Analysis & PDE

Volume 19 No. 7 2026

msp

Analysis & PDE

msp.org/apde

EDITORS-IN-CHIEF

Colin Guillarmou	Université Paris-Saclay, France colin.guillarmou@universite-paris-saclay.fr
Anna L. Mazzucato	Penn State University, USA alm24@psu.edu
Clément Mouhot	Cambridge University, UK c.mouhot@dpmms.cam.ac.uk

BOARD OF EDITORS

Massimiliano Berti	Sc. Intern. Sup. di Studi Avan., Italy berti@sissa.it	Omar Mohsen	Université Paris-Cité, France omar.mohsen.fr@gmail.com
Zbigniew Błocki	Uniwersytet Jagielloński, Poland zbigniew.blocki@uj.edu.pl	Igor Rodnianski	Princeton University, USA irod@math.princeton.edu
Yu Deng	University of Chicago, USA yudeng@uchicago.edu	Xavier Ros Oton	Catalan Inst. for Res. and Adv. Std., Spain xros@icrea.cat
Thierry Gallay	Université Grenoble Alpes, France Thierry.Gallay@univ-grenoble-alpes.fr	Nicolas Rougerie	ENS Lyon, France nicolas.rougerie@ens-lyon.fr
David Gérard-Varet	Université de Paris, France david.gerard-varet@imj-prg.fr	Dimitri Shlyakhtenko	University of California, Los Angeles, USA shlyakht@math.ucla.edu
Ursula Hamenstaedt	Universität Bonn, Germany ursula@math.uni-bonn.de	Yum-Tong Siu	Harvard University, USA siu@math.harvard.edu
Sebastian Herr	Universität Bielefeld, Germany herr@math.uni-bielefeld.de	Jian Song	Rutgers University, USA jiansong@math.rutgers.edu
Peter Hintz	ETH Zurich, Switzerland peter.hintz@math.ethz.ch	Xavier Tolsa	Catalan Inst. for Res. and Adv. Std., Spain xavier.tolsa@uab.cat
Vadim Kaloshin	Institute of Science and Technology, Austria vadim.kaloshin@gmail.com	Gunther Uhlmann	University of Washington, USA gunther@math.washington.edu
Carlangelo Liverani	Università degli Studi di Roma Tor Vergata, Italy liverani@mat.uniroma2.it	András Vasy	Stanford University, USA andras@math.stanford.edu
Jonathan Wing-hong Luk	Standford University jluk@stanford.edu	Jim Wright	University of Edinburgh, UK j.r.wright@ed.ac.uk
Richard B. Melrose	Massachussets Inst. of Tech., USA rbm@math.mit.edu	Ruixiang Zhang	University of California, Berkeley, USA ruixiang@berkeley.edu
Frank Merle	Université de Cergy-Pontoise, France merle@ihes.fr	Maciej Zworski	University of California, Berkeley, USA zworski@math.berkeley.edu
William Minicozzi II	Johns Hopkins University, USA minicozz@math.jhu.edu		

Cover image: Eric J. Heller: "Linear Ramp"

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

The subscription price for 2026 is US $500/year for the electronic version, and $780/year (+$75, 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, 2000 Allston Way #59, Berkeley, CA 94701-4004, 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/

© 2026 Mathematical Sciences Publishers

ANALYSIS AND PDE
Vol. 19 (2026), No. 7, pp. 1251–1294
DOI: 10.2140/apde.2026.19.1251

msp

EXISTENCE, UNIQUENESS AND CHARACTERISATION OF LOCAL MINIMISERS IN SECOND-ORDER CALCULUS OF VARIATIONS IN L^∞

NIKOS KATZOURAKIS AND ROGER MOSER

We study variational problems for second-order supremal functionals $\mathrm{F}_\infty(u) = \|F(\cdot, u, \mathrm{D}u, \mathrm{A}{:}\mathrm{D}^2u)\|_{\mathrm{L}^\infty(\Omega)}$, where F satisfies certain natural assumptions, A is a positive symmetric matrix, and $\Omega \Subset \mathbb{R}^n$. Higher-order problems are very novel in the calculus of variations in L^∞ and exhibit a strikingly different behaviour compared to first-order problems, for which there exists an established theory, pioneered by Aronsson in 1960s. The aim of this paper is to develop a complete theory for F_∞. We prove that, under appropriate conditions, "localised" minimisers can be characterised as solutions to a nonlinear system of PDEs, which is different from the corresponding Aronsson equation for F_∞; the latter is only a necessary but not a sufficient condition for minimality. We also establish the existence and uniqueness of localised minimisers subject to Dirichlet conditions on $\partial\Omega$, as well as their partial regularity outside a singular set of codimension 1, which may be nonempty even if $n = 1$.

