \leqslant C\frac{1}{\lambda^{1/2}}\|A\|_\infty\|f\|_{X^{1/2}_{\lambda,\rho}(B)}\|g\|_{Y^{1/2}_{\lambda,-\rho}(B)}$$

whenever $(f, g) \in \dot{X}^{1/2}_\rho(B) \times \dot{X}^{1/2}_{-\rho}(B)$.

It remains to show that

$$\Big|\int_{\mathbb{R}^n} A^\sharp\cdot\nabla(f_B g_B)\Big| \leqslant C\frac{e^{\lambda R^2/2}}{|\rho|^{1/2}}\|A\|_\infty\|f\|_{X^{1/2}_{\lambda,\rho}(B)}\|g\|_{Y^{1/2}_{\lambda,-\rho}(B)}. \tag{32}$$

Using the product rule, we can separate into two similar terms,

$$\int_{\mathbb{R}^n} A^\sharp\cdot\nabla(f_B g_B) = \int_{\mathbb{R}^n} A^\sharp\cdot\nabla f_B g_B + \int_{\mathbb{R}^n} A^\sharp\cdot\nabla g_B f_B, \tag{33}$$

and initially treat the first term on the right-hand side (the second term will eventually be dealt with by symmetry). We decompose the integral as

$$\int_{\mathbb{R}^n} A^\sharp\cdot\nabla f_B g_B = \int_{\mathbb{R}^n} A^\sharp\cdot\nabla L f_B L g_B + \int_{\mathbb{R}^n} A^\sharp\cdot\nabla L f_B H g_B + \int_{\mathbb{R}^n} A^\sharp\cdot\nabla H f_B g_B,$$

where $L$ denotes the low-frequency filter defined by

$$Lf(x) := \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\chi\Big(\frac{\xi}{4|\rho|}\Big)\hat{f}(\xi)\,\mathrm{d}\xi$$

and $H := I - L$. By the properties of $\chi$, the frequency supports of $\nabla L f_B L g_B$ and $A^\sharp$ are disjoint, so that by Plancherel's identity the first term is in fact 0, yielding

$$\int_{\mathbb{R}^n} A^\sharp\cdot\nabla f_B g_B = \int_{\mathbb{R}^n} A^\sharp\cdot\nabla L f_B H g_B + \int_{\mathbb{R}^n} A^\sharp\cdot\nabla H f_B g_B.$$

Then, by the Cauchy–Schwarz inequality (writing $\|\bullet\|_2 := \|\bullet\|_{L^2(\mathbb{R}^n)}$),

$$\left|\int_{\mathbb{R}^n} A^\sharp \cdot \nabla f_B g_B\right| \leqslant \|A^\sharp\|_\infty (\|\nabla L f_B\|_2 \|H g_B\|_2 + \|\nabla H f_B\|_2 \|g_B\|_2).$$

Now as $\|A^\sharp\|_\infty \leqslant C\|A\|_\infty$ and

$$\|\nabla L f_B\|_2 \|H g_B\|_2 \leqslant C|\rho| \|L f_B\|_2 \|H g_B\|_2 \leqslant C\|f_B\|_2 \|\nabla H g_B\|_2,$$

we find that

$$\left|\int_{\mathbb{R}^n} A^\sharp \cdot \nabla f_B g_B\right| \leqslant C\|A\|_\infty (\|f_B\|_2 \|\nabla H g_B\|_2 + \|\nabla H f_B\|_2 \|g_B\|_2).$$

Since the right-hand side is symmetric in the roles of $f_B$ and $g_B$, we can conclude the same bound for the second term on the right-hand side of (33), yielding

$$\left|\int_{\mathbb{R}^n} A^\sharp \cdot \nabla (f_B g_B)\right| \leqslant C\|A\|_\infty (\|f_B\|_2 \|\nabla H g_B\|_2 + \|\nabla H f_B\|_2 \|g_B\|_2). \tag{34}$$

Now clearly we have that

$$\|f_B\|_2 \leqslant \|f\|_{L^2(B, e^{\lambda(\theta\cdot x)^2})} \quad \text{and} \quad \|g_B\|_2 \leqslant e^{\lambda R^2/2} \|g\|_{L^2(B, e^{-\lambda(\theta\cdot x)^2})}.$$

On the other hand, by Lemma 2.2 of [Haberman and Tataru 2013], we have

$$\|\nabla H f_B\|_2 \leqslant C\|\tilde{f}\|_{\dot{X}^{1/2}_\rho} \quad \text{and} \quad \|\nabla H g_B\|_2 \leqslant C\|\tilde{g}\|_{\dot{X}^{1/2}_{-\rho}},$$

where $(\tilde{f}, \tilde{g}) \in \dot{X}^{1/2}_\rho \times \dot{X}^{1/2}_{-\rho}$ denotes any pair of extensions of $(f, g)$. Substituting these inequalities into (34) and taking the infimum over extensions yields

$$\left|\int_{\mathbb{R}^n} A^\sharp \cdot \nabla (f_B g_B)\right| \leqslant C\|A\|_\infty (\|f\|_{L^2(B, e^{\lambda(\theta\cdot x)^2})} \|g\|_{\dot{X}^{1/2}_{-\rho}(B)} + e^{\lambda R^2/2} \|f\|_{\dot{X}^{1/2}_\rho(B)} \|g\|_{L^2(B, e^{-\lambda(\theta\cdot x)^2})}).$$

Recalling the weightings in the norms, this completes the proof of (32). □

From this we can deduce our estimate for $\mathrm{M}_q : f \mapsto qf$, where $q$ is defined in (4).

**Corollary 8.2.** *Consider $\rho \in \mathbb{C}^n$ such that $\rho \cdot \rho = 0$ and $\lambda > 1$. Then there is a $C > 1$, depending on $\|\nabla \log \sigma\|_\infty$ and the radius $R$ of $B$, such that*

$$\|\mathrm{M}_q f\|_{Y^{-1/2}_{\lambda,\rho,\mathrm{c}}(B)} \leqslant C\left(\frac{1}{\lambda^{1/2}|\rho|} + \frac{1}{\lambda^{1/2}} + \frac{e^{\lambda R^2/2}}{|\rho|^{1/2}}\right) \|f\|_{X^{1/2}_{\lambda,\rho}(B)}$$

*whenever $f \in \dot{X}^{1/2}_\rho(B)$.*

*Proof.* By an application of the product rule, the definition (4) can be rewritten as

$$\langle q, \psi\rangle = \frac{1}{4}\int_\Omega |\nabla \log \sigma|^2 \psi - \frac{1}{2}\int_\Omega \nabla \log \sigma \cdot \nabla \psi.$$

Using our a priori assumptions that $\sigma$ is bounded below and $\nabla\sigma$ is bounded above almost everywhere (which follows from Lipschitz continuity), $\nabla\log\sigma = \sigma^{-1}\nabla\sigma$ is a vector of bounded functions. Thus, by taking

$$a_0 = \tfrac{1}{4}|\nabla\log\sigma|^2 \quad \text{and} \quad A = -\tfrac{1}{2}\nabla\log\sigma,$$

we can write

$$\langle \mathrm{M}_q\, f, g\rangle := \langle qf, g\rangle := \langle q, fg\rangle = \mathcal{B}(f, g)$$

for all $(f, g) \in \dot{X}^{1/2}_{\rho}(B) \times \dot{X}^{1/2}_{-\rho}(B)$. Then, by Proposition 8.1, we find that

$$|\langle \mathrm{M}_q\, f, g\rangle| \leqslant C\left(\frac{1}{\lambda^{1/2}|\rho|} + \frac{1}{\lambda^{1/2}} + \frac{e^{\lambda R^2/2}}{|\rho|^{1/2}}\right)\sum_{j=0}^{n}\|a_j\|_\infty\|f\|_{X^{1/2}_{\lambda,\rho}(B)}\|g\|_{Y^{1/2}_{\lambda,-\rho}(B)}$$

for all $(f, g) \in \dot{X}^{1/2}_{\rho}(B) \times \dot{X}^{1/2}_{-\rho}(B)$. Finally, using the identification

$$(\dot{X}^{1/2}_{-\rho}(B), \|\bullet\|_{Y^{1/2}_{\lambda,-\rho}(B)})^* \cong (L^2(B) + \dot{X}^{-1/2}_{\rho,\mathrm{c}}(B), \|\bullet\|_{Y^{-1/2}_{\lambda,\rho,\mathrm{c}}(B)}),$$

we obtain the desired inequality. □

## 9. Locally defined CGO solutions via Neumann series

Let $X^{1/2}_{\lambda,\rho}(B)$ and $Y^{-1/2}_{\lambda,\rho,\mathrm{c}}(B)$ denote the Banach spaces defined in (27) and (30), respectively. Recall that $f \in C^\infty_{\mathrm{c}}(B) \mapsto \Delta^{-1}_{\rho} f|_B$ can be extended as a bounded linear operator

$$\mathrm{T}^B_\rho : Y^{-1/2}_{\lambda,\rho,\mathrm{c}}(B) \to X^{1/2}_{\lambda,\rho}(B)$$

using Corollary 7.1 and that $\mathrm{M}_q : f \mapsto qf$, with $q$ defined in (4), is bounded as

$$\mathrm{M}_q : X^{1/2}_{\lambda,\rho}(B) \to Y^{-1/2}_{\lambda,\rho,\mathrm{c}}(B)$$

by Corollary 8.2. The contraction will follow by choosing $|\rho|$ and $\lambda$ appropriately so that the product of the operator norms is small.

**Theorem 9.1.** *Consider $\rho \in \mathbb{C}^n$ such that $\rho\cdot\rho = 0$ and $\lambda > 1$. Then there is a $C_0 > 1$, depending on $\|\nabla\log\sigma\|_\infty$ and the radius $R$ of $B$, such that*

$$\|\mathrm{T}^B_\rho \circ \mathrm{M}_q\|_{\mathcal{L}(X^{1/2}_{\lambda,\rho}(B))} \leqslant \tfrac{1}{2} \tag{35}$$

*whenever $|\rho| > \lambda e^{\lambda R^2}$ and $\lambda = 36C_0^2$. For all $f \in Y^{-1/2}_{\lambda,\rho,\mathrm{c}}(B)$, there is a $w \in X^{1/2}_{\lambda,\rho}(B)$ such that*

$$(\mathrm{I} - \mathrm{T}^B_\rho \circ \mathrm{M}_q)w = \mathrm{T}^B_\rho[f]. \tag{36}$$

*Moreover, there is a $C > 1$, depending only on $R$, such that*

$$\|w\|_{X^{1/2}_{\lambda,\rho}(B)} \leqslant C\|f\|_{Y^{-1/2}_{\lambda,\rho,\mathrm{c}}(B)}. \tag{37}$$

*Proof.* By combining Corollaries 7.1 and 8.2, we have that $\mathrm{T}^B_\rho \circ \mathrm{M}_q$ is a bounded operator whenever $|\rho| \geqslant 4\lambda R$. Furthermore, there exists a constant $C_0 > 1$ such that

$$\|\mathrm{T}^B_\rho \circ \mathrm{M}_q\|_{\mathcal{L}(X^{1/2}_{\lambda,\rho}(B))} \leqslant C_0\left(\frac{1}{\lambda^{1/2}|\rho|} + \frac{1}{\lambda^{1/2}} + \frac{e^{\lambda R^2/2}}{|\rho|^{1/2}}\right) \leqslant \frac{1}{2}$$

whenever $|\rho|^{1/2} > 6C_0 e^{\lambda R^2/2}$ and $\lambda^{1/2} = 6C_0$. Then, by Neumann series, $\mathrm{I} - \mathrm{T}^B_\rho \circ \mathrm{M}_q$ has a bounded inverse,

$$(\mathrm{I} - \mathrm{T}^B_\rho \circ \mathrm{M}_q)^{-1} = \sum_{k\geqslant 0} (\mathrm{T}^B_\rho \circ \mathrm{M}_q)^k$$

on $X^{1/2}_{\lambda,\rho}(B)$, and so $w = (\mathrm{I} - \mathrm{T}^B_\rho \circ \mathrm{M}_q)^{-1}\mathrm{T}^B_\rho[f]$ satisfies (36). Moreover,

$$\|w\|_{X^{1/2}_{\lambda,\rho}(B)} \leqslant \sum_{k\geqslant 0} \|(\mathrm{T}^B_\rho \circ \mathrm{M}_q)^k \mathrm{T}^B_\rho[f]\|_{X^{1/2}_{\lambda,\rho}(B)} \leqslant 2\|\mathrm{T}^B_\rho[f]\|_{X^{1/2}_{\lambda,\rho}(B)}$$

by the triangle inequality, the contraction (35), and summing the geometric series. Then (37) follows by a final application of Corollary 7.1. ☐

Recall that we can also use the trivial inequality (14) to extend $f \in C^\infty_{\mathrm{c}}(B) \mapsto \Delta^{-1}_\rho f$ as a bounded linear operator

$$\Delta^{-1}_\rho : \dot{X}^{-1/2}_\rho \to \dot{X}^{1/2}_\rho.$$

In the following corollary we clarify that the restriction of this extension to the ball $B$ and the previous locally defined extension $\mathrm{T}^B_\rho$ are the same. We also record the properties of our CGO solutions that we will need in the remaining sections.

**Corollary 9.2.** *Consider $\rho \in \mathbb{C}^n$ and $\lambda > 1$ as in Theorem 9.1. Then*

$$\|\Delta^{-1}_\rho \circ \mathrm{M}_q\|_{\mathcal{L}(X^{1/2}_{\lambda,\rho}(B))} \leqslant \tfrac{1}{2}, \tag{38}$$

*there is a $w \in H^1(B)$ that solves* (7), *and there is a $C > 1$, depending on $\|\nabla \log \sigma\|_\infty$ and the radius $R$ of $B$, such that*

$$\|w\|_{\dot{X}^{1/2}_\rho(B)} \leqslant C\|q\|_{\dot{X}^{-1/2}_\rho}. \tag{39}$$

*Moreover, $v = e_\rho(1 + w) \in H^1(B)$ solves the Lippmann–Schwinger-type equation*

$$(\mathrm{I} - \mathrm{S}_q)v = e_\rho, \quad \text{where } \mathrm{S}_q := e_\rho \Delta^{-1}_\rho \circ \mathrm{M}_q[e_{-\rho}\,\bullet] \tag{40}$$

*as elements of $H^1(B)$, and is also a weak solution to the Schrödinger equation* (3).

*Proof.* By Corollary 7.1, the equivalence of norms (26), and the trivial inequality (14),

$$\begin{aligned}\|\mathrm{T}^B_\rho g - \Delta^{-1}_\rho g\|_{X^{1/2}_{\lambda,\rho}(B)} &\leqslant \|\mathrm{T}^B_\rho[g - g_j]\|_{X^{1/2}_{\lambda,\rho}(B)} + \|\Delta^{-1}_\rho[g_j - g]\|_{X^{1/2}_{\lambda,\rho}(B)} \\ &\leqslant C(\|g - g_j\|_{Y^{-1/2}_{\lambda,\rho,\mathrm{c}}(B)} + \|g_j - g\|_{\dot{X}^{-1/2}_\rho}),\end{aligned}$$

and given that the dual norms are also equivalent, by (28), we can choose $g_j \in C_c^\infty(B)$ such that the right-hand side converges to 0. Then, combining with Corollary 8.2, the contraction (38) follows directly from the previous contraction (35).

Taking $f = \mathrm{M}_q[1]$ in Theorem 9.1, we find $w \in X^{1/2}_{\lambda,\rho}(B)$ solving

$$w = \mathrm{T}^B_\rho \circ \mathrm{M}_q[1+w].$$

Again by Corollary 8.2, we have $\mathrm{M}_q[1+w] \in Y^{-1/2}_{\lambda,\rho,\mathrm{c}}(B)$, so that, taking this as the function $g$ above, we can also write

$$w = \Delta_\rho^{-1} \circ \mathrm{M}_q[1+w] \tag{41}$$

as elements of $X^{1/2}_{\lambda,\rho}(B)$. Thus, combining with the norm equivalences (25) and (26), we find that $w \in H^1(B)$ solves (7). Moreover, the inequality (39) follows from the previous inequality (37) combined with (26) and the dual version of (28).