1. Introduction and main results

Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open set with Lipschitz boundary $\partial\Omega$, $n \in \mathbb{N}$. Suppose also that

$$F : \Omega \times (\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}) \to \mathbb{R}$$

is a given Carathéodory function, whose arguments will be denoted by $(x, \eta, \mathrm{p}, \xi)$. In this paper we are interested in studying L^∞ variational problems for supremal functionals involving higher-order derivatives of admissible functions. Specifically, we will consider the second-order functional

$$\mathrm{F}_\infty(u) := \|F(\cdot, u, \mathrm{D}u, \mathrm{A}{:}\mathrm{D}^2u)\|_{\mathrm{L}^\infty(\Omega)}, \quad u \in \mathcal{W}^{2,\infty}(\Omega), \tag{1-1}$$

where the class $\mathcal{W}^{2,\infty}(\Omega)$ of weakly twice differentiable functions is defined as

$$\mathcal{W}^{2,\infty}(\Omega) := \bigcap_{1<p<\infty} \{u \in \mathrm{W}^{2,p}(\Omega) \mid \mathrm{A}{:}\mathrm{D}^2u \in \mathrm{L}^\infty(\Omega)\}. \tag{1-2}$$

In (1-1)–(1-2) $\mathrm{A} \in \mathbb{R}^{n\times n}$ denotes a positive symmetric matrix, and ":" denotes the Euclidean inner product in $\mathbb{R}^{n\times n}$ (for details on all the notation we employ in this paper, we refer to Section 2.1). Variational problems for second-order (or higher-order) supremal functionals are very novel in the calculus of variations in L^∞, and no general theory is yet available. *The purpose of the present paper is to develop a complete theory for the general second-order functional* (1-1).

Katzourakis has been partially financially supported through the EPSRC grant EP/X017206/1. Moser has been partially financially supported through the EPSRC grant EP/X017109/1.

MSC2020: primary 35A15, 35B38, 35D99, 35J94, 49K20; secondary 49J27.

Keywords: calculus of variations in L^∞, higher-order problems, local minimisers, absolute minimisers, Euler–Lagrange equations, Aronsson equations.

© 2026 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.

Since F_∞ is not Gateaux differentiable, general critical points cannot be defined in the classical sense. The lack of smoothness is a standard intricacy of all supremal functionals and persists even if, e.g., F is C^∞ and we consider only the restriction $\mathrm{F}_\infty|_{\mathrm{C}^2(\bar{\Omega})}$; by Danskin's theorem [1966], the Gateaux derivative exists only when the argmax set is a singleton. We therefore turn our attention to the *existence*, *uniqueness and regularity of appropriately defined minimisers*, *as well in their variational characterisation through PDEs.*

Before introducing our assumptions and results, and for the sake of comparison, let us discuss briefly the more classical case of first-order L^∞ functionals, which presents some striking differences. In this case, at least when restricted to scalar-valued admissible functions, there exists a rather complete theory. Let $\mathrm{H} : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be given, and consider

$$\mathrm{E}_\infty(u, \mathcal{O}) := \|\mathrm{H}(\,\cdot\,, u, \mathrm{D}u)\|_{\mathrm{L}^\infty(\mathcal{O})}, \quad u \in \mathrm{W}^{1,\infty}(\Omega), \quad \mathcal{O} \Subset \Omega. \tag{1-3}$$

Variational problems for (1-3) were pioneered in the 1960s by Aronsson, who is considered the founder of the theory; see, e.g., [Aronsson 1965; 1966; 1968; 1969]. The trademark feature of (1-3) is that minimisers of $\mathrm{E}_\infty(\,\cdot\,, \Omega)$ are in general highly nonunique and may not satisfy any PDE, in the sense of some analogue of the Euler–Lagrange equations of integral functionals. If however one builds locality in the minimality notion — namely require *by definition* that $\mathrm{E}_\infty(\,\cdot\,, \mathcal{O})$ is minimised for all $\mathcal{O} \Subset \Omega$ with respect to the function's own boundary values on $\partial\mathcal{O}$ — then any such function u satisfies the *Aronsson equation* in the viscosity sense and is called an *absolute minimiser* of (1-3); for a pedagogical introduction, we refer, e.g., to [Katzourakis 2015]. The Aronsson equation arising from (1-3) is the quasilinear second-order degenerate elliptic PDE

$$\mathrm{H}_\mathrm{p}(\,\cdot\,, u, \mathrm{D}u)\mathrm{D}(\mathrm{H}(\,\cdot\,, u, \mathrm{D}u)) = 0 \quad \text{in } \Omega. \tag{1-4}$$