Finally, writing $v = e_\rho(1+w)$, we can multiply (41) by $e_\rho$ to find

$$v - e_\rho = e_\rho \Delta_\rho^{-1} \circ \mathrm{M}_q[e_{-\rho} v] =: \mathrm{S}_q[v]$$

as elements of $H^1(B)$. Then, by integration by parts and Plancherel's identity (31), cancelling the Fourier multipliers,

$$\begin{aligned} -\int_{\mathbb{R}^n} \nabla \mathrm{S}_q[v] \cdot \nabla \psi &= \int_{\mathbb{R}^n} \Delta_\rho^{-1} \circ \mathrm{M}_q[e_{-\rho} v] e_\rho \Delta[e_{-\rho} e_\rho \psi] \\ &= \int_{\mathbb{R}^n} m_\rho^{-1} \widehat{\mathrm{M}_q[e_{-\rho} v]} m_\rho (e_\rho \psi)^\vee \\ &= \langle q v, \psi \rangle \end{aligned} \tag{42}$$

whenever $\psi \in C_c^\infty(B)$. Given that $e_\rho$ is harmonic, we see that $v \in H^1(B)$ is also a weak solution to the Schrödinger equation (3). □

**Remark 9.3.** The CGO solutions $v = e_\rho(1+w)$ given by Corollary 9.2 also satisfy

$$v = \sum_{k \geqslant 0} e_\rho (\Delta_\rho^{-1} \circ \mathrm{M}_q)^k [1], \tag{43}$$

with convergence in $H^1(B)$. On the other hand, we have that

$$e_\rho (\Delta_\rho^{-1} \circ \mathrm{M}_q)^k [1] = (e_\rho \Delta_\rho^{-1} \circ \mathrm{M}_q[e_{-\rho} \bullet])^k [e_\rho] = \mathrm{S}_q^k [e_\rho]$$

as elements of $H^1(B)$. Substituting this into (43), we find that

$$v = \sum_{k \geqslant 0} \mathrm{S}_q^k [e_\rho]$$

again in the $H^1(B)$-sense. If we had proven that $\mathrm{S}_q$ is contractive on $H^1(B)$, we could have solved (40) more directly by Neumann series, and the solution would have taken this form.

## 10. The boundary integral identities

Here we use the divergence theorem to equate the boundary integral to an integral over the domain. Identities similar to the first identity of Lemma 10.1, often known as Alessandrini identities, are foundational for the Calderón problem. Recall that our boundary integral $BI_{\Lambda_\sigma}: H^{1/2}(\partial\Omega)\times H^1(\Omega)\to\mathbb{C}$ is defined by

$$BI_{\Lambda_\sigma}(\phi,\psi):=\int_{\partial\Omega}(\sigma^{-1/2}\Lambda_\sigma[\sigma^{-1/2}\phi]-\nu\cdot\nabla P_0[\phi])\psi, \tag{44}$$

where $P_0[\phi]$ denotes the harmonic extension of $\phi$. A key idea of [Nachman 1988] and [Novikov 1988] was to take the Faddeev fundamental solution within boundary integrals similar to this, yielding similar formulas to the second identity in Lemma 10.1.

**Lemma 10.1.** *Let $q$ be defined by* (4), *and let $BI_{\Lambda_\sigma}$ be defined by* (44). *Then*

$$BI_{\Lambda_\sigma}(v|_{\partial\Omega},\psi)=\langle qv,\psi\rangle$$

*whenever $\psi$ is harmonic on $\Omega$ and $v\in H^1(\Omega)$ solves the Schrödinger equation* (3). *Moreover*,

$$BI_{\Lambda_\sigma}(v|_{\partial\Omega},G_\rho(\mathrm{x},\bullet))=\mathrm{S}_q[v](\mathrm{x})$$

*as elements of $H^1(B\setminus\bar{\Omega})$, where $\mathrm{S}_q$ is defined in* (40).

*Proof.* For the first identity, consider the weak solution to the conductivity equation given by $u=\sigma^{-1/2}v$. Recalling that $\nabla\sigma$ is bounded almost everywhere, by an application of the product rule, we find that $\Delta u=-\sigma^{-1}\nabla\sigma\cdot\nabla u\in L^2(\Omega)$. Thus the normal traces can be defined so that the divergence theorem can be applied to $\sigma\nabla u\sigma^{-1/2}\psi-\nabla P_0[\phi]\psi$, yielding

$$BI_{\Lambda_\sigma}(v|_{\partial\Omega},\psi)=\int_\Omega(\sigma\nabla u\cdot\nabla(\sigma^{-1/2}\psi)-\nabla P_0[v|_{\partial\Omega}]\cdot\nabla\psi); \tag{45}$$

see for example [Kim and Kwon 2022, Proposition 2.4]. Now, as $\psi$ is harmonic on $\Omega$, we have

$$\int_\Omega\nabla(P_0[v|_{\partial\Omega}]-\sigma^{1/2}u)\cdot\nabla\psi=0,$$

which can be substituted in (45) to find that

$$BI_{\Lambda_\sigma}(v|_{\partial\Omega},\psi)=\int_\Omega(\sigma\nabla u\cdot\nabla(\sigma^{-1/2}\psi)-\nabla(\sigma^{1/2}u)\cdot\nabla\psi).$$

Then, after applying the product rule again, terms cancel and one finds that the right-hand side of this identity is equal to $\langle q\sigma^{1/2}u,\psi\rangle=\langle qv,\psi\rangle$, as desired.

For the second identity, recall that $G_{-\rho}:=e_{-\rho}F_{-\rho}$ is a fundamental solution for the Laplacian. In particular $\Delta G_{-\rho}(\bullet,x)=0$ on $\Omega$ for all $x\in B\setminus\bar{\Omega}$. On the other hand, $G_\rho$ inherits a skew symmetry from $F_\rho$,

$$G_{-\rho}(y,x):=e_{-\rho}(y-x)F_{-\rho}(y-x)=e_\rho(x-y)F_\rho(x-y)=:G_\rho(x,y), \tag{46}$$

so we can reinterpret this as $\Delta G_\rho(x,\bullet)=0$ on $\Omega$ for all $x\in B\setminus\overline{\Omega}$. Thus, we can substitute this into the first identity to find that

$$BI_{\Lambda_\sigma}(v|_{\partial\Omega}, G_\rho(x,\bullet)) = \langle qv, G_\rho(x,\bullet)\rangle \tag{47}$$

for all $x\in B\setminus\overline{\Omega}$.

Now, for any $f\in H^1(\Omega)$ and any smooth $\psi_x$ supported in a small ball centred at $x$ and properly contained in $B\setminus\overline{\Omega}$, we have that

$$\int_B \langle qf, G_\rho(y,\bullet)\rangle \psi_x(y)\,\mathrm{d}y = \left\langle qf, \int_B G_\rho(y,\bullet)\psi_x(y)\,\mathrm{d}y\right\rangle. \tag{48}$$

This follows by interchanging the integral and the gradient, using Lebesgue's dominated convergence theorem, and applying Fubini's theorem. Then using the skew symmetry (46) again and the kernel representation (9) of $\Delta_\rho^{-1}$, the right-hand side of (48) can be rewritten as

$$\langle qf, e_{-\rho}\Delta_{-\rho}^{-1}[e_\rho\psi_x]\rangle = \int_B e_\rho\Delta_\rho^{-1}[qfe_{-\rho}](y)\psi_x(y)\,\mathrm{d}y. \tag{49}$$

Here we have considered $\Delta_\rho^{-1}$ to be the globally defined extension given by (14), and the identity follows by moving the Fourier multiplier $m_\rho^{-1}$ from one term to the other after an application of Plancherel's identity (31). Combining (48) with (49) and recalling the definition (40) of $\mathrm{S}_q$, we find that

$$\int_B \langle qf, G_\rho(y,\bullet)\rangle \psi_x(y)\,\mathrm{d}y = \int_B \mathrm{S}_q[f](y)\psi_x(y)\,\mathrm{d}y.$$

Now by the bounds of the previous section, we have that $\mathrm{S}_q[f]\in H^1(B)$, and so, letting $\psi_x$ approximate the Dirac delta $\delta_x$, we find that

$$\langle qf, G_\rho(x,\bullet)\rangle = \mathrm{S}_q[f](x) \tag{50}$$

for almost every $x\in B\setminus\overline{\Omega}$ by a suitable version of the Lebesgue differentiation theorem; see for example [Muscalu and Schlag 2013, Theorem 2.12]. Taking $f=v$ and combining (47) with (50) yields the second identity. □

## 11. The proofs of Theorems 3.1 and 3.2

The second identity of Lemma 10.1 allows us to define $\Gamma_{\Lambda_\sigma}: H^{1/2}(\partial\Omega)\to H^{1/2}(\partial\Omega)$ by taking the outer trace $\mathrm{T}_{\partial\Omega}: H^1(B\setminus\overline{\Omega})\to H^{1/2}(\partial\Omega)$ of the boundary integral:

$$\Gamma_{\Lambda_\sigma}[\phi] := \mathrm{T}_{\partial\Omega}[BI_{\Lambda_\sigma}(\phi, G_\rho(\mathrm{x},\bullet))] \tag{51}$$

for all $\phi\in H^{1/2}(\partial\Omega)$. Moreover, it gives us the alternative representation

$$\Gamma_{\Lambda_\sigma}[\phi] = \mathrm{T}_{\partial\Omega}\circ \mathrm{S}_q\circ P_q[\phi], \tag{52}$$

where $\mathrm{S}_q$ is defined in (40) and $P_q[\phi]$ denotes the solution to (3) with Dirichlet data $\phi$.

We restate the main theorems from Section 3 before proving them. The proof of the second part of the following theorem bears some resemblance to the argument of [Astala et al. 2016, Theorem 3.1], allowing us to avoid the use of double layer potentials.

**Theorem 3.1.** *Consider $\rho \in \mathbb{C}^n$ such that $\rho \cdot \rho = 0$ and $|\rho|^2 = \rho \cdot \bar{\rho}$ is sufficiently large. Let $\Gamma_{\Lambda_\sigma}$ be defined by* (12). *Then*

(i) $\Gamma_{\Lambda_\sigma} : H^{1/2}(\partial\Omega) \to H^{1/2}(\partial\Omega)$ *is bounded compactly,*

(ii) *if $\Gamma_{\Lambda_\sigma}[\phi] = \phi$, then $\phi = 0$,*

(iii) $\mathrm{I} - \Gamma_{\Lambda_\sigma}$ *has a bounded inverse on $H^{1/2}(\partial\Omega)$,*

*and if $v = e_\rho(1 + w)$, where $w \in H^1(B)$ is a solution to* (7), *then*

(iv) $v|_{\partial\Omega} = (\mathrm{I} - \Gamma_{\Lambda_\sigma})^{-1}[e_\rho|_{\partial\Omega}]$.

*Proof.* By hypothesis $(\mathrm{I} - \mathrm{S}_q)v = e_\rho$, so part (iv) follows from the alternative representation (52) and part (iii), which in turn will follow from parts (i) and (ii) by the Fredholm alternative.

To see (i), note first that the trace operator $\mathrm{T}_{\partial\Omega} : H^1(B \setminus \overline{\Omega}) \to H^{1/2}(\partial\Omega)$ and solution operator $P_q : H^{1/2}(\partial\Omega) \to H^1(\Omega)$ are bounded. Combining this with the alternative representation (52), it will suffice to show that $\mathrm{S}_q : H^1(\Omega) \to H^1(B \setminus \overline{\Omega})$ is bounded compactly. For this we recall that, on $B \setminus \overline{\Omega}$, we have the representation (50), and so by applications of the product rule we can divide the operator into three parts $\mathrm{S}_q = \mathrm{S}_1 + \mathrm{S}_2 + \mathrm{S}_3$, where

$$\mathrm{S}_1[f] := \frac{1}{4} \int_\Omega |\nabla \log \sigma(y)|^2 f(y) G_\rho(\bullet - y)\, \mathrm{d}y,$$
$$\mathrm{S}_2[f] := -\frac{1}{2} \int_\Omega \nabla \log \sigma(y) \cdot \nabla f(y) G_\rho(\bullet - y)\, \mathrm{d}y,$$
$$\mathrm{S}_3[f] := \frac{1}{2} \int_\Omega \nabla \log \sigma(y) \cdot \nabla G_\rho(\bullet - y) f(y)\, \mathrm{d}y.$$

By our a priori assumptions, $\nabla \log \sigma = \sigma^{-1}\nabla\sigma \in L^\infty(\Omega)^n$, and on the other hand $G_\rho$ and $\nabla G_\rho$ are locally integrable by (10). Thus, by Young's convolution inequality,

$$\mathrm{S}_1 : L^2(\Omega) \to L^2(B \setminus \overline{\Omega}), \quad \mathrm{S}_2 : H^1(\Omega) \to L^2(B \setminus \overline{\Omega}), \quad \text{and} \quad \mathrm{S}_3 : L^2(\Omega) \to L^2(B \setminus \overline{\Omega})$$

are bounded. Moreover, by Lebesgue's dominated convergence theorem, we can take derivatives under the integral, and by (10) we have that

$$\partial_{\mathrm{x}_j} \partial_{\mathrm{x}_i} G_\rho(\mathrm{x} - \mathrm{y}) = c_n n \frac{(\mathrm{x}_j - \mathrm{y}_j)(\mathrm{x}_i - \mathrm{y}_i)}{|\mathrm{x} - \mathrm{y}|^{n+2}} + \partial_{\mathrm{x}_j} \partial_{\mathrm{x}_i} H_\rho(\mathrm{x} - \mathrm{y}).$$

On the one hand, the second-order Riesz transforms are easily bounded in $L^2$ noting that the Fourier multipliers $-\xi_j \xi_i / |\xi|^2$ are uniformly bounded; see for example [Muscalu and Schlag 2013, Section 7.2]. On the other hand, the operator corresponding to the second term can be bounded in $L^2(B \setminus \overline{\Omega})$ by Young's inequality again. Together we find that

$$\mathrm{S}_1 : L^2(\Omega) \to H^2(B \setminus \overline{\Omega}), \quad \mathrm{S}_2 : H^1(\Omega) \to H^2(B \setminus \overline{\Omega}), \quad \text{and} \quad \mathrm{S}_3 : L^2(\Omega) \to H^1(B \setminus \overline{\Omega})$$

are bounded. Thus, by Rellich's theorem, all three operators are bounded from $H^1(\Omega)$ to $H^1(B \setminus \overline{\Omega})$ compactly. Altogether we find that $\mathrm{S}_q$ maps $H^1(\Omega)$ to $H^1(B \setminus \overline{\Omega})$ compactly, which completes the proof of (i).