It turns out that the opposite is true as well, and the PDE (1-4) is sufficient for absolute minimality, whilst the corresponding Dirichlet problem has a unique solution. All these deep results require appropriate assumptions to be imposed, typically involving convexity of H in p, and no explicit dependence in u. Without any attempt to be exhaustive, for some important contributions in the area, we refer, e.g., to [Ansini and Prinari 2015; Barron et al. 2001; Brizzi and De Pascale 2023; Bungert and Korolev 2022; Champion et al. 2004; Jensen et al. 2008; Kreisbeck and Zappale 2020; Miao et al. 2017; Peng et al. 2021; Prinari and Zappale 2020; Ribeiro and Zappale 2014].

The phenomena arising in the study of higher-order supremal functionals are radically different from those governing the first-order case of (1-3)–(1-4). Firstly, at least for the case of (1-1) where the dependence on second derivatives is through the elliptic operator $u \mapsto \mathrm{A}{:}\mathrm{D}^2u$, absolute minimisers are not a fundamental notion, as *usual* (*global*) *minimisers of* (1-1) *satisfy astounding uniqueness properties*, even with u-dependence being allowed and under very weak convexity requirements. In fact, one can take this one step further, and show that even *localised minimisers* — namely those which minimise only in an appropriate neighbourhood of $\mathcal{W}^{2,\infty}(\Omega)$ — are also locally unique. On the other hand, the corresponding higher-order Aronsson equation for (1-1), although it can be easily derived, plays only an ancillary role in this setting, as it *is a necessary but not a sufficient condition for minimality*, even if we restrict our attention to classical solutions. Additionally, it is fully nonlinear and of third-order, failing to be elliptic in

any reasonable sense. Furthermore, the very definition of generalised solutions for the Aronsson equation is a nontrivial matter, as neither weak nor viscosity solutions apply.

In this paper, following an entirely different approach, *we identify a different set of equations which completely characterises minimisers of* (1-1), *even localised ones*, *and it allows us to ascertain their uniqueness for given prescribed Dirichlet boundary conditions* under natural assumptions. We may also include explicit $(u, \mathrm{D}u)$-dependence, a fact which is interesting in itself, as u-dependence typically precludes uniqueness assertions in the first-order case; see, e.g., [Yu 2006]. Our investigations are in part inspired by curvature problems in Riemannian geometry, specifically the Yamabe and Nirenberg problems; see [Moser and Schwetlick 2012; Sakellaris 2017] and subsequently. However, our results herein have a different focus, and as they stand they do not directly apply to differential geometry.

We now present the main contributions of this work, beginning with the minimality notions. To state them concisely, let us define the following subspace of (1-2):

$$\mathcal{W}_0^{2,\infty}(\Omega) := \mathcal{W}^{2,\infty}(\Omega) \cap \mathrm{W}_0^{2,2}(\Omega). \tag{1-5}$$

Then, $\mathcal{W}_0^{2,\infty}(\Omega)$ becomes a dual Banach space with separable predual when endowed with the norm (we postpone a thorough discussion until Section 2.4)

$$\|u\|_{\mathcal{W}_0^{2,\infty}(\Omega)} := \|\mathrm{A}{:}\mathrm{D}^2 u\|_{\mathrm{L}^\infty(\Omega)}.$$

Definition 1 (global, narrow, local minimisers). Consider a function $u \in \mathcal{W}^{2,\infty}(\Omega)$.

(i) We say that u is a (*global*) *minimiser* of (1-1) when

$$\mathrm{F}_\infty(u) \le \mathrm{F}_\infty(u+\psi) \quad \text{for all } \psi \in \mathcal{W}_0^{2,\infty}(\Omega).$$

(ii) We say that u is a *narrow minimiser* of (1-1) when there exists $\rho > 0$ such that

$$\mathrm{F}_\infty(u) \le \mathrm{F}_\infty(u+\psi) \quad \text{for all } \psi \in \mathbb{B}_\rho^{\mathrm{W}^{1,\infty}(\Omega)} \cap \mathcal{W}_0^{2,\infty}(\Omega).$$

(iii) We say that u is a *local minimiser* of (1-1) when there exists $\rho > 0$ such that

$$\mathrm{F}_\infty(u) \le \mathrm{F}_\infty(u+\psi) \quad \text{for all } \psi \in \mathbb{B}_\rho^{\mathcal{W}_0^{2,\infty}(\Omega)}.$$

(iv) If additionally the inequality in (i), (ii), or (iii) is strict when $\psi \neq 0$, then u will be called a *strict global*, *narrow*, *or local minimiser*, respectively.

One can define global/narrow/local minimisers of integral functionals in an analogous fashion. Evidently, every global minimiser is narrow, and as a consequence of Lemma 9 that follows, every narrow minimiser is local, but in general neither of the converse implications are true. The significance of the concept of narrow minimisers is justified by the fact that they satisfy stronger existence, uniqueness and stability properties than local minimisers. Narrow minimisers relate to the concept of strong local minimisers (see, e.g., [Kristensen and Taheri 2003]), but they are different. The space $\mathrm{W}^{1,\infty}(\Omega)$ appearing in their definition is not essential and can be replaced by any Banach space $(\mathfrak{X}, |||\cdot|||)$ into which $\mathcal{W}_0^{2,\infty}(\Omega)$ has

a compact embedding (e.g., $\mathrm{L}^\infty(\Omega)$, $\mathrm{C}^{0,\gamma}(\overline{\Omega})$, $\mathrm{C}^{1,\gamma}(\overline{\Omega})$, $\mathrm{W}^{2-\varepsilon,p}(\Omega)$, etc.), and analogous results are obtainable in this case too. Let us now state our assumptions:

- $F : \Omega \times (\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}) \to \mathbb{R}$ is a Carathéodory function satisfying
$$F(x, \cdot\,) \in \mathrm{C}^2(\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}) \text{ for a.e. } x \in \Omega \text{ and } F(\,\cdot\,, \eta, \mathrm{p}, \xi) \in \mathrm{L}^\infty(\Omega) \text{ uniformly when } (\eta, \mathrm{p}, \xi) \in \mathcal{M}$$
for any bounded $\mathcal{M} \subseteq \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$. (1-6)

- There exist $C \in \mathrm{C}^{\nearrow}[0, \infty)$ and $c > 0$ such that
$$\begin{aligned} \partial_\xi F(x, \eta, \mathrm{p}, \xi) &\geq c, \\ F(x, \eta, \mathrm{p}, \cdot\,)^{-1}(\{0\}) &\subseteq \big[-C(|\eta| + |\mathrm{p}|), C(|\eta| + |\mathrm{p}|)\big], \\ |\partial_* F(x, \eta, \mathrm{p}, \xi)| &\leq C(|\eta| + |\mathrm{p}| + |\xi|), \\ |\partial^2_{**} F(x, \eta, \mathrm{p}, \xi)| &\leq C(|\eta| + |\mathrm{p}| + |\xi|) \end{aligned}$$
for a.e. $x \in \Omega$ and all $(\eta, \mathrm{p}, \xi) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$. (1-7)

In (1-7) we used the notation $\mathrm{C}^{\nearrow}[0, \infty)$ for the set of continuous increasing functions on $[0, \infty)$ and also the following compact notation for the *partial* gradient $\partial_* F$ and the *partial* Hessian $\partial^2_{**} F$ with respect to the variables $(\eta, \mathrm{p}, \xi) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$:

$$\partial_* F := (0, \partial_\eta F, \partial_{\mathrm{p}} F, \partial_\xi F), \quad \partial^2_{**} F := \left[\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & \partial^2_{\eta\eta} F & \partial^2_{\eta\mathrm{p}} F & \partial^2_{\eta\xi} F \\ 0 & \partial^2_{\mathrm{p}\eta} F & \partial^2_{\mathrm{pp}} F & \partial^2_{\mathrm{p}\xi} F \\ 0 & \partial^2_{\xi\eta} F & \partial^2_{\xi\mathrm{p}} F & \partial^2_{\xi\xi} F \end{array}\right].$$