In order to see (ii), we combine its hypothesis with the alternative representation (52), obtaining

$$\phi = \Gamma_{\Lambda_\sigma}[\phi] = \mathrm{T}_{\partial\Omega} \circ \mathrm{S}_q \circ P_q[\phi]. \tag{53}$$

By the bounds of Section 9, we know that $\mathrm{S}_q \circ P_q[\phi] = e_\rho \Delta_\rho^{-1} \circ \mathrm{M}_q[e_{-\rho} P_q[\phi]] \in H^1(B)$, so we can replace the outer trace on $\partial\Omega$ with the inner trace as they both extend the restriction to $\partial\Omega$ of smooth functions, which are dense in $H^1(B)$. On the other hand, combining the calculation (42) with the defining property of $P_q[\phi]$, we have that

$$-\int_{\mathbb{R}^n} \nabla(\mathrm{S}_q \circ P_q[\phi]) \cdot \nabla\psi = \langle q P_q[\phi], \psi\rangle = -\int_{\mathbb{R}^n} \nabla P_q[\phi] \cdot \nabla\psi$$

whenever $\psi \in C_c^\infty(\Omega)$, and so $\Delta[\mathrm{S}_q \circ P_q[\phi] - P_q[\phi]] = 0$ in $\Omega$ in the weak sense. Combining this with our hypothesis (53) and the uniqueness of solutions for the Dirichlet problem with zero boundary data, we find that

$$\mathrm{S}_q \circ P_q[\phi] = P_q[\phi] \quad \text{in } \Omega. \tag{54}$$

If we had a contraction for $\mathrm{S}_q$, it would be easier to conclude that $\phi = 0$. In any case, we can use the contraction we have by considering

$$\eta := e_{-\rho}\mathrm{S}_q \circ P_q[\phi] = \Delta_\rho^{-1} \circ \mathrm{M}_q[e_{-\rho} P_q[\phi]] = \Delta_\rho^{-1} \circ \mathrm{M}_q[\eta],$$

where the final identity follows from the definition of $\eta$ and (54). Then our contraction (38) implies that $\eta$ must be the zero element of $X^{1/2}_{\lambda,\rho}(B)$, so, by the equivalence of the norms, $e_\rho\eta$ must be the zero element of $H^1(B)$. Then, by the definition of $\eta$ and (54) again, $P_q[\phi]$ is the zero element of $H^1(\Omega)$. Finally, by uniqueness of the Dirichlet problem, $\phi$ is the zero element of $H^{1/2}(\partial\Omega)$, which completes the proof of the injectivity. □

**Remark 11.1.** Much of the previous argument is insensitive to the choice of fundamental solutions used to invert $\Delta$ and $\Delta_\rho$. Rather than troubling ourselves to invert $\Delta_\rho$ using the Faddeev fundamental solution, we could have more easily inverted the operator using the a priori estimates proved in the uniqueness result of [Caro and Rogers 2016]. Indeed, we were able to use those estimates to find a different fundamental solution $K_\rho$ and $w$ such that

$$w(\mathrm{x}) - \langle qw, K_\rho(\mathrm{x}, \bullet)\rangle = \langle q, K_\rho(\mathrm{x}, \bullet)\rangle \quad \text{in } B \setminus \overline{\Omega}.$$

The associated CGO solutions $v = e_\rho(1 + w)$ satisfy

$$v(\mathrm{x}) - \langle qv, L_\rho(\mathrm{x}, \bullet)\rangle = e_\rho(\mathrm{x}) \quad \text{in } B \setminus \overline{\Omega},$$

where $L_\rho(\mathrm{x}, \mathrm{y}) := e_\rho(\mathrm{x} - \mathrm{y})K_\rho(\mathrm{x}, \mathrm{y})$ as before. However, not only are these fundamental solutions less explicitly defined, they also fail to satisfy the skew symmetry law (46): that is $K_{-\rho}(\mathrm{x}, \mathrm{y}) = K_\rho(\mathrm{y}, \mathrm{x})$. Thus, even though we know that $L_\rho(\bullet, y)$ is harmonic on $\mathbb{R}^n \setminus \{y\}$, one is unable to conclude that $L_\rho(x, \bullet)$ is harmonic on $\mathbb{R}^n \setminus \{x\}$, which is what allowed us to take it in the boundary integral identity. We attempted to modify the fundamental solution so that the skew symmetry law is satisfied as in [Nachman and Street 2010]; however, we were unable to do this while maintaining the contraction.

We are now ready to complete the formula for the Fourier transform $\hat{q}(\xi) := \langle q, e^{-i\xi\cdot x}\rangle$, with $q$ defined in (4). The proof makes use of the boundary integral identity again combined with the averaging argument due to [Haberman and Tataru 2013].

**Theorem 3.2.** *Let $\Pi$ be a two-dimensional linear subspace orthogonal to $\xi \in \mathbb{R}^n$, and define*

$$S^1 := \Pi \cap \{\theta \in \mathbb{R}^n : |\theta| = 1\}.$$

*For $\theta \in S^1$, let $\vartheta \in S^1$ be such that $\{\theta, \vartheta\}$ is an orthonormal basis of $\Pi$, and define*

$$\rho := \tau\theta + i\left(-\frac{\xi}{2} + \left(\tau^2 - \frac{|\xi|^2}{4}\right)^{1/2}\vartheta\right), \quad \rho' := -\tau\theta + i\left(-\frac{\xi}{2} - \left(\tau^2 - \frac{|\xi|^2}{4}\right)^{1/2}\vartheta\right),$$

*where $\tau > 1$. Let $BI_{\Lambda_\sigma}$ and $\Gamma_{\Lambda_\sigma}$ be defined by* (11) *and* (12)*, respectively. Then*

$$\hat{q}(\xi) = \lim_{T\to\infty} \frac{1}{2\pi T}\int_T^{2T}\int_{S^1} BI_{\Lambda_\sigma}((\mathrm{I} - \Gamma_{\Lambda_\sigma})^{-1}[e_\rho|_{\partial\Omega}], e_{\rho'})\, \mathrm{d}\theta\, \mathrm{d}\tau.$$

*Proof.* Noting that $\rho\cdot\rho = \rho'\cdot\rho' = 0$, we can take the CGO solution $v = e_\rho(1+w) \in H^1(B)$ given by Corollary 9.2 and $\psi = e_{\rho'}$ in the first boundary integral identity of Lemma 10.1. Noting also that $\rho + \rho' = -i\xi$, the right-hand side of the identity can be written as $\hat{q}(\xi)$ plus a remainder term. Indeed, we find that

$$BI_{\Lambda_\sigma}(v|_{\partial\Omega}, e_{\rho'}) = \hat{q}(\xi) + \langle qw, e^{-i\xi\cdot x}\rangle. \tag{55}$$

Now, for any extension $\tilde{w} \in \dot{X}_\rho^{1/2}$ of $w$ and smooth $\chi_B$ equal to 1 on $\Omega$ and supported on $B$, by duality we have that

$$|\langle qw, e^{-i\xi\cdot x}\rangle| \leqslant \|q\|_{\dot{X}_\rho^{-1/2}} \|\chi_B e^{-i\xi\cdot x}\tilde{w}\|_{\dot{X}_\rho^{1/2}} \leqslant C\|q\|_{\dot{X}_\rho^{-1/2}}\|\tilde{w}\|_{\dot{X}_\rho^{1/2}},$$

where the constant $C > 1$ depends on $|\xi|$ and $R$; see [Haberman and Tataru 2013, Lemma 2.2] or [Caro et al. 2013, (3.17)]. Taking the infimum over extensions, we find

$$|\langle qw, e^{-i\xi\cdot x}\rangle| \leqslant C\|q\|_{\dot{X}_\rho^{-1/2}}\|w\|_{\dot{X}_\rho^{1/2}(B)}.$$

Then, using the estimate (39) for the remainder in Corollary 9.2 and taking an average over $\rho$, we find that

$$\frac{1}{2\pi T}\int_T^{2T}\int_{S^1} |\langle qw, e^{-i\xi\cdot x}\rangle|\, \mathrm{d}\theta\, \mathrm{d}\tau \leqslant \frac{C}{2\pi T}\int_T^{2T}\int_{S^1} \|q\|^2_{\dot{X}_\rho^{-1/2}}\, \mathrm{d}\theta\, \mathrm{d}\tau,$$

where $C > 1$ depends on $|\xi|$, the radius $R$, and $\|\nabla\log\sigma\|_\infty$. Now, [Haberman and Tataru 2013, Lemma 3.1] proved that the right-hand side converges to 0 as $T \to \infty$. Combining with (55), noting that $\hat{q}(\xi)$ is unchanged by the average, yields

$$\hat{q}(\xi) = \lim_{T\to\infty}\frac{1}{2\pi T}\int_T^{2T}\int_{S^1} BI_{\Lambda_\sigma}(v|_{\partial\Omega}, e_{\rho'})\, \mathrm{d}\theta\, \mathrm{d}\tau.$$

Finally, we can use our formula for the values of $v$ on the boundary given by Theorem 3.1, which completes the proof. □

**Remark 11.2.** In [Haberman 2015; Ham et al. 2021; Ponce-Vanegas 2021], the contraction was found after taking similar averages over $\rho$, which yields the existence of a sequence of CGO solutions

$$\{v_j = e_{\rho_j}(1 + w_j)\}_{j\geqslant 1} \quad \text{with } |\rho_j| \to \infty \text{ as } j \to \infty.$$

The authors of the aforementioned works were able to take advantage of the existence of these solutions to prove uniqueness; however, in order to reconstruct in terms of these solutions, one would need to know which values of $\rho_j \in \mathbb{C}^n$ to take.

## 12. Reconstruction in practice

There is an extensive literature dedicated to the real-world practicalities of the Calderón problem, such as stability, partial data and numerical implementation; see for example [Caro et al. 2016; Delbary et al. 2012; Kenig et al. 2007]. Here we suggest some simplifications that would make things easier to measure and calculate without dwelling on how much the simplifications would corrupt the image.

**12.1.** ***What to measure.*** An approximation of the conductivity on the surface $\sigma|_{\partial\Omega}$ could be measured directly by placing real potential differences over pairs of adjacent electrodes, measuring the induced current, and applying Ohm's law. Earlier reconstruction algorithms also required the perpendicular gradient of the conductivity on the surface, which seems harder to measure directly. We would also need to measure an approximation of

$$\mathrm{Meas}_T(\xi) := \frac{1}{2\pi T}\int_T^{2T}\int_{S^1}\int_{\partial\Omega} \Lambda_\sigma[\sigma^{-1/2}e_\rho]\sigma^{-1/2}e_{\rho'}\,\mathrm{d}\theta\,\mathrm{d}\tau$$

for all $\xi \in R^{-1}\mathbb{Z}^n \cap [-cT, cT]^n$, where $cT > 1$ and $R$ is approximately twice the diameter of $\Omega$. For the complex integrand one can place two separate real electric potentials. Given sufficient access to a large enough part of the surface, one would hope to approximate the inner integral with some accuracy; however, applying the oscillating electric potentials could prove to be the more difficult technical challenge. The outer averaged integrals seem less important and a more rudimentary finite sum approximation could be sufficient.

**12.2.** ***What to calculate.*** Given $\mathrm{Meas}_T$ and $\sigma|_{\partial\Omega}$, one could then employ a triangular finite element method to calculate an approximate solution to

$$\begin{cases} \Delta v = (\mathrm{Re}\, q_T)v & \text{in } \Omega, \\ v = \sigma|_{\partial\Omega}^{1/2} & \text{on } \partial\Omega, \end{cases}$$

where, letting $\mathbf{1}_\Omega$ denote the characteristic function of the domain, $q_T$ is defined by

$$q_T(\mathrm{x}) := \frac{1}{(2\pi R)^n}\sum_{\xi\in R^{-1}\mathbb{Z}^n\cap[-cT,cT]^n} e^{i\mathrm{x}\cdot\xi}\left(\mathrm{Meas}_T(\xi) + \frac{|\xi|^2}{2}\widehat{\mathbf{1}_\Omega}(\xi)\right).$$

Then the grayscale image is given by $v^2$, taking $T$ as large as is practicable.

**12.3.** ***Justification of the simplifications.*** A loose interpretation of Theorem 3.1 is that $v|_{\partial\Omega}$ is not so different from $e_\rho|_{\partial\Omega}$ (this is known as the Born approximation; see [Delbary et al. 2012; Knudsen and Mueller 2011; Siltanen et al. 2000] for numerical implementations). Indeed, if the conductivity were constant, then $\Gamma_{\Lambda_\sigma}$ would be identically 0 and so part of the reconstruction integral from Theorem 3.2 could be rewritten using the divergence theorem:

$$\int_{\partial\Omega} \partial_\nu P_0[e_\rho] e_{\rho'} = \int_\Omega \nabla e_\rho \cdot \nabla e_{\rho'} = \rho \cdot \rho' \int_\Omega e^{-i\xi\cdot x} = -\frac{|\xi|^2}{2}\widehat{\mathbf{1}_\Omega}(\xi).$$

Note also that, by the uncertainty principle, $\hat{q}$ and $\widehat{\mathbf{1}_\Omega}$ are essentially constant at scale $R^{-1}$. Thus the reconstruction formula approximately simplifies to $\hat{q} \approx \lim_{T\to\infty} \hat{q}_T$ pointwise. Note that the cutoff of the frequencies serves to mollify, so that $q_T$ is a function even though it approximately converges to $q$ in the distributional sense. Finally, one observes that $\sigma^{1/2}$ is the unique solution to the Schrödinger equation with $v|_{\partial\Omega} = \sigma|_{\partial\Omega}^{1/2}$.

## Acknowledgements

This work was partially supported by Ikerbasque and BERC 2022-2025, the MICINN grants PID2021-125021NAI00, PID2021-122154NB-I00, PID2021-122156NB-I00 and PID2021-124195NB-C33, the Severo Ochoa grants CEX2021-001142-S and CEX2023-001347-S, and the ERC grant AdG-834728.

K. Rogers thanks Kari Astala, Daniel Faraco, Peter Rogers and Jorge Tejero for helpful conversations.

## References

[Astala and Päivärinta 2006a] K. Astala and L. Päivärinta, "A boundary integral equation for Calderón's inverse conductivity problem", *Collect. Math.* (2006), 127–139. MR

[Astala and Päivärinta 2006b] K. Astala and L. Päivärinta, "Calderón's inverse conductivity problem in the plane", *Ann. of Math.* (2) **163**:1 (2006), 265–299. MR

[Astala et al. 2016] K. Astala, D. Faraco, and K. M. Rogers, "Unbounded potential recovery in the plane", *Ann. Sci. Éc. Norm. Supér.* (4) **49**:5 (2016), 1027–1051. MR

[Bennett 1974] C. Bennett, "Banach function spaces and interpolation methods, I: The abstract theory", *J. Functional Analysis* **17** (1974), 409–440. MR

[Brown 1996] R. M. Brown, "Global uniqueness in the impedance-imaging problem for less regular conductivities", *SIAM J. Math. Anal.* **27**:4 (1996), 1049–1056. MR

[Brown 2001] R. M. Brown, "Recovering the conductivity at the boundary from the Dirichlet to Neumann map: a pointwise result", *J. Inverse Ill-Posed Probl.* **9**:6 (2001), 567–574. MR

[Brown and Torres 2003] R. M. Brown and R. H. Torres, "Uniqueness in the inverse conductivity problem for conductivities with 3/2 derivatives in $L^p$, $p > 2n$", *J. Fourier Anal. Appl.* **9**:6 (2003), 563–574. MR

[Calderón 2006] A. P. Calderón, "On an inverse boundary value problem", *Comput. Appl. Math.* **25**:2-3 (2006), 133–138. MR

[Caro and Rogers 2016] P. Caro and K. M. Rogers, "Global uniqueness for the Calderón problem with Lipschitz conductivities", *Forum Math. Pi* **4** (2016), art. id. e2. MR

[Caro et al. 2013] P. Caro, A. García, and J. M. Reyes, "Stability of the Calderón problem for less regular conductivities", *J. Differential Equations* **254**:2 (2013), 469–492. MR

[Caro et al. 2016] P. Caro, D. Dos Santos Ferreira, and A. Ruiz, "Stability estimates for the Calderón problem with partial data", *J. Differential Equations* **260**:3 (2016), 2457–2489. MR

[Delbary et al. 2012] F. Delbary, P. C. Hansen, and K. Knudsen, "Electrical impedance tomography: 3D reconstructions using scattering transforms", *Appl. Anal.* **91**:4 (2012), 737–755. MR

[Faddeev 1965] L. Faddeev, "Increasing solutions of the Schrödinger equation", *Dokl. Akad. Nauk SSSR* **165** (1965), 514–517. In Russian; translated in *Sov. Phys. Dokl.* **10** (1966), 1033–1035.

[García and Zhang 2016] A. García and G. Zhang, "Reconstruction from boundary measurements for less regular conductivities", *Inverse Problems* **32**:11 (2016), art. id. 115015. MR

[Gilbarg and Trudinger 1983] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, 2nd ed., Grundl. Math. Wissen. **224**, Springer, 1983. Corrected reprint, 1998. MR

[Haberman 2015] B. Haberman, "Uniqueness in Calderón's problem for conductivities with unbounded gradient", *Comm. Math. Phys.* **340**:2 (2015), 639–659. MR

[Haberman and Tataru 2013] B. Haberman and D. Tataru, "Uniqueness in Calderón's problem with Lipschitz conductivities", *Duke Math. J.* **162**:3 (2013), 497–516. MR

[Ham et al. 2021] S. Ham, Y. Kwon, and S. Lee, "Uniqueness in the Calderón problem and bilinear restriction estimates", *J. Funct. Anal.* **281**:8 (2021), art. id. 109119. MR

[Jerison and Kenig 1995] D. Jerison and C. E. Kenig, "The inhomogeneous Dirichlet problem in Lipschitz domains", *J. Funct. Anal.* **130**:1 (1995), 161–219. MR

[Kenig et al. 2007] C. E. Kenig, J. Sjöstrand, and G. Uhlmann, "The Calderón problem with partial data", *Ann. of Math.* (2) **165**:2 (2007), 567–591. MR

[Kim and Kwon 2022] H. Kim and H. Kwon, "Dirichlet and Neumann problems for elliptic equations with singular drifts on Lipschitz domains", *Trans. Amer. Math. Soc.* **375**:9 (2022), 6537–6574. MR

[Knudsen and Mueller 2011] K. Knudsen and J. L. Mueller, "The Born approximation and Calderón's method for reconstruction of conductivities in 3-D", *Discrete Contin. Dyn. Syst.* (2011), 844–853. MR

[Kohn and Vogelius 1984] R. Kohn and M. Vogelius, "Determining conductivity by boundary measurements", *Comm. Pure Appl. Math.* **37**:3 (1984), 289–298. MR

[Kohn and Vogelius 1985] R. V. Kohn and M. Vogelius, "Determining conductivity by boundary measurements, II: Interior results", *Comm. Pure Appl. Math.* **38**:5 (1985), 643–667. MR

[Mitrea and Taylor 1999] M. Mitrea and M. Taylor, "Boundary layer methods for Lipschitz domains in Riemannian manifolds", *J. Funct. Anal.* **163**:2 (1999), 181–251. MR

[Muscalu and Schlag 2013] C. Muscalu and W. Schlag, *Classical and multilinear harmonic analysis*, *I*, Cambridge Studies in Advanced Mathematics **137**, Cambridge University Press, 2013. MR

[Nachman 1988] A. I. Nachman, "Reconstructions from boundary measurements", *Ann. of Math.* (2) **128**:3 (1988), 531–576. MR

[Nachman and Street 2010] A. Nachman and B. Street, "Reconstruction in the Calderón problem with partial data", *Comm. Partial Differential Equations* **35**:2 (2010), 375–390. MR

[Nachman et al. 1988] A. Nachman, J. Sylvester, and G. Uhlmann, "An $n$-dimensional Borg–Levinson theorem", *Comm. Math. Phys.* **115**:4 (1988), 595–605. MR

[Newton 1989] R. G. Newton, *Inverse Schrödinger scattering in three dimensions*, Springer, 1989. MR

[Novikov 1988] R. G. Novikov, "A multidimensional inverse spectral problem for the equation $-\Delta\psi + (v(x) - Eu(x))\psi = 0$", *Funktsional. Anal. i Prilozhen.* **22**:4 (1988), 11–22. In Russian; translated in *Funct. Anal. Appl.* **22**:4 (1988), 263–272. MR

[Päivärinta et al. 2003] L. Päivärinta, A. Panchenko, and G. Uhlmann, "Complex geometrical optics solutions for Lipschitz conductivities", *Rev. Mat. Iberoamericana* **19**:1 (2003), 57–72. MR

[Ponce-Vanegas 2021] F. Ponce-Vanegas, "A bilinear strategy for Calderón's problem", *Rev. Mat. Iberoam.* **37**:6 (2021), 2119–2160. MR

[Siltanen et al. 2000] S. Siltanen, J. Mueller, and D. Isaacson, "An implementation of the reconstruction algorithm of A Nachman for the 2D inverse conductivity problem", *Inverse Problems* **16**:3 (2000), 681–699. MR

[Sylvester and Uhlmann 1987] J. Sylvester and G. Uhlmann, "A global uniqueness theorem for an inverse boundary value problem", *Ann. of Math.* (2) **125**:1 (1987), 153–169. MR

[Uhlmann 1998] G. Uhlmann, "Inverse boundary value problems for partial differential equations", pp. 77–86 in *Proceedings of the International Congress of Mathematicians, III* (Berlin, 1998), edited by G. Fischer and U. Rehmann, Deutsche Mathematiker Vereinigung, Berlin, 1998. MR

[Zhang and Bao 2012] W. Zhang and J. Bao, "Regularity of very weak solutions for elliptic equation of divergence form", *J. Funct. Anal.* **262**:4 (2012), 1867–1878. MR

Received 17 Jan 2024. Revised 5 Jul 2024. Accepted 20 Sep 2024.

PEDRO CARO: pcaro@bcamath.org
*Basque Center for Applied Mathematics, Bilbao, Spain*

MARÍA ÁNGELES GARCÍA-FERRERO: garciaferrero@icmat.es
*Current address*: *Instituto de Ciencias Matemáticas, CSIC-UAM-UC3M-UCM, Madrid, Spain*
*Universitat de Barcelona, Barcelona, Spain*

KEITH M. ROGERS: keith.rogers@icmat.es
*Instituto de Ciencias Matemáticas, CSIC-UAM-UC3M-UCM, Madrid, Spain*

ANALYSIS AND PDE
Vol. 18 (2025), No. 8, pp. 2061–2080
DOI: 10.2140/apde.2025.18.2061

# WEAKLY TURBULENT SOLUTION TO THE SCHRÖDINGER EQUATION ON THE TWO-DIMENSIONAL TORUS WITH REAL POTENTIAL DECAYING TO ZERO AT INFINITY

AMBRE CHABERT

We build a smooth time-dependent real potential on the two-dimensional torus, decaying as time tends to infinity in Sobolev norms along with all its time derivatives, and we exhibit a smooth solution to the associated Schrödinger equation on the two-dimensional torus whose $H^s$ norms nevertheless grow logarithmically as time tends to infinity. We use Fourier decomposition in order to exhibit a discrete resonant system of interactions, which we are further able to reduce to a sequence of finite-dimensional linear systems along which the energy propagates to higher and higher frequencies. The constructions are very explicit, and we can thus obtain lower bounds on the growth rate of the solution.

## 1. Introduction

**1.1.** ***Main result.*** In this paper, we build an explicit $\mathcal{C}^\infty$ solution to the Schrödinger equation on the two-dimensional torus $\mathbb{T}^2 := \mathbb{R}^2/(2\pi\mathbb{Z})^2$,

$$i\partial_t u(t,x) = -\Delta u(t,x) + V(t,x)u(t,x), \quad (t,x) \in [0,+\infty) \times \mathbb{T}^2, \tag{1-1}$$

where the potential $V(t,x)$ is real, smooth on the interval $[0,+\infty)\times\mathbb{T}^2$, and decaying at infinity in Sobolev norms. With a carefully chosen $V$, we are able to exhibit *weakly turbulent* behaviour; that is, we are able to prove the following theorem.

**Theorem 1.1.** *There exist a real smooth potential $V(t,x)$ and a smooth function $u(t,x)$ such that*

$$i\partial_t u(t,x) = -\Delta u(t,x) + V(t,x)u(t,x), \quad (t,x) \in [0,+\infty) \times \mathbb{T}^2. \tag{1-2}$$

*Furthermore, given any small constant $\delta > 0$ and any order $s > 0$, there exists $c_{\delta,s} > 0$ such that, as $t \to \infty$,*

$$\|u(t)\|_{H^s} \geqslant c_{\delta,s}(\log t)^{s(1-\delta)}. \tag{1-3}$$

*Finally, the potential $V$ satisfies the bound,*

$$\text{for all } k \in \mathbb{N}, \text{ for all } s \geqslant 0, \quad \lim_{t\to\infty} \|\partial_t^k V(t,\cdot)\|_{H^s} = 0. \tag{1-4}$$

We will in Section 5 explore possible upper bounds for the decay rate of $V$, which is subpolynomial; see (5-16).

*MSC2020:* primary 35B40; secondary 35Q41.
*Keywords:* linear Schrödinger equation, weak turbulence, resonant system, forward cascade of energy, backward integration.

© 2025 The Author, under license to MSP (Mathematical Sciences Publishers). Distributed under the Creative Commons Attribution License 4.0 (CC BY). Open Access made possible by subscribing institutions via Subscribe to Open.

**1.2.** ***Earlier work.*** The first example of unbounded growth of the Sobolev norms for the Schrödinger equation (1-1) on the torus $\mathbb{T}^2$ was given in [Bourgain 1999a], although there the potential $V$ was chosen to be *quasiperiodic*. Bourgain proves that a logarithmic growth of the Sobolev norms can be achieved in this setting and that it is optimal. Bourgain [1999b] also studied the case of a random behaviour in time with certain smoothness conditions. Furthermore, Bourgain proves in those articles that, with a bounded smooth potential $V$, the growth in any norm $H^s$ is bounded by $t^\varepsilon$ for all $\varepsilon > 0$ (with a constant that depends upon $s$, $V$, $\varepsilon$) and that, for a potential *analytic* in time, the bound can be refined to $(\log t)^\alpha$.

With regards to the logarithmic growth rate we are able to achieve in the present article, it is necessarily subpolynomial as $V$ is assumed to be smooth and bounded, but we may not use the logarithmic a priori bound as $V(t)$ is not analytic in $t$ in our construction. Still, logarithmic growth rate is nearly optimal as the optimal growth is necessarily subpolynomial.

The study of upper bounds on the possible growth rate of Sobolev norms of the solutions to the linear Schrödinger equation has a long history. The general question can be formulated as follow: consider $u$ a regular solution to

$$i\partial_t u = Hu + P(t)u, \tag{1-5}$$

where $H$ is either the Laplacian $-\Delta$ on a $d$-dimensional torus, or, more generally, when the domain is $\mathbb{R}^d$ or even a manifold, a time-independent self-adjoint nonnegative operator with some assumptions on its spectrum, and $P(t)$ is a smooth time-dependent family of pseudodifferential operators of order strictly less than 2. Then one can try and prove an upper bound on the growth rate of $\|u(t)\|_{H^s}$ as $t \to \infty$.

Maspero and Robert [2017, Theorem 1.9] proved, along with global well-posedness, $t^\varepsilon$ upper bounds on the growth rate in the case where $H$ has an increasing spectral gap (as is the case for the Laplacian on Zoll manifolds) and $P(t)$ is a smooth perturbation with suitable assumptions. They also proved polynomial upper bounds in broader settings. Under the increasing spectral gap assumption, the bound can be improved to $(\log t)^\gamma$ for some $\gamma > 0$ when $P(t)$ is analytic in time, which is reminiscent of Bourgain's bound. Using those results, [Bambusi et al. 2021] proved $t^\varepsilon$ upper bounds on the growth rate of solutions to (1-5) in an abstract setting, which includes in particular the case where $H$ is the harmonic oscillator in $\mathbb{R}^d$ and $P(t)$ is a pseudodifferential operator of order strictly lower than $H$ depending in a quasiperiodic way on time. The first result of a $t^\varepsilon$ upper bound with an *unbounded* $P(t)$ was obtained in [Bambusi et al. 2022] on the torus $\mathbb{T}^d$ with $H = -\Delta$. Finally, $t^\varepsilon$ upper bounds have been proved for general hamiltonians of quantum integrable systems in [Bambusi and Langella 2022].

Regarding the dual question of exhibiting growth of Sobolev norms in solutions to (1-5), Maspero [2022; 2023] proved the existence of solutions with (unbounded) polynomial growth in the case where $H$ has a fixed spectral gap and $P(t)$ is a potential *periodic* in time using a resonance phenomenon. Loosening the time smoothness hypothesis, Erdoğan, Killip and Schlag [Erdoğan et al. 2003] showed genericity of Sobolev norm growth when the potential is a stationary Markov process. See also [Delort 2010; Eliasson and Kuksin 2009; Nersesyan 2009; Wang 2008].

Regarding potentials whose Sobolev norms decay to 0 with time more specifically, Faou and Raphaël [2023] were able to exhibit logarithmic growth in the context where

$$H = -\Delta + |x|^2$$

is the harmonic oscillator on $\mathbb{R}^2$. Their method relies on *quasiconformal modulations* of so-called *bubble* solutions of the unperturbed Schrödinger equation. It is not surprising that we are able to exhibit logarithmic growth on the torus as the setting is similar. Indeed, both the harmonic oscillator on $\mathbb{R}^2$ and the Laplacian on the torus are operators with compact resolvent and a spectrum with geometric properties (as it is formed of points in a lattice) which allows for explicit resonance mechanism. Let us note that the author was able to prove in [Chabert 2024] that their method extends to the case where the cubic nonlinear term $u|u|^2$ is added to the equation, using an approximation scheme similar to the one found in the present article.

The method we shall use here is inspired by the seminal work [Colliander et al. 2010] refined by [Guardia and Kaloshin 2015]. Indeed, we use that, on the two-dimensional torus, eigenfunctions of the Laplacian are given by $e^{in\cdot x}$ for $n \in \mathbb{Z}^2$ with eigenvalue $|n|^2$. The lattice structure is then used to produce resonance phenomena between carefully chosen frequencies of the Fourier decomposition of the solution $u$. The idea is that only certain *resonant* interactions will dominate the behaviour of the solution; thus, using an arbitrarily small potential, we are able to transfer the energy of the solution to higher and higher frequencies, leading to the growth of Sobolev norms.

**1.3.** ***Main ideas of the proof.*** The first step of the proof is directly inspired by [Colliander et al. 2010]. In Section 2, we decompose (1-1) into Fourier frequencies, thus reducing it to an infinite-dimensional ODE on the Fourier frequencies $(a_n(t))$ of the solution. This enables us to exhibit some *resonant interactions* between Fourier frequencies, which will dominate the behaviour of the solution in terms of Sobolev norms. In that spirit, we first study a *resonant Fourier system* where we drop the nonresonant interactions. We then build a family of Fourier frequencies $(m_n)_{n\geqslant 0}$, satisfying carefully computed *orthogonality properties*, along which we are able to transfer energy to higher frequencies (as $|m_n| \to \infty$) with a well-tailored potential $V$ for a solution $(a_n(t))$ whose Fourier frequencies are almost supported on the $(m_n)$.

In Section 3, we give a detailed construction of a potential allowing said energy transfer to higher frequencies, thanks to the crucial point that, as we only consider resonant interactions, we may light up only specific Fourier frequencies in the potential, which further reduces the resonant system to a *sequence of finite-dimensional linear systems* which we can explicitly solve.

In Sections 4 and 5, we prove that the solution to the resonant system yields a solution to the full system up to a perturbation thanks to a Cauchy sequence scheme, thus concluding that the perturbation decays to 0 as $t \to \infty$. We finally use the explicit construction of the solution to the resonant system to deduce lower bounds on the growth of the Sobolev norm of the full solution, thus concluding to the proof of Theorem 1.1.