For notational convenience, let us also introduce the following notation:

$$\begin{aligned} \mathrm{J}^2 u &:= (\,\cdot\,, u, \mathrm{D}u, \mathrm{A}{:}\mathrm{D}^2 u) : \Omega \to \Omega \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}, \\ \mathrm{J}^2_* u &:= (u, \mathrm{D}u, \mathrm{A}{:}\mathrm{D}^2 u) : \Omega \to \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}. \end{aligned}$$

We will call $\mathrm{J}^2 u$ the (second-order) *jet* of u and $\mathrm{J}^2_* u$ the (second-order) *reduced jet* of u. In this notation, the functional (1-1) can be compactly rewritten as

$$\mathrm{F}_\infty(u) = \|F(\mathrm{J}^2 u)\|_{\mathrm{L}^\infty(\Omega)}.$$

Finally, let $\mathrm{sgn} : \mathbb{R} \to [-1, 1]$ be the sign function, defined by $\mathrm{sgn}(t) := t/|t|$ for $t \neq 0$ and $\mathrm{sgn}(0) := 0$. We may now state our first main result, which identifies a system of equations as a sufficient condition for minimality. Given $u_\infty \in \mathcal{W}^{2,\infty}(\Omega)$ and $f_\infty \in \mathrm{L}^1(\Omega)$, we consider the differential equations

$$F(\mathrm{J}^2 u_\infty) = \mathrm{F}_\infty(u_\infty)\,\mathrm{sgn}(f_\infty) \quad \text{a.e. on } \Omega, \tag{1-8}$$

$$\mathrm{div}\left(\mathrm{A}\mathrm{D} f_\infty - \frac{\partial_{\mathrm{p}} F(\mathrm{J}^2 u_\infty)}{\partial_\xi F(\mathrm{J}^2 u_\infty)} f_\infty\right) + \frac{\partial_\eta F(\mathrm{J}^2 u_\infty)}{\partial_\xi F(\mathrm{J}^2 u_\infty)} f_\infty = 0 \quad \text{in } (\mathrm{C}^\infty_c(\Omega))^*. \tag{1-9}$$

In the above strongly coupled system, (1-8) is a second-order fully nonlinear PDE satisfied in the strong a.e. sense by u_∞, and (1-9) is a second-order linear divergence elliptic PDE for f_∞ satisfied in the sense of distributions. The function f_∞ should be regarded as a "parameter" corresponding to u_∞ (this will be the minimiser).

Theorem 2 (sufficiency of the equations). *Consider the functional* (1-1), *where the supremand F is assumed to satisfy* (1-6)–(1-7). *Let $u_\infty \in \mathcal{W}^{2,\infty}(\Omega)$ and $f_\infty \in \mathrm{L}^1(\Omega) \setminus \{0\}$ be given functions which satisfy the system of equations* (1-8)–(1-9). *Then, for any compact set $\mathcal{K} \subseteq \mathbb{S}^{\mathcal{W}^{2,\infty}_0(\Omega)}$, there exists $\rho > 0$ such that*

$$\mathrm{F}_\infty(u_\infty) < \mathrm{F}_\infty(u_\infty + t\psi), \quad \psi \in \mathcal{K}, \ \ 0 < t < \rho. \tag{1-10}$$

We note that the conclusion (1-10) is weaker than local minimisers (compare Definition 1(iii)), as it requires a compact set of directions of variations. However, since $-\mathcal{K}$ is also compact if $\mathcal{K}$ is compact, inequality (1-10) is in fact true for all $t \in (-\rho, \rho) \setminus \{0\}$ (perhaps for a smaller $\rho > 0$). In order to obtain that u_∞ is a global/narrow/local minimiser, we need some stronger hypotheses on F; see Theorem 3 that follows. Nevertheless, in Section 2.5 we provide a necessary and sufficient condition for strong compactness in $\mathcal{W}^{2,\infty}_0(\Omega)$, along the lines of the classical Ascoli–Arzela and Fréchet–Kolmogorov theorems. In Corollary 15 (Section 2.4) we demonstrate particular interesting choices of compact sets of variations.

We now state our second main result, which gives sufficient conditions for u_∞ to be a local/narrow/global minimiser for F_∞ under additional assumptions on F.