## 2. Fourier decomposition and resonant system

**2.1.** ***Reduction to a resonant Fourier system.*** We now show how (1-1) can be heuristically approximated by an easier equation, focussing on the *resonant* interactions. Indeed, as we wish to find smooth solutions of (1-1), we may write

$$u(t,x) = \sum_{n\in\mathbb{Z}^2} a_n(t) e^{i(n\cdot x - |n|^2 t)}. \tag{2-1}$$

We now set the potential to take the form

$$V(t,x) = -\sum_{n\in\mathbb{Z}^2} 2v_n(t)\sin(|n|^2 t)e^{in\cdot x}, \tag{2-2}$$

where $v_{-n} = v_n$ is real. Thus, we need only find a solution to the $l^2$ system

$$\partial_t a_n = \sum_{m\in\mathbb{Z}^2} a_m(t)v_{n-m}(t)(e^{-i\omega^+_{m,n}t} - e^{-i\omega^-_{m,n}t}), \tag{$\mathcal{FS}$}$$

where we set

$$\omega^+_{m,n} := |m|^2 + |m-n|^2 - |n|^2,$$
$$\omega^-_{m,n} := |m|^2 - |m-n|^2 - |n|^2.$$

Now, in the spirit of [Colliander et al. 2010], we expect that the resonant interaction will dominate, that is, the interaction between frequencies $m, n$ such that one of $\omega^+_{m,n}$ or $\omega^-_{m,n}$ is 0. We thus define, for $n \in \mathbb{Z}^2$,

$$\Gamma^+_{\text{res}}(n) := \{m \in \mathbb{Z}^2 : |m|^2 + |m-n|^2 - |n|^2 = 0\},$$
$$\Gamma^-_{\text{res}}(n) := \{m \in \mathbb{Z}^2 : |m|^2 - |m-n|^2 - |n|^2 = 0\}$$

and define the approximated system

$$\partial_t a_n = \sum_{m\in\Gamma^+_{\text{res}}(n)} a_m(t)v_{n-m}(t) - \sum_{m\in\Gamma^-_{\text{res}}(n)} a_m(t)v_{n-m}(t). \tag{$\mathcal{RFS}$}$$

We observe that ($\mathcal{RFS}$) conserves the $l^2$ norm. Indeed,

$$\frac{d}{dt}\|(a_n)\|^2_{l^2} = 2\,\text{Re}\Bigg(\sum_{n\in\mathbb{Z}^2}\sum_{m\in\Gamma^+_{\text{res}}(n)} \overline{a_n(t)}a_m(t)v_{n-m}(t) - \sum_{n\in\mathbb{Z}^2}\sum_{m\in\Gamma^-_{\text{res}}(n)} \overline{a_n(t)}a_m(t)v_{n-m}(t)\Bigg).$$

However, $m \in \Gamma^+_{\text{res}}(n)$ if and only if $n \in \Gamma^-_{\text{res}}(m)$. Using moreover that $v_{-k} = v_k$, we see that the right-hand side equals 0.

**2.2.** ***Geometric interpretation of the resonant frequencies.*** Now, we turn our attention to the geometric interpretation of the equation $\omega^{+/-}_{m,n} = 0$: we first see that $\omega^+_{m,n} = 0$ if and only if

$$\begin{cases} m + (n-m) = n, \\ |m|^2 + |n-m|^2 = |n|^2, \end{cases} \tag{2-3}$$

which means that $m$ is orthogonal to $n-m$. This can be reformulated by saying that $m$ resonates with $m+l$, where $l \in \mathbb{Z}^2$ is orthogonal to $m$.

Similarly we see that $\omega^-_{m,n} = 0$ if and only if $(n-m)$ is orthogonal to $n$, which finally means that $m$ and $n$ are resonant frequencies if one of $m$ or $n$ is the sum of the other one and an orthogonal vector. We may sum these facts up in a lemma.

**Lemma 2.1.** *For all $n, m \in \mathbb{Z}^2$, we have $m \in \Gamma^+_{\text{res}}(n)$ if and only if $m$ and $n-m$ are orthogonal. Moreover, $m \in \Gamma^-_{\text{res}}(n)$ if and only if $n$ and $n-m$ are orthogonal.*

**2.3.** ***Explicit family of resonant frequencies and further reduction.*** We shall now build a potential $(v_m(t))$ and a specific solution to $(\mathcal{RFS})$ by constructing two families $(m_k)$ and $(l_k)$, $k \geqslant 0$, of vectors of $\mathbb{Z}^2$ which satisfy good orthogonality properties. Namely, in some sense, we require that there are no exceptional resonances.

**Lemma 2.2.** *There exist two families $(m_k)_{k\geqslant 0}$ and $(l_k)_{k\geqslant 0}$ of vectors of $\mathbb{Z}^2$ such that*

(P1) $m_k \neq 0,\ l_k \neq 0$;

(P2) $m_k \perp l_{k'} \Leftrightarrow k = k'$;

(P3) $m_{k+1} = m_k + l_k$;

(P4) $m_k$ *is not orthogonal to* $m_{k'}$ *and is not orthogonal to* $m_{k'} - l_{k'}$ *for all* $k, k'$;

(P5) $m_k - l_k$ *is not orthogonal to* $l_{k'}$ *for all* $k, k'$;

(P6) $m_{k'} - l_k$ *is not orthogonal to* $l_k$ *for all* $k' \neq k+1$;

(P7) $m_{k'} - l_{k'} - l_k$ *is not orthogonal to* $l_k$ *for all* $k, k'$;

(P8) $l_k + m_{k'}$ *is not orthogonal to* $l_k$ *for all* $k, k'$;

(P9) $l_k + m_{k'} - l_{k'}$ *is not orthogonal to* $l_k$ *for all* $k \neq k'$;

(P10) $|l_{k+1}| > |l_k| + 1$.

*Moreover, we can find families such that there exist universal constants $C > 1 > c$ such that, for all $n \geqslant 1$, we have*

$$c(n-1)! \leqslant |m_n| \leqslant C^n (n-1)!,$$
$$cn! \leqslant |l_n| \leqslant C^n n!.$$

At first glance these properties may seem overwhelming, but it follows quite directly from geometric observations that they greatly reduce the system if we choose the potential with nonzero Fourier frequencies supported in the set $\{\pm l_k\}_{k\geqslant 0}$. More precisely, before proving Lemma 2.2, we will state and prove the following lemma.

**Lemma 2.3.** *Set $\Lambda := \{\pm l_k,\ k \geqslant 0\}$ and $\Lambda' := \{m_k,\ k \geqslant 0\}$. Set moreover $\Sigma := \{m_k - l_k,\ k \geqslant 0\}$. Assume $(a_n(t))_{n\in\mathbb{Z}^2}$ is a solution to $(\mathcal{RFS})$ with potential $(v_n(t))_n$ such that $(a_n(0))$ is supported in $\Lambda' \cup \Sigma$ (in the sense that $a_n(0) = 0$ whenever $n \notin \Lambda' \cup \Sigma$). If $(v_n(t))_n$ is supported in $\Lambda$ for all $t \geqslant 0$, then $(a_n(t))$ is supported in $\Lambda' \cup \Sigma$ for all $t \geqslant 0$.*

*Moreover, define $p_k(t) := a_{m_k}(t)$, $s_k(t) := a_{m_k - l_k}(t)$ and $r_k(t) := v_{l_k}(t)$ (with the convention that $p_{-1} = r_{-1} = 0$). The system $(\mathcal{RFS})$ reduces to,*

$$\text{for all } k \geqslant 0, \qquad \begin{cases} \partial_t p_k = p_{k-1} r_{k-1} - p_{k+1} r_k - s_k r_k, \\ \partial_t s_k = p_k r_k. \end{cases} \tag{2-4}$$

*Proof.* As $v_n(t) = 0$ whenever $n \notin \Lambda$, $(\mathcal{RFS})$ reduces to

$$\partial_t a_n = \sum_{\substack{m\in\Gamma^+_{\text{res}}(n)\\ n-m\in\Lambda}} a_m(t) v_{n-m}(t) - \sum_{\substack{m\in\Gamma^-_{\text{res}}(n)\\ n-m\in\Lambda}} a_m(t) v_{n-m}(t). \tag{2-5}$$

In order to prove the first part of the lemma, we need only show that, whenever $n \notin \Lambda' \cup \Sigma$, those $m$ that appear on the right-hand side of (2-5) are also not in $\Lambda' \cup \Sigma$. Indeed, the system then reduces to a linear system with zero initial condition on $\mathbb{Z}^2 \backslash \Lambda' \cup \Sigma$, so by uniqueness we have $a_n(t) = 0$ for all $t$ whenever $n \notin \Lambda' \cup \Sigma$.

Take $n \notin \Lambda' \cup \Sigma$. We claim that if $m \in \Lambda' \cup \Sigma$ satisfies $n - m \in \Lambda$, then $m \notin \Gamma^+_{\text{res}}(n) \cup \Gamma^-_{\text{res}}(n)$. Indeed, assume first that $m = m_k$ for some $k$ and $n - m \in \Lambda$. Then there exists $k' \geqslant 0$ such that $n - m = \pm l_{k'}$.

(i) If $n = m_k + l_{k'}$, then $k \neq k'$ as otherwise $n = m_{k+1} \in \Lambda'$; but then $m_k$ is not orthogonal to $n - m_k = l_{k'}$ thanks to (P2), thus $m_k \notin \Gamma^+_{\text{res}}(n)$. Similarly, $n - m_k = l_k$ is not orthogonal to $n = m_k + l_{k'}$ thanks to (P8).

(ii) If $n = m_k - l_{k'}$, then $k' \neq k$ as otherwise $n \in \Sigma$ and $k' \neq k - 1$ as otherwise $n = m_{k-1} \in \Lambda'$. Thus, $m_k \notin \Gamma^+_{\text{res}}(n)$ as $m_k$ is not orthogonal to $-l_{k'}$ by (P2), and $m_k \notin \Gamma^-_{\text{res}}(n)$ as $l_{k'} = n - m_k$ is not orthogonal to $m_k - l_{k'} = n$ thanks to (P6).

Now, assume that $m = m_k - l_k$ for some $k \geqslant 0$ and $n - m = \pm l_{k'}$ for some $k' \geqslant 0$.

(i) If $n = m_k - l_k + l_{k'}$, then $k \neq k'$ as $n \notin \Lambda'$, so $m_k - l_k$ is not orthogonal to $l_{k'}$ thanks to (P5) — thus $m_k - l_k \notin \Gamma^+_{\text{res}}(n)$ — and $l_{k'}$ is not orthogonal to $m_k - l_k + l_{k'}$ thanks to (P9) — thus $m_k - l_k \notin \Gamma^-_{\text{res}}(n)$.

(ii) Finally, if $n = m_k - l_k - l_{k'}$, then — as $m_k - l_k$ is not orthogonal to $-l_{k'}$ thanks to (P5) — we find that $m_k - l_k \notin \Gamma^+_{\text{res}}(n)$ and — as $m_k - l_k - l_{k'}$ is not orthogonal to $-l_{k'}$ thanks to (P7) — we also find that $m_k - l_k \notin \Gamma^-_{\text{res}}(n)$.

In order to prove the second part of the lemma, we follow the same steps. Take $k \geqslant 0$. First, let $m \in \Gamma^+_{\text{res}}(m_k) \cap (\Lambda' \cup \Sigma)$ such that $m_k - m = \pm l_{k'}$. As $m$ is orthogonal to $l_{k'}$, properties (P2) and (P5) yield that $m = m_{k'}$, and thus $m_k = m_{k'} \pm l_{k'}$. As $m_k$ is orthogonal to $l_k$, (P5) yields that necessarily $m_k = m_{k'} + l_{k'} = m_{k'+1}$, and thus $k' = k - 1$ (as from (P3) and (P10) we have $|m_{i+1}| > |m_i|$), which yields the contribution $p_{k-1} r_{k-1}$ to the right-hand side of the first equation.

Now, let $m \in \Gamma^-_{\text{res}}(m_k) \cap (\Lambda' \cup \Sigma)$ such that $m_k - m = \pm l_{k'}$. As $l_{k'}$ is orthogonal to $m_k$, (P2) yields that $k = k'$ — thus $m = m_k \pm l_k$ — and both are in $\Gamma^-_{\text{res}}(m_k) \cap (\Lambda' \cup \Sigma)$. This yields the contribution $-p_{k+1} r_k - s_k r_k$ to the right-hand side of the first equation.

Finally, take $k \geqslant 0$ and $m \in \Gamma^+_{\text{res}}(m_k - l_k) \cap (\Lambda' \cup \Sigma)$ such that $m_k - l_k - m = \pm l_{k'}$: as $m$ is orthogonal to $l_{k'}$, we find again that $m = m_{k'}$, and thus that $m_k = l_k + m_{k'} \pm l_{k'}$. If the sign is a minus, property (P9) yields that $k' = k$, and thus $m = m_k$, which gives the contribution $p_k r_k$ to the right-hand side of the second equation (as we recall that $v_{-n} = v_n$ for all $n$). If the sign is a plus, we find that $m_k - l_k = m_{k'+1}$ is orthogonal to $l_{k'+1}$, which contradicts property (P5).

We see moreover that there is not a $m \in \Gamma^-_{\text{res}}(m_k - l_k) \cap (\Lambda' \cup \Sigma)$ such that $m_k - l_k - m \in \Lambda$. Indeed by definition of $\Gamma^-_{\text{res}}$, this would mean that there is a $k'$ such that $m_k - l_k$ is orthogonal to $l_{k'}$ thus contradicting property (P5). □

We now turn to the proof of Lemma 2.2. Choose $m_0 \in \mathbb{Z}^2 \backslash \{0\}$ arbitrarily; for example, $m_0 = (1, 0)$. As $m_{k+1} = m_k + l_k$, we need only construct the $l_k$ for $k \geqslant 0$. We will do so by induction. Assume the sequence $(m_k)$ is constructed up to $k = n$ and satisfies properties (P1)–(P10) (which means that $l_0, \dots, l_{n-1}$ have been constructed). We need to construct $l_n \in \mathbb{Z}^2$ (and thus $m_{n+1} = m_n + l_n$) such that the properties still hold up to $k = n+1$. Define $m$ the vector obtained from $m_n$ by applying a rotation by angle $\pi/2$ (which is orthogonal to $m_n$ and which has the same Euclidean norm). We will show that there is $a \in \mathbb{N}$ such that $n+1 \leqslant a \leqslant C(n+1)$, with $C$ a universal constant such that setting $l_n := am$ will suffice.

- (P1) always holds, and (P3) holds by definition.

- In order for (P2) to hold, observe first that, by construction, $l_n = am$ is orthogonal to $m_n$. Moreover, we need, on the one hand, $m_k$ to be nonorthogonal to $l_n$ for $k \leqslant n-1$. However, since $m_k \perp l_k$ up to $k = n$ by induction and since we are in dimension 2, this amounts to asking $l_k$ to be nonorthogonal to $m_n$ for $k \leqslant n-1$, which is true by induction from (P2). On the other hand, we need $m_{n+1}$ to be nonorthogonal to $l_k$ up to $k = n$; that is, since $l_k \perp m_k$ and since we are in dimension 2, we need only prove that $m_{n+1} = m_n + am$ is not parallel to $m_k$ for $k \leqslant n$. It is always true for $k = n$ as $a > 0$, and, for each $k \leqslant n-1$, there is at most one value of $a$ for which $m_{n+1}$ could be parallel to $m_k$ (as $m$ is not parallel to $m_k$, otherwise $m_k$ would be orthogonal to $m_n$, thus contradicting (P4)). This excludes at most $n$ possible values for $a$.

- In order for (P4) to hold, we need $m_{n+1} \cdot m_k \neq 0$ for $k \leqslant n$. It is always true for $k = n$, and for $k < n$ it means that $m_n \cdot m_k + am \cdot m_k \neq 0$. Now, $m \cdot m_k \neq 0$, otherwise this would contradict (P2). Thus, at most $n$ possible values of $a$ are to be excluded. We also need $m_{n+1} \cdot (m_k - l_k) \neq 0$ for $k \leqslant n$, which is always true for $k = n$ if we set $a \geqslant 2$, and it follows from the construction of (P5) that $m \cdot (m_k - l_k) \neq 0$, as $m_n$ is not parallel to $m_k - l_k$; hence this excludes at most $n$ values of $a$. We finally need $m_k \cdot (m_n - am) \neq 0$ for $k \leqslant n-1$, which excludes at most $n$ values of $a$, as $m \cdot m_k \neq 0$.

- In order for (P5) to hold, we need, on the one hand, $m_k - l_k$ to be nonparallel to $m_{n+1}$ for $k \leqslant n$, which excludes at most $n$ values of $a$, as this is always true for $k = n$ and as $m_k - l_k$ is not parallel to $m$ for $k < n$ thanks to (P4). On the other hand, we need $m_n - l_n = m_n - am$ to be nonparallel to $m_k$ for $k \leqslant n$, which again excludes at most $n$ values for $a$ as $m$ is not parallel to $m_k$ for $k < n$ thanks to (P2).

- In order for (P6) to hold, we need, on the one hand, $(m_{n+1} - l_k) \cdot l_k \neq 0$ for $k \leqslant n-1$, which is equivalent to $am \cdot l_k \neq$ constant. As we know that $m \cdot l_k \neq 0$ (otherwise $m_k$ is orthogonal to $m_n$), this excludes at most $n$ values for $a$. One the other hand, we need $m_k - am$ to be nonorthogonal to $am$ for $k \leqslant n-1$, which is ensured by the fact that $|m| > |m_k|$.

- In order for (P7) to hold, we need, on the one hand, $m_n - am - l_k$ to be nonorthogonal to $l_k$ for $k \leqslant n-1$, which excludes at most $n$ values for $a$, as $m \cdot l_k \neq 0$. On the other hand, we need $m_k - l_k - am$ to be nonorthogonal to $am$ for $k \leqslant n-1$, and once again this excludes at most $n$ values for $a$.

- In order for (P8) to hold, we need, on the one hand, $am + m_k$ to be nonorthogonal to $am$ for $k \leqslant n$, thus excluding at most $n$ values for $a$, and, on the other hand, $l_k + m_n + am$ to be nonorthogonal to $l_k$ for $k \leqslant n-1$, which excludes at most $n$ values for $a$, as $m \cdot l_k \neq 0$.

- In order for (P9) to hold, we need, on the one hand, $am + m_k - l_k$ to be nonorthogonal to $am$ for $k \leqslant n-1$, thus excluding at most $n$ values for $a$, and, on the other hand, $l_k + m_n - am$ to be nonorthogonal to $l_k$, excluding once again at most $n$ values for $a$.

We thus finally see that any $a \geqslant 1$ except maybe at most $C(n+1)$ values can be chosen, where $C \geqslant 2$. Up to taking $C$ a little larger we may thus find $n+1 \leqslant a \leqslant C(n+1)$ such that setting $l_n = am$ enables the induction hypothesis to be satisfied.

By this procedure we are able to construct sequences for which the desired properties hold. Moreover, we have $n|m_n| \leqslant |l_n| \leqslant Cn|m_n|$, and thus

$$n|m_n| \leqslant |m_{n+1}| \leqslant C'n|m_n|$$

for $C' = \sqrt{C+1}$; thus proving the last part of the lemma.

## 3. Solution to the resonant system

Thanks to the previous section, we are now able to exhibit explicit potential $(r_k(t))$ and an explicit solution $(p_k(t))$, $(s_k(t))$ to (2-4) for which we control precisely the energy transfer between Fourier frequencies. We turn to the explicit study of the mechanism that will allow energy transfer between frequencies. We start at $t=0$ with well-chosen values for $p_0$, $p_1$, $s_0$ and set the other $p_k$ and $s_k$ to be 0. The idea is then to locally fully transfer the energy from $(p_k, s_k, p_{k+1})$ to $(p_{k+1}, s_{k+1}, p_{k+2})$ in finite time, thus ensuring that, for all given $n$, after a time $T_n$, we have $p_k = s_k = 0$ for all $k \leqslant n$. Now, as $(\mathcal{RFS})$ conserves the $l^2$ norm, this ensures that the Sobolev $H^s$ norm is greater than $|m_n|^s$ for $t \geqslant T_n$.

**3.1.** ***General form of the solution to the linear system.*** Explicitly, fix an interval $I = [t_0, t_1]$ and a smooth function $\phi$ on $I$. Fix $k \geqslant 1$. We look at the system

$$\begin{cases} \partial_t p_{k+1} = \phi(t) p_k, \\ \partial_t p_k = -\phi(t) p_{k+1} - \phi(t) s_k, \\ \partial_t s_k = \phi(t) p_k, \end{cases} \tag{3-1}$$

which corresponds to (2-4) when we only light up $r_k(t) = \phi(t)$; that is, we set $r_{k'}(t) = 0$ for $k' \neq k$ on $I$. In comparison, (2-4) is a system on *all* different values of $k$, whereas in system (3-1) we have *fixed* a particular value for $k$. Hence, the equation for $p_{k+1}$ corresponds to the first line of (2-4), where $k$ is replaced by $k+1$ and $r_l = 0$ for $l \neq k$. The system can then be written in the form of a simple linear system:

$$\partial_t \begin{pmatrix} p_{k+1} \\ p_k \\ s_k \end{pmatrix} = \phi(t) A \begin{pmatrix} p_{k+1} \\ p_k \\ s_k \end{pmatrix}, \tag{3-2}$$

where we set

$$A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{3-3}$$

Now, the solution with initial condition

$$\begin{pmatrix} p_{k+1}(t_0) \\ p_k(t_0) \\ s_k(t_0) \end{pmatrix}$$

is given by

$$\begin{pmatrix} p_{k+1}(t) \\ p_k(t) \\ s_k(t) \end{pmatrix} = \exp\left(\left(\int_{t_0}^{t} \phi(s)\,ds\right) A\right) \begin{pmatrix} p_{k+1}(t_0) \\ p_k(t_0) \\ s_k(t_0) \end{pmatrix}. \tag{3-4}$$

Now, one can compute

$$\exp(TA) = \begin{pmatrix} \frac{1}{2}(\cos(T\sqrt{2})+1) & \frac{1}{\sqrt{2}}\sin(T\sqrt{2}) & \frac{1}{2}(\cos(T\sqrt{2})-1) \\ -\frac{1}{\sqrt{2}}\sin(T\sqrt{2}) & \cos(T\sqrt{2}) & -\frac{1}{\sqrt{2}}\sin(T\sqrt{2}) \\ \frac{1}{2}(\cos(T\sqrt{2})-1) & \frac{1}{\sqrt{2}}\sin(T\sqrt{2}) & \frac{1}{2}(\cos(T\sqrt{2})+1) \end{pmatrix}. \tag{3-5}$$

This explicit matrix allows us to build three moves in order to transfer a specific configuration from $(p_k, s_k, p_{k+1})$ to $(p_{k+1}, s_{k+1}, p_{k+2})$ in finite time.

**3.1.1.** *First move.* Start with

$$\begin{pmatrix} p_{k+1}(t_0) \\ p_k(t_0) \\ s_k(t_0) \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1/\sqrt{2} \\ 1/2 \end{pmatrix}. \tag{3-6}$$

We set $\phi$, a nonnegative $\mathcal{C}^\infty$ function with support in $[t_0, t_1]$ such that $\int \phi = 7\pi/(4\sqrt{2})$. We have

$$\begin{pmatrix} p_{k+1}(t_1) \\ p_k(t_1) \\ s_k(t_1) \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}. \tag{3-7}$$

**3.1.2.** *Second move.* We now set

$$\begin{pmatrix} p_{k+1}(t_0) \\ p_k(t_0) \\ s_k(t_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{3-8}$$

With the integral of $\phi$ being $\pi/(2\sqrt{2})$, we have

$$\begin{pmatrix} p_{k+1}(t_1) \\ p_k(t_1) \\ s_k(t_1) \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}. \tag{3-9}$$

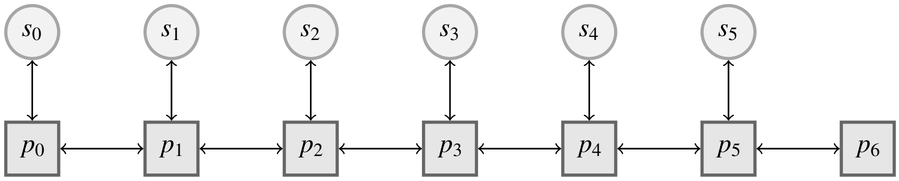
**3.1.3.** *Third move.* If finally we set

$$\begin{pmatrix} p_{k+1}(t_0) \\ p_k(t_0) \\ s_k(t_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{3-10}$$

and set the integral of $\phi$ to be $\pi/\sqrt{2}$, we have

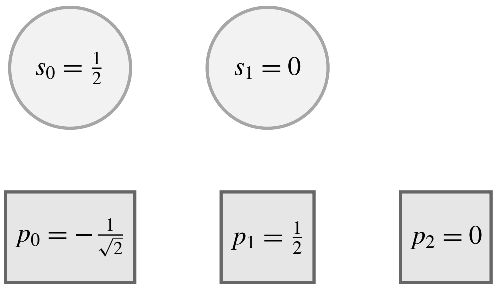
$$\begin{pmatrix} p_{k+1}(t_1) \\ p_k(t_1) \\ s_k(t_1) \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}. \tag{3-11}$$

**3.2.** ***Idea of the construction of the potential and the resonant solution.*** These easy observations yield the construction both of the potential $(r_k(t))$ and of the solution $(p_k(t)),\ (s_k(t))$. We may represent the solution $(p_k(t)),\ (s_k(t))$ as points in the semi-infinite chain

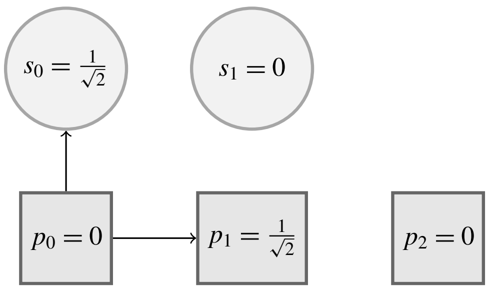



where the arrows represent the possible interactions between the Fourier frequencies induced by the potential $(r_k(t))$.

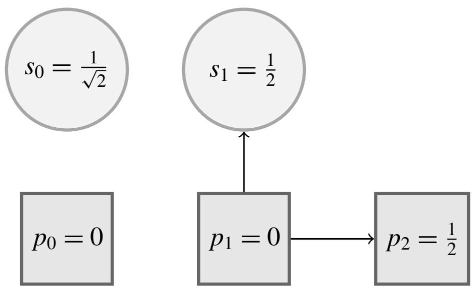
Assume that, at $t = 0$, we start with the configuration



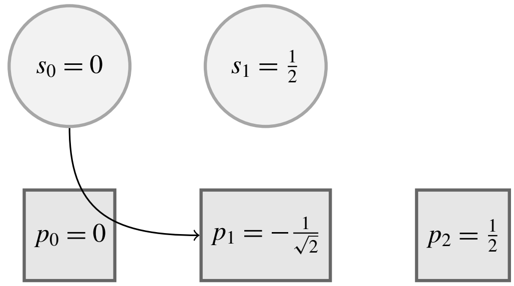

Then, using the first move, if we light up only $r_0$ during an appropriate time, we may fully transfer the mass from $p_0$ to $s_0$ and $p_1$ equally:




Now, we clear $p_1$ using the second move; that is, lighting up only $r_1(t)$, we can fully transfer the mass from $p_1$ to $s_1$ and $p_2$ equally:



Finally, we use the third move to transfer fully the remaining mass from $s_0$ to $p_1$ lighting only $r_0$ again:




Thus, we find exactly the same situation as at the start, with indexes incremented by 1. This enables us to start a recursive scheme so that, as time goes by, we repeat these three moves to transfer the mass to higher frequencies. The strategy to ensure that the potential $V$ decreases in Sobolev norms as $t \to \infty$ is that, up to lighting $r_k$ for a longer time, we may at each step choose it to be arbitrarily small.

**3.3.** ***Explicit computation of the potential and of the resonant solution.*** We now make the previous argument rigorous. We first find a smooth function $\phi$ on $\mathbb{R}$, nonnegative and nondecreasing, such that $\phi = 0$ on $(-\infty, 0]$, $\phi = 1$ on $[1, +\infty)$, and we set $\alpha = \int_0^1 \phi$. Take $(\beta_k)_{k\geqslant 0}$ to be a sequence of positive real numbers such that $\beta_k \ll 1$. The $(\beta_k)$ will control the amplitude to which we light up $r_k$, and we will fix them later in order to control the decay of the potential $V$ in Sobolev norms.

**3.3.1.** *Initialising the induction.* We choose, at $t = 0$,

$$\begin{pmatrix} p_1(0) \\ p_0(0) \\ s_0(0) \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1/\sqrt{2} \\ 1/2 \end{pmatrix} \tag{3-12}$$

(and the other $p_k$, $s_k$ are set to 0). We now set

$$r_0(t) = \begin{cases} \frac{7\pi}{4\sqrt{2}\alpha}\beta_0\phi(t), & 0 \leqslant t \leqslant 1, \\ \frac{7\pi}{4\sqrt{2}\alpha}\beta_0, & 1 \leqslant t \leqslant 1 + t_0, \\ \frac{7\pi}{4\sqrt{2}\alpha}\beta_0\phi(t_0 + 2 - t), & 1 + t_0 \leqslant t \leqslant t_0 + 2, \end{cases} \tag{3-13}$$

where we set $t_0$ such that $\int_0^{t_0+2} r_0 = 7\pi/(4\sqrt{2})$, which means $t_0 = \alpha(\beta_0^{-1} - 2)$. We set $r_k(t) = 0$ on $[0, t_0 + 2]$ for all $k \geqslant 1$.

Now, at $t = t_0 + 2$, we find

$$\begin{pmatrix} p_1(t_0+2) \\ p_0(t_0+2) \\ s_0(t_0+2) \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}. \tag{3-14}$$

Set now

$$r_1(t) = \begin{cases} \frac{\pi}{2\sqrt{2}\alpha}\beta_1\phi(t - (t_0 + 2)), & t_0 + 2 \leqslant t \leqslant t_0 + 3, \\ \frac{\pi}{2\sqrt{2}\alpha}\beta_1, & t_0 + 3 \leqslant t \leqslant t_0 + 3 + t_1, \\ \frac{\pi}{2\sqrt{2}\alpha}\beta_1\phi(4 + t_0 + t_1 - t), & 3 + t_0 + t_1 \leqslant t \leqslant 4 + t_0 + t_1, \end{cases} \tag{3-15}$$

with $t_1$ such that the integral of $r_1$ on $[2+t_0, 4+t_0+t_1]$ is equal to $\pi/(2\sqrt{2})$, which means $t_1 = \alpha(\beta_1^{-1}-2)$. Set $r_k(t) = 0$ on $[2+t_0, 4+t_0+t_1]$ for all $k \neq 1$. Now, at $t = 4+t_0+t_1$, we have

$$\begin{pmatrix} p_2(4+t_0+t_1) \\ p_1(4+t_0+t_1) \\ s_1(4+t_0+t_1) \\ p_0(4+t_0+t_1) \\ s_0(4+t_0+t_1) \end{pmatrix} = \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \tag{3-16}$$

and the other $p_k$, $s_k$ are equal to 0. To finish the cycle, we need to transfer all the mass from $s_0$ to $p_1$, and we will end up with $(p_2, p_1, s_1) = (1/2, -1/\sqrt{2}, 1/2)$ which was exactly the initial state on $(p_1, p_0, s_0)$. This enables us to start a recursive process. More precisely, set

$$r_0(t) = \begin{cases} \frac{\pi}{\sqrt{2}\alpha}\beta_0\phi(t-(4+t_0+t_1)), & 4+t_0+t_1 \leqslant t \leqslant 5+t_0+t_1, \\ \frac{\pi}{\sqrt{2}\alpha}\beta_0, & 5+t_0+t_1 \leqslant t \leqslant 5+t_0+t_1+t_0, \\ \frac{\pi}{\sqrt{2}\alpha}\beta_0\phi(6+t_0+t_1+t_0-t), & 5+t_0+t_1+t_0 \leqslant t \leqslant 6+t_0+t_1+t_0, \end{cases} \tag{3-17}$$

with once again $t_0 = \alpha(\beta_0^{-1}-2)$, and $r_k(t) = 0$ on $[4+t_0+t_1, 6+t_0+t_1+t_0]$ for $k \neq 1$. We have, at $t = 6+2t_0+t_1$,

$$\begin{pmatrix} p_2(t) \\ p_1(t) \\ s_1(t) \\ p_0(t) \\ s_0(t) \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1/\sqrt{2} \\ 1/2 \\ 0 \\ 0 \end{pmatrix}, \tag{3-18}$$

as was expected.