Theorem 3 (sufficiency of the equations and strictness of minimisers). *In the setting of Theorem 2 and under the same assumptions, let $L, R \subseteq \mathbb{R}$ denote arbitrary open intervals, let $M \subseteq \mathbb{R}^n$ denote an open bounded convex set, and consider the following hypotheses*:

- *There exists L, M, R such that* $\operatorname{ess}(\mathrm{J}^2_* u_\infty(\Omega)) \subseteq L \times M \times R$,
$$\text{and there exists } \bar{p} \geq 2 \text{ such that } |F(x,\cdot)|^{\bar{p}} \text{ is convex on } L \times M \times R \text{ for a.e. } x \in \Omega. \tag{1-11}$$
- *There exists L, M with* $\operatorname{ess}((u_\infty, \mathrm{D}u_\infty)(\Omega)) \subseteq L \times M$, *and, for all $R \subseteq \mathbb{R}$,*
$$\text{there exists } \bar{p} \geq 2 \text{ such that } |F(x,\cdot)|^{\bar{p}} \text{ is convex on } L \times M \times R \text{ for a.e. } x \in \Omega. \tag{1-12}$$
- *For all L, M, R,*
$$\text{there exists } \bar{p} \geq 2 \text{ such that } |F(x,\cdot)|^{\bar{p}} \text{ is convex on } L \times M \times R \text{ for a.e. } x \in \Omega. \tag{1-13}$$

Then, the following are true:

(i) *If* (1-11) *is also satisfied, then u_∞ is a strict local minimiser of* F_∞, *and hence it is unique within the neighbourhood of $\mathcal{W}^{2,\infty}_{u_\infty}(\Omega)$ that minimises.*

(ii) *If* (1-12) *is also satisfied, then u_∞ is a strict narrow minimiser of* F_∞, *and hence it is unique within the neighbourhood of $\mathcal{W}^{2,\infty}_{u_\infty}(\Omega)$ that minimises.*

(iii) *If* (1-13) *is also satisfied, then u_∞ is a strict global minimiser of* F_∞, *and hence it is globally unique in the space $\mathcal{W}^{2,\infty}_{u_\infty}(\Omega)$.*

The meaning of the inclusions in (1-11) and (1-12) are understood in the sense of the essential range of measurable functions, whilst the statements (i) and (ii) are meant in the sense that there is a neighbourhood around u_∞ within which it is the unique local/narrow minimiser.

In Section 2.2.1 we provide equivalent and sufficient conditions for the convexity of $|F(x,\cdot)|^{\bar{p}}$ for any $\bar{p} \geq 2$; see (2-1) and (2-3). In particular, this requirement does *not* impose convexity on $|F(x,\cdot)|$, but rather a semiconvexity-type condition, but it does however impose level-convexity with respect to the leading term involving the Hessian, which is natural for this class of problems.

In Section 2.2.3 we discuss explicit examples of functions which satisfy our assumptions. For convenience, we define the affine space counterpart of $\mathcal{W}^{2,\infty}_0(\Omega)$: namely for a given $u_0 \in \mathcal{W}^{2,\infty}(\Omega)$, we set

$$\mathcal{W}^{2,\infty}_{u_0}(\Omega) := u_0 + \mathcal{W}^{2,\infty}_0(\Omega). \tag{1-14}$$

Now we consider the opposite problem regarding the necessity of equations (1-8)–(1-9) for minimality as well as the related problem of existence of minimisers given appropriate Dirichlet boundary conditions on $\partial\Omega$. With this aim, we need to strengthen our assumptions on F. In addition to (1-6)–(1-7), we will suppose that

- there exists $c > 0$ and $\alpha \in (0,1)$ such that
$$\partial^2_{\xi\xi} F(x,\eta,\mathrm{p},\xi) \geq -\frac{1}{c} \text{ and } F(x,\eta,\mathrm{p},\cdot)^{-1}(\{0\}) \subseteq \left[-\frac{(|\eta|+|\mathrm{p}|)^\alpha+1}{c}, +\frac{(|\eta|+|\mathrm{p}|)^\alpha+1}{c}\right]$$
for a.e. $x \in \Omega$ and all $(\eta,\mathrm{p},\xi) \in \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}$, (1-15)

and also that

- there exists $C \in \mathrm{C}^\uparrow[0,\infty)$ such that
$$|\partial_\xi F(x,\eta,\mathrm{p},\xi)| \leq C(|\eta|+|\mathrm{p}|),$$
$$|\partial_{\mathrm{p}} F(x,\eta,\mathrm{p},\xi)| \leq C(|\eta|+|\mathrm{p}|)(1+|\xi|),$$
$$|\partial_\eta F(x,\eta,\mathrm{p},\xi)| \leq C(|\eta|+|\mathrm{p}|)(1+|\xi|)$$
for a.e. $x \in \Omega$ and all $(\eta,\mathrm{p},\xi) \in \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}$. (1-16)