**3.3.2.** *Recursive scheme.* Set $t_n := \alpha(\beta_n^{-1}-2)$. Suppose, for $T_n = 6n+2t_0+3t_1+3t_2+\cdots+3t_{n-1}+t_n$, we have

$$\begin{pmatrix} p_{n+1} \\ p_n \\ s_n \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1/\sqrt{2} \\ 1/2 \end{pmatrix}, \tag{3-19}$$

with the other $p_k$, $s_k$ being equal to 0. We set now

$$r_n(t) = \begin{cases} \frac{7\pi}{8\sqrt{2}\alpha}\beta_n\phi(t), & T_n \leqslant t \leqslant 1+T_n, \\ \frac{7\pi}{8\sqrt{2}\alpha}\beta_n, & 1+T_n \leqslant t \leqslant 1+T_n+t_n, \\ \frac{7\pi}{8\sqrt{2}\alpha}\beta_n\phi(2+T_n+t_n-t), & 1+T_n+t_n \leqslant t \leqslant T_n+2+t_n, \end{cases} \tag{3-20}$$

all the other $r_k$ being set to 0 on $[T_n, T_n+2+t_n]$. Now we have, at $t = T_n+2+t_n$,

$$\begin{pmatrix} p_{n+1} \\ p_n \\ s_n \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}. \tag{3-21}$$

Now set

$$r_{n+1}(t)=\begin{cases}\frac{\pi}{2\sqrt{2}\alpha}\beta_{n+1}\phi(t-(T_n+2+t_n)), & T_n+2+t_n\leqslant t\leqslant T_n+3+t_n,\\ \frac{\pi}{2\sqrt{2}\alpha}\beta_{n+1}, & T_n+3+t_n\leqslant t\leqslant T_n+3+t_n+t_{n+1},\\ \frac{\pi}{2\sqrt{2}\alpha}\beta_{n+1}\phi(T_n+4+t_n+t_{n+1}-t), & T_n+3+t_n+t_{n+1}\leqslant t\leqslant T_n+4+t_n+t_{n+1},\end{cases} \tag{3-22}$$

the other $r_k$ being set to 0 on $[T_n+2+t_n,\, T_n+4+t_n+t_{n+1}]$. We have, at $t=T_n+4+t_n+t_{n+1}$,

$$\begin{pmatrix}p_{n+2}(T_n+4+t_n+t_{n+1})\\ p_{n+1}(T_n+4+t_n+t_{n+1})\\ s_{n+1}(T_n+4+t_n+t_{n+1})\\ p_n(T_n+4+t_n+t_{n+1})\\ s_n(T_n+4+t_n+t_{n+1})\end{pmatrix}=\begin{pmatrix}1/2\\ 0\\ 1/2\\ 0\\ 1/\sqrt{2}\end{pmatrix}. \tag{3-23}$$

Set finally

$$r_n(t)=\begin{cases}\frac{\pi}{\sqrt{2}\alpha}\beta_n\phi(t-(T_n+4+t_n+t_{n+1})), & T_n+4+t_n+t_{n+1}\leqslant t\leqslant T_n+5+t_n+t_{n+1},\\ \frac{\pi}{\sqrt{2}\alpha}\beta_n, & T_n+5+t_n+t_{n+1}\leqslant t\leqslant T_n+5+t_n+t_{n+1}+t_n,\\ \frac{\pi}{\sqrt{2}\alpha}\beta_n\phi(T_n+6+t_n+t_{n+1}+t_n-t), & T_n+5+t_n+t_{n+1}+t_n\leqslant t\leqslant T_n+6+t_n+t_{n+1}+t_n.\end{cases}$$

We now have, at $T_{n+1}=T_n+6+t_n+t_{n+1}+t_n$,

$$\begin{pmatrix}p_{n+2}(t)\\ p_{n+1}(t)\\ s_{n+1}(t)\\ p_n(t)\\ s_n(t)\end{pmatrix}=\begin{pmatrix}1/2\\ -1/\sqrt{2}\\ 1/2\\ 0\\ 0\end{pmatrix}. \tag{3-24}$$

We may now induce this construction for all $n\geqslant 1$, which yields a solution $(p_k(t), s_k(t))$ to (2-4), thus leading to a solution $(a_n(t))$ of $(\mathcal{RFS})$ which we control very explicitly.

**Remark 3.1.** Provided the $\beta_k$ are small enough, the explicit construction yields firstly that $|a_n(t)|\leqslant 1$ for all $n,\ t$, and secondly the following behaviour for $(a_n(t))$: for each n, observe that $a_n(t)=0$ outside of a finite interval. Moreover, this interval can be divided into a bounded number of subintervals, so that either those subintervals are of length 2 (corresponding to the time we take to light up an $r_k$ or turn it off), or $a_n(t)$ is a finite linear combination of oscillating factors $e^{ift}$, where the frequency $f$ is of the order of $\beta_k$ for some $k$ and hence is arbitrarily small.

**3.4.** ***Explicit choice for $\beta_k$ in order for $V$ to decay.*** In order to prove Theorem 1.1, we need to ensure that $V$ and all its derivatives decay with respect to all Sobolev norms as $t\to\infty$. Now, from the construction, we see that, for all $t\geqslant 0$, there is a unique $k(t)$ such that $v_n=0$ for all $n\neq\pm l_{k(t)}$. Now, for any $m\in\mathbb{N}$ and any $s\geqslant 0$, we have

$$\|\partial_t^m V(t,\cdot)\|_{H^s}\simeq\beta_{k(t)}|l_{k(t)}|^{s+2m}. \tag{3-25}$$

As $k(t) \to +\infty$ when $t \to +\infty$, and thus as $|l_{k(t)}| \to +\infty$, we need to ensure that $\beta_k$ decays faster with respect to $k$ than any power of $l_k$. A natural choice is

$$\beta_k := |l_k|^{-|l_k|}, \tag{3-26}$$

and we will see that this choice indeed enables us to close the estimates.

## 4. Approximation

**4.1. *Resonant solution and perturbation decomposition.*** In order to construct a solution to the full system ($\mathcal{FS}$), we try and approximate it by the solution $(a_n(t))$ built in the previous section. In this spirit, we set the solution $(b_n(t))$ to have the a priori form $b_n(t) = a_n(t) + c_n(t)$, where $a_n(t)$ is the solution to ($\mathcal{RFS}$) built above and $c_n$ is a perturbation. We may thus write

$$\partial_t(a_n + c_n) = \sum_{m\in\mathbb{Z}^2} (a_m + c_m)(t) v_{n-m}(t) (e^{-i\omega^+_{m,n}t} - e^{-i\omega^-_{m,n}t}), \tag{4-1}$$

and we already know that

$$\partial_t a_n = \sum_{m\in\Gamma^+_{\text{res}}(n)} a_m(t) v_{n-m}(t) - \sum_{m\in\Gamma^-_{\text{res}}(n)} a_m(t) v_{n-m}(t). \tag{4-2}$$

Thus we need $(c_n)$ to solve

$$\partial_t c_n = \sum_{m\in\mathbb{Z}^2} c_m(t) v_{n-m}(t) (e^{-i\omega^+_{m,n}t} - e^{-i\omega^-_{m,n}t}) + \sum_{m\notin\Gamma^+_{\text{res}}(n)} a_m(t) v_{n-m}(t) e^{-i\omega^+_{m,n}t} - \sum_{m\notin\Gamma^-_{\text{res}}(n)} a_m(t) v_{n-m}(t) e^{-i\omega^-_{m,n}t}. \tag{4-3}$$

Our goal is now to build a solution $(c_n)$ to (4-3) which decays as $t \to \infty$. We will use a Cauchy sequence method: equation (4-3) is globally well-posed in $l^1(\mathbb{Z})$, so we may set, for a given integer $N > 0$, the solution $(c_n^N)$ on $\mathbb{R}_+$ with initial condition $c^N(T_N) = 0$. We have

$$c_n^N(t) = -\sum_{m\in\mathbb{Z}^2} \int_t^{T_N} c_m^N(s) v_{n-m}(s) (e^{-i\omega^+_{m,n}s} - e^{-i\omega^-_{m,n}s})\, ds - \sum_{m\notin\Gamma^+_{\text{res}}(n)} \int_t^{T_N} a_m(s) v_{n-m}(s) e^{-i\omega^+_{m,n}s}\, ds + \sum_{m\notin\Gamma^-_{\text{res}}(n)} \int_t^{T_N} a_m(s) v_{n-m}(s) e^{-i\omega^-_{m,n}s}\, ds,$$

from which we infer, for $t \leqslant T_N$,

$$\|(c_n^N(t))\|_{l^1} \leqslant 2\int_t^{T_N} \|(c_n^N(s))\|_{l^1} \|(v_n(s))\|_{l^1}\, ds + \sum_n \sum_{m\notin\Gamma^+_{\text{res}}(n)} \left| \int_t^{T_N} a_m(s) v_{n-m}(s) e^{i\omega^+_{m,n}s}\, ds \right| + \sum_n \sum_{m\notin\Gamma^-_{\text{res}}(n)} \left| \int_t^{T_N} a_m(s) v_{n-m}(s) e^{i\omega^-_{m,n}s}\, ds \right|,$$

which we rewrite as the inequality, for $t \leqslant T_N$,

$$\|(c_n^N(t))\|_{l^1} \leqslant \alpha(t) + \int_t^{T_N} \|(c_n^N(s))\|_{l^1} \beta(s)\, ds. \tag{4-4}$$

By Gronwall's lemma,

$$\text{for all } t \leqslant T_N, \quad \|(c_n^N(t))\|_{l^1} \leqslant \alpha(t) + \int_t^{T_N} \alpha(s)\beta(s)\exp\left(\int_t^s \beta(\sigma)\,d\sigma\right)ds. \tag{4-5}$$

**4.2.** ***Estimates on*** $\boldsymbol{\alpha(t)}$**.** First, let us study $\alpha(t)$. The set of pairs $(m, n-m)$, $n \in Z^2$ and $m \notin \Gamma_{\text{res}}^+(n)$ (resp. $m \notin \Gamma_{\text{res}}^-(n)$), is equal to the set of pairs $(n_1, n_2) \in \mathbb{Z}^2$ such that $n_1$ and $n_2$ aren't orthogonal (resp. $n_2$ and $n_1 + n_2$ aren't orthogonal). Moreover, we have $v_n(s) = 0$ for all $n \neq \pm l_k$ for a given $k \geqslant 0$, and we recall that $v_{-n} = v_n$. Finally, we know that $a_n(s)v_{l_k}(s) = 0$ as soon as

$$n \notin \{m_k,\, m_k - l_k,\, m_{k+1},\, m_{k+1} - l_{k+1},\, m_{k+2}\} =: E_k.$$

We may then write

$$\alpha(t) = \sum_{k \geqslant 0} \sum_{n \in E_k} I(k, n, t), \tag{4-6}$$

where $I(k, n)$ is a sum of at most four quantities of the form

$$J(k, n, \omega, t) := \left| \int_t^{T_n} a_n(s) r_k(s) e^{i\omega s}\, ds \right| \tag{4-7}$$

and $\omega$ is a frequency belonging to $\mathbb{Z}\backslash\{0\}$, thus ensuring $|\omega| \geqslant 1$. (It is here that we use the nonresonance of the interactions).

We may now write

$$\int_t^{T_n} a_n(s) r_k(s) e^{i\omega s}\, ds = \left[\left(\int_s^t a_n(\sigma) e^{i\omega\sigma}\, d\sigma\right) r_k(s)\right]_t^{T_N} - \int_t^{T_N} \left(\int_s^t a_n(\sigma) e^{i\omega\sigma}\, d\sigma\right) r_k'(s)\, ds.$$

The bracket term is equal to 0 as $r_k$ is 0 at $T_N$ for all $k$. Moreover, we may infer from the construction of $r_k$ that

$$\int_{\mathbb{R}_+} |r_k'(s)|\, ds \leqslant C\beta_k, \tag{4-8}$$

with $C$ a universal constant independent of $k$ (indeed, we use that $r_k$ is a constant except maybe on a finite number of intervals of length 2 where its derivative is bounded by $c\beta_k \|\phi'\|_\infty$).

Finally, we have

$$\left| \int_s^t a_n(\sigma) e^{i\omega\sigma}\, d\sigma \right| \leqslant C, \tag{4-9}$$

with $C$ a universal constant independent of $s,\ n,\ t,\ \omega$. Indeed, for any $n$, using Remark 3.1, we know that, on the one hand, $|a_n| \leqslant 1$ on $\mathbb{R}_+$, and, on the other hand, that, outside of a fixed finite number of intervals of length 2 (yielding a bounded contribution to the integral), $a_n$ is either equal to 0 or equal to a finite linear combination with a bounded number of terms of oscillating exponentials $e^{ift}$, with frequency $f = C'\beta_l$, where $C'$ is a universal constant and $l \geqslant 0$. Thus, up to choosing $|m_0|$ larger, we can require that we always have $|f| < 1/2$. Hence, we are left with integrating oscillating exponentials $e^{i(f+\omega)\sigma}$ where $|f + \omega| \geqslant 1/2$ (since $|\omega| > 1$). A simple integration is enough to conclude the proof of the claim.

This yields the bound

$$J(k,n,\omega,t) \leqslant C\beta_k, \tag{4-10}$$

where $C$ is a universal constant. Moreover, we see that $r_k(s)=0$ for all $s \geqslant T_{k+1}$; thus we have

$$J(k,n,\omega,t) = 0 \quad \text{for all } t \geqslant T_{k+1}. \tag{4-11}$$

From this we may infer the bound

$$\alpha(t) \leqslant C \sum_{k \geqslant k(t)} \beta_k, \tag{4-12}$$

where we set $k(t)$ to be the smallest nonnegative integer such that $t \leqslant T_{k+1}$. Using moreover the fast decay of $\beta_k$, we may further bound, up to taking a larger $C$,

$$\alpha(t) \leqslant C\beta_{k(t)}. \tag{4-13}$$

**4.3.** ***Estimates on $\boldsymbol{\beta(t)}$.*** As for $\beta(t)$, we see that, for all $t$, there is a unique $l(t)$ such that $r_k(t)=0$ as soon as $k \neq l(t)$; thus we find that

$$\beta(t) = 4r_{l(t)}(t). \tag{4-14}$$

This yields the bound

$$\int_t^s \beta(\sigma)\,d\sigma \leqslant 4\int_0^s r_{l(\sigma)}(\sigma)\,d\sigma \leqslant C(k(s)+1);$$

indeed, we see that the integral of $r_k$ over $\mathbb{R}_+$ is a constant independent of $k$.

**4.4.** ***Conclusion of the estimates on $\boldsymbol{c^N}$.*** We may thus bound, for $t \leqslant T_N$,

$$\|(c_n^N(t))\|_{l^1} \leqslant C\biggl(\beta_{k(t)} + \int_t^{T_N} \beta_{k(s)}\beta_{l(s)} \exp(C(k(s)+1))\,ds\biggr). \tag{4-15}$$

Now, from the construction of $(r_k(s))$, we have $l(s) \geqslant k(s)$, and thus $\beta_{l(s)} \leqslant \beta_{k(s)}$. Therefore, for $t \leqslant T_N$, we have

$$\|(c_n^N(t))\|_{l^1} \leqslant C\biggl(\beta_{k(t)} + \int_t^{T_N} \beta_{k(s)}^2 \exp(C(k(s)+1))\,ds\biggr). \tag{4-16}$$

Now, $k(s)$ is equal to $k$ on an interval with measure $l_k$ such that $l_k\beta_k$ is equal to a constant, yielding the bound

$$\|(c_n^N(t))\|_{l^1} \leqslant C\biggl(\beta_{k(t)} + \sum_{k \geqslant k(t)} \beta_k e^{Ck}\biggr). \tag{4-17}$$

As $\beta_k$ is decaying faster than a double exponential, we finally have,

$$\text{for all } t \leqslant T_N, \quad \|(c_n^N(t))\|_{l^1} \leqslant C\beta_{k(t)} e^{Ck(t)}. \tag{4-18}$$

## 5. Cauchy sequence and conclusion

**5.1. *Cauchy sequence.*** We now prove that $(c^N)$ is a Cauchy sequence in $l^1(\mathbb{Z})$. Set $M > N$. We look at the equation satisfied by $c^M - c^N$:

$$(c_n^M - c_n^N)(t) = -\sum_{m\in\mathbb{Z}^2}\int_t^{T_N}(c_m^M - c_m^N)(s)v_{n-m}(s)(e^{i\omega_{m,n}^+ s} - e^{i\omega_{m,n}^- s})\,ds + c_n^M(T_N)$$
$$-\sum_{m\notin\Gamma_{\mathrm{res}}^+(n)}\int_{T_N}^{T_M} a_m(s)v_{n-m}(s)e^{-i\omega_{m,n}^+ s}\,ds + \sum_{m\notin\Gamma_{\mathrm{res}}^-(n)}\int_{T_N}^{T_M} a_m(s)v_{n-m}(s)e^{-i\omega_{m,n}^- s}\,ds.$$

Thus

$$\begin{aligned}\|((c_n^M - c_n^N)(t))\|_{l^1} &\leqslant 2\int_t^{T_N}\|((c_n^M - c_n^N)(s))\|_{l^1}\|(v_n(s))\|_{l^1}\,ds + \|(c_n^M(T_N))\|_{l^1}\\ &\quad+\sum_n\sum_{m\notin\Gamma_{\mathrm{res}}^+(n)}\left|\int_{T_N}^{T_M} a_m(s)v_{n-m}(s)e^{-i\omega_{m,n}^+ s}\,ds\right|\\ &\quad+\sum_n\sum_{m\notin\Gamma_{\mathrm{res}}^-(n)}\left|\int_{T_N}^{T_M} a_m(s)v_{n-m}(s)e^{-i\omega_{m,n}^- s}\,ds\right|\\ &\leqslant 2\int_t^{T_N}\|((c_n^M - c_n^N)(s))\|_{l^1}\|(v_n(s))\|_{l^1}\,ds + C\beta_{k(T_N)}e^{Ck(T_N)} + C\beta_{k(T_N)}\\ &\leqslant 2\int_t^{T_N}\|((c_n^M - c_n^N)(s))\|_{l^1}\|(v_n(s))\|_{l^1}\,ds + C\beta_{N-1}e^{C(N-1)}.\end{aligned}$$

Using the backward Gronwall lemma,

$$\|((c_n^M - c_n^N)(t))\|_{l^1} \leqslant C\beta_{N-1}e^{C(N-1)}\left(1+\int_t^{T_N}\beta(s)\exp\left(\int_t^s\beta(\sigma)\,d\sigma\right)ds\right), \tag{5-1}$$

where $\beta(s) = 2\|(v_n(s))\|_{l^1}$. We know that $\beta(s) = 4r_{l(s)}(s)$, and thus $\int_s^t \beta(\sigma)\,d\sigma \leqslant C(k(s)+1)$. We have

$$\|((c_n^M - c_n^N)(t))\|_{l^1} \leqslant C\beta_{N-1}e^{C(N-1)}\left(1+\int_t^{T_N}\beta_{k(s)}\exp(Ck(s))\,ds\right). \tag{5-2}$$

This upper bound decays to 0 as $N, M \to \infty$ if we fix $t$. This shows that $(c^N(t))$ is a Cauchy sequence in $l^1(\mathbb{Z}^2)$, and it thus converges to a $c(t)$ such that, using integral form of the differential equation, $b = a + c$ is a solution to $(\mathcal{RFS})$. We have, moreover,

$$\|(c_n(t))\|_{l^1} \leqslant C\beta_{k(t)}e^{Ck(t)}, \tag{5-3}$$

and this upper bound decays to 0 as $t \to +\infty$, as expected.

**5.2. *Growth of the Sobolev norm: qualitative result.*** In order to conclude, we recall that $\|(a_n(t))\|_{l^2}$ is preserved and that, for all $t \geqslant 0$, there are at most five of the $a_n$ that are nonzero. Therefore, we have, on the one hand, that $a_k = 0$ for $|k| < |m_n|$ and for all $t \geqslant T_n$, and, on the other hand, that there exists $|k| \geqslant |m_n|$ such that $|a_k(t)| \geqslant \varepsilon$, where $\varepsilon > 0$ is a universal constant. Now, if we set $N$ large enough, we can ensure that $\|(c_n(t))\|_{l^1} \leqslant \varepsilon/2$ for $t \geqslant T_N$. Therefore, for all $t \geqslant T_N$ with $N$ large enough, there exists $|k| \geqslant |m_N|$ such that $b_k = a_k + c_k$ satisfies $|b_k| \geqslant \varepsilon/2$. Now, this ensures that,

$$\text{for all } t \geqslant T_N, \quad \|(b_n(t))\|_{H^s} \geqslant |k|^s |b_k| \geqslant (\varepsilon/2)|m_N|^s. \tag{5-4}$$

This already yields a qualitative result for Theorem 1.1, as we already proved in Section 3.4 that the potential $V$ along with all its time derivatives are decaying in all Sobolev norms when $t \to +\infty$.

**5.3.** ***Quantitative estimates on the growth rate.*** We now investigate the quantitative bounds that we can hope to get on the rate of growth.

We first see that $T_n \leqslant C\beta_n^{-1}$ using the fast decay of $\beta_n$. Moreover, as $|l_n| \leqslant C^n n!$, we find that

$$T_n \leqslant \exp(C^n n! \log(C^n n!)). \tag{5-5}$$

This yields the lower bound

$$\|(b_n(t))\|_{H^s} \geqslant \delta |m_{n(t)}|^s, \tag{5-6}$$

where $\delta > 0$ is a constant and $n(t)$ is the largest integer $n$ such that $\exp(C^n n! \log(C^n n!)) \leqslant t$. Now, we know moreover that $|m_n| \geqslant c(n-1)!$, thus leading to the lower bound

$$\|u(t)\|_{H^s} \geqslant \varepsilon c^s ((n(t)-1)!)^s. \tag{5-7}$$

In order to obtain better bounds, take $\eta > 0$. We first use Stirling's formula

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}, \tag{5-8}$$

which ensures that, provided $n$ is large enough,

$$C^n n! \log(C^n n!) \leqslant ((1+\eta)n)^{(1+\eta)n}. \tag{5-9}$$

Now set $f(x) := x^x$. We find that, provided

$$f((1+\eta)n) \leqslant \log t \tag{5-10}$$

and provided $n$ is large enough, we have $n \leqslant n(t)$. Now, provided $n$ is large enough, we also have

$$(n-1)! \geqslant ((1-\eta)n)^{(1-\eta)n} = f((1-\eta)n). \tag{5-11}$$

Thus, setting $E(x)$ to be the largest integer $k$ such that $k \leqslant x$, we can find a lower bound of the form

$$\begin{aligned}
\|u(t)\|_{H^s} &\geqslant \left(cf\left(\frac{1-\eta}{1+\eta}E(f^{-1}(\log t))\right)\right)^s \\
&\geqslant c_{s,\eta} \exp\left(s\frac{(1-\eta)^2}{1+\eta} f^{-1}(\log t) \log\left(\frac{(1-\eta)^2}{1+\eta} f^{-1}(\log t)\right)\right) \quad \text{(provided } t \text{ is large enough)} \\
&\geqslant c_{s,\eta} \exp\left(s\frac{(1-\eta)^3}{1+\eta} f^{-1}(\log t) \log(f^{-1}(\log t))\right) \quad \text{(provided } t \text{ is large enough)} \\
&\geqslant c_{s,\eta} (\log t)^{s(1-\eta)^3/(1+\eta)}.
\end{aligned}$$

As we may choose $\eta$ arbitrarily, we find that, given any $\delta$, $s > 0$, there exists $c_{\delta,s} > 0$ such that, for $t > 1$,

$$\|u(t)\|_{H^s} \geqslant c_{\delta,s}(\log t)^{s(1-\delta)}, \tag{5-12}$$

thus concluding the proof of Theorem 1.1.

**5.4.** ***Estimates on the decay rate of V.*** We now prove similar upper bounds on the decay rate of the potential $V(t)$. Fix $s \geqslant 0$ and $m \in \mathbb{N} \cup \{0\}$. Thanks to (1-4), we may bound

$$\|\partial_t^m V(t,\cdot)\|_{H^s} \leqslant c|l_{k(t)}|^{M-|l_{k(t)}|}, \tag{5-13}$$

where $M = M_{m,s} > 0$ and $k(t)$ is the unique $k \geqslant 0$ such that $r_{k(t)} \neq 0$. We may furthermore infer from the previous subsection that, given $\delta > 0$, there exists $c_\delta > 0$ such that

$$|l_{k(t)}| \geqslant c_\delta (\log t)^{1-\delta}. \tag{5-14}$$

Thus

$$\|\partial_t^m V(t,\cdot)\|_{H^s} \leqslant C_\delta \exp((M_{m,s} - (\log t)^{1-\delta})(1-\delta)\log\log t). \tag{5-15}$$

As this holds for all $\delta > 0$, we may conclude that, for all $\delta > 0$, there exists $C_{\delta,m,s}$ such that

$$\|\partial_t^m V(t,\cdot)\|_{H^s} \leqslant C_{\delta,m,s} \exp(-(\log t)^{1-\delta})\log\log t). \tag{5-16}$$

As this yields a quantitative bounds for the decay of $V$, it should be noted that it is subpolynomial in the sense that the upper bound decays slower than $t^{-\varepsilon}$ for all $\varepsilon > 0$. It doesn't seem that we can improve the bound, as, on $[T_N, T_{N+1}]$, $\|V(t)\|_{H^1}$ is of order $\beta_N$ and $T_{N+1}$ is of order $\beta_{N+1}^{-1}$. As for all $\varepsilon > 0$ asymptotically we have $\beta_{N+1}^{\varepsilon} \ll \beta_N$, we thus cannot hope for a better bound.

## Acknowledgements

This work was completed while I was a Ph.D. student at DMA, École normale supérieure, Université PSL, CNRS, 75005 Paris, France.

I would like to express my deepest thanks to my Ph.D. advisors, Professors Pierre Germain and Isabelle Gallagher, who helped me greatly to format the present article through many discussions and proofreading. I would also like to thank Professor Pierre Raphaël for asking me the question solved here.

## References

[Bambusi and Langella 2022] D. Bambusi and B. Langella, "Growth of Sobolev norms in quasi integrable quantum systems", 2022. Zbl arXiv 2202.04505

[Bambusi et al. 2021] D. Bambusi, B. Grébert, A. Maspero, and D. Robert, "Growth of Sobolev norms for abstract linear Schrödinger equations", *J. Eur. Math. Soc.* (*JEMS*) **23**:2 (2021), 557–583. MR

[Bambusi et al. 2022] D. Bambusi, B. Langella, and R. Montalto, "Growth of Sobolev norms for unbounded perturbations of the Schrödinger equation on flat tori", *J. Differential Equations* **318** (2022), 344–358. MR

[Bourgain 1999a] J. Bourgain, "Growth of Sobolev norms in linear Schrödinger equations with quasi-periodic potential", *Comm. Math. Phys.* **204**:1 (1999), 207–247. MR Zbl

[Bourgain 1999b] J. Bourgain, "On growth of Sobolev norms in linear Schrödinger equations with smooth time dependent potential", *J. Anal. Math.* **77** (1999), 315–348. MR Zbl

[Chabert 2024] A. Chabert, "A weakly turbulent solution to the cubic nonlinear harmonic oscillator on $\mathbb{R}^2$ perturbed by a real smooth potential decaying to zero at infinity", *Comm. Partial Differential Equations* **49**:3 (2024), 185–216. MR

[Colliander et al. 2010] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation", *Invent. Math.* **181**:1 (2010), 39–113. MR

[Delort 2010] J.-M. Delort, "Growth of Sobolev norms of solutions of linear Schrödinger equations on some compact manifolds", *Int. Math. Res. Not.* **2010**:12 (2010), 2305–2328. MR Zbl

[Eliasson and Kuksin 2009] H. L. Eliasson and S. B. Kuksin, "On reducibility of Schrödinger equations with quasiperiodic in time potentials", *Comm. Math. Phys.* **286**:1 (2009), 125–135. MR Zbl

[Erdoğan et al. 2003] M. B. Erdoğan, R. Killip, and W. Schlag, "Energy growth in Schrödinger's equation with Markovian forcing", *Comm. Math. Phys.* **240**:1-2 (2003), 1–29. MR Zbl

[Faou and Raphaël 2023] E. Faou and P. Raphaël, "On weakly turbulent solutions to the perturbed linear harmonic oscillator", *Amer. J. Math.* **145**:5 (2023), 1465–1507. MR Zbl

[Guardia and Kaloshin 2015] M. Guardia and V. Kaloshin, "Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation", *J. Eur. Math. Soc.* (*JEMS*) **17**:1 (2015), 71–149. MR

[Maspero 2022] A. Maspero, "Growth of Sobolev norms in linear Schrödinger equations as a dispersive phenomenon", *Adv. Math.* **411** (2022), art. id. 108800. MR Zbl

[Maspero 2023] A. Maspero, "Generic transporters for the linear time-dependent quantum harmonic oscillator on $\mathbb{R}$", *Int. Math. Res. Not.* **2023**:14 (2023), 12088–12118. MR Zbl

[Maspero and Robert 2017] A. Maspero and D. Robert, "On time dependent Schrödinger equations: global well-posedness and growth of Sobolev norms", *J. Funct. Anal.* **273**:2 (2017), 721–781. MR

[Nersesyan 2009] V. Nersesyan, "Growth of Sobolev norms and controllability of the Schrödinger equation", *Comm. Math. Phys.* **290**:1 (2009), 371–387. MR Zbl

[Wang 2008] W.-M. Wang, "Logarithmic bounds on Sobolev norms for time dependent linear Schrödinger equations", *Comm. Partial Differential Equations* **33**:10-12 (2008), 2164–2179. MR

Received 10 Apr 2024. Revised 14 Aug 2024. Accepted 20 Sep 2024.

AMBRE CHABERT: ambre.chabert@ens.psl.eu
*Département de Mathématiques et Applications, Ecole Normale Supérieure, Paris, France*

## 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 18 No. 8 2025