Finally, we will need to assume that the restriction $u_0|_{\partial\Omega}$ of the boundary data determined by $u_0 \in \mathcal{W}^{2,\infty}_0(\Omega)$, as well as the boundary $\partial\Omega$ itself, are more regular. So we also assume that

- the boundary $\partial\Omega$ is C^2 and the restriction $u_0|_{\partial\Omega}$ is C^2 on $\partial\Omega$. (1-17)

We will discuss the exact meaning of the regularity assumption (1-17) in Section 2.2.2. Below is our third main result, which concerns the necessity of equations (1-8)–(1-9) for minimality.

Theorem 4 (necessity of the equations). *Suppose that* $\{F, \partial\Omega, u_0\}$ *satisfy* (1-6)–(1-7) *and also satisfy* (1-15)–(1-17). *Let* $u_\infty \in \mathcal{W}^{2,\infty}_{u_0}(\Omega)$ *be a narrow minimiser of* (1-1). *Then, there exists* $f_\infty \in \mathrm{L}^1(\Omega)\setminus\{0\}$ *such that the pair of functions* (u_∞, f_∞) *satisfies equations* (1-8)–(1-9) *and* $\{f_\infty = 0\}$ *is countably* $(n-1)$*-rectifiable* (*and hence a Lebesgue nullset*). *In addition,* (1-9) *is satisfied weakly in* $(\mathcal{W}^{2,\infty}_0(\Omega))^*$, *whilst* $f_\infty \in \mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega)$ *for all* $p \in (1,\infty)$, *and the set* $\{f_\infty = 0\}$ *is closed.*

By combining Theorems 3 and 4, we have the following important uniqueness assertion for narrow and global minimisers (Definition 1).

Theorem 5 (uniqueness of narrow and global minimisers). *Suppose that* $\{F, \partial\Omega, u_0\}$ *satisfy* (1-6)–(1-7) *and also* (1-15)–(1-17).

(I) *If* (1-12) *holds true, then any narrow minimiser* $u_\infty \in \mathcal{W}^{2,\infty}_{u_0}(\Omega)$ *of* (1-1) *is the unique narrow minimiser within* $\mathbb{B}_\rho^{\mathrm{W}^{1,\infty}(\Omega)}(u_\infty) \cap \mathcal{W}^{2,\infty}_{u_0}(\Omega)$ (*perhaps for a smaller* $\rho > 0$).

(II) *If* (1-13) *holds true, then global minimisers of* (1-1) *in the space* $\mathcal{W}^{2,\infty}_{u_0}(\Omega)$ *are unique.*

Now we consider the question of regularity of local/narrow/global minimisers and of solutions to the system of equations (1-8)–(1-9). In [Katzourakis and Pryer 2020, Section 8], *it was proved that even when* $n = 1$ *and no lower-order terms are present, the minimisers for* $u \mapsto \|u''\|_{L^\infty(a,b)}$ *in* $\mathrm{W}^{2,\infty}_{u_0}(a, b)$ *are absolute, unique, but non-*C^2. More precisely, *minimisers are always piecewise quadratic with exactly one singular point* $\bar{x} \in (a, b)$, *and they are smooth if and only if the boundary data* $\{u_0(a), u_0'(a), u_0(b), u_0'(b)\}$ *can be interpolated by a quadratic polynomial, in which case this is the minimiser*. In the generality of this work, we have the following (partial) regularity result.

Corollary 6 (partial regularity). *Suppose that* F *satisfies* (1-6)–(1-7) *and also that* $F \in \mathrm{C}^k(\Omega \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R})$ *for some* $k \in \mathbb{N}$.

(a) *Let* $(u_\infty, f_\infty) \in \mathcal{W}^{2,\infty}(\Omega) \times (\mathrm{L}^1(\Omega) \setminus \{0\})$ *satisfy equations* (1-8)–(1-9). *Then,*

$$u_\infty \in \mathrm{C}^{k+2,\gamma}(\Omega \setminus \{f_\infty = 0\}) \quad \textit{for all } \gamma \in (0, 1),$$

and the set $\{f_\infty = 0\}$ *is closed, countably* $(n-1)$*-rectifiable, and, in particular, a Lebesgue nullset in* Ω.

(b) *Under the additional assumptions of Theorem 5, any narrow/global minimiser* u_∞ *satisfies the above* $C^{k+2,\gamma}$ *partial regularity on* $\Omega \setminus \{f_\infty = 0\}$.

We now connect the system of equations (1-8)–(1-9) to the Aronsson equation which is associated with (1-1). By a standard formal computational argument based on L^p approximations and an appropriate rescaling of the Euler–Lagrange equations (as, e.g., in [Katzourakis and Pryer 2020; Aronsson and Barron 2012]), we can derive the Aronsson equation for (1-1), which is a fully nonlinear third-order PDE:

$$\partial_\xi F(\mathrm{J}^2 u_\infty)\mathrm{A}{:}\mathrm{D}(|F(\mathrm{J}^2 u_\infty)|) \otimes \mathrm{D}(|F(\mathrm{J}^2 u_\infty)|) = 0 \quad \text{in } \Omega. \tag{1-18}$$

Since by assumption (1-7) we have $\partial_\xi F > 0$, and also $\mathrm{A} \in \mathbb{R}^{n\times n}$ is a (strictly) positive symmetric matrix, (1-18) is (at least formally) equivalent to

$$\mathrm{D}(|F(\mathrm{J}^2 u_\infty)|) = 0 \quad \text{in } \Omega.$$

Therefore, *the Aronsson equation* (1-18) *corresponding to* (1-1) *is equivalent to the statement that* $|F(\mathrm{J}^2 u_\infty)|$ *is constant on connected components of* Ω. *This is a genuinely weaker statement than the statement encoded in the system of equations* (1-8)–(1-9). Indeed, on the one hand, (1-8) guarantees that $|F(\mathrm{J}^2 u_\infty)|$ is not only constant throughout Ω, but in fact it is everywhere equal to the exact value

$\mathrm{F}_\infty(u_\infty)$ (even if Ω has more than one connected component). On the other hand, (1-9) provides sharper information about the sign of $F(\mathrm{J}^2 u_\infty)$, which is determined by the nodal set of the weak solution f_∞ to this divergence PDE. By the results in [Hardt and Simon 1989], the nodal set is actually a lower-dimensional set and in particular has zero Lebesgue measure. *Therefore, the Aronsson equation is only a necessary but not a sufficient condition for minimality.* More precisely, directly from Theorem 4, we have that any narrow minimiser of (1-1) is a strong solution to (1-18) in *contracted* form (since $|F(\mathrm{J}^2 u_\infty)|$ is constant throughout Ω and hence is differentiable). Further, by using [Katzourakis 2017, Theorem 30, p. 665] and by arguing as in [Katzourakis and Pryer 2020, Section 7, pp. 18-20], one can additionally show that u_∞ is in fact a (generalised) $\mathcal{D}$-solution to the *expanded* counterpart of the third-order fully nonlinear Aronsson equation (1-18).

Now we supplement Theorem 4 with certain approximability properties, originating from the method of construction through penalised L^p-approximating functionals. Our method is based on a novel adaptation of the trademark method of L^p approximations, which consists of the introduction of an extra penalisation term. To the best of our knowledge, this adaptation first appeared in the literature in [Moser 2022]. *The advantage of this modification is that it forces the approximating* L^p *minimisers to converge to a preselected* L^∞ *minimiser as* $p \to \infty$. For any $u_0 \in \mathcal{W}^{2,\infty}(\Omega)$, $\bar{u} \in \mathcal{W}^{2,\infty}_{u_0}(\Omega)$, $p > n$, and $\epsilon > 0$, we set

$$\mathrm{F}_{p,\epsilon,\bar{u}}(u) := \left(\fint_\Omega |F(\mathrm{J}^2 u)|^p \, \mathrm{d}\mathcal{L}^n \right)^{\frac{1}{p}} + \frac{\epsilon}{2} \fint_\Omega |u - \bar{u}|^2 \, \mathrm{d}\mathcal{L}^n \tag{1-19}$$

for all $u \in \mathrm{W}^{2,p}_{u_0}(\Omega)$ $(= \mathrm{W}^{2,p}_{\bar{u}}(\Omega))$.

Theorem 7 (necessity of the equations and L^p approximation with penalisation). *In the setting of Theorem 4 and under the same hypotheses, we have that the pair* (u_∞, f_∞) *satisfies the following approximation properties*:

(i) *There exists a family of functions* $\{u_p\}_{n<p<\infty} \subseteq \mathrm{W}^{2,n}_{u_0}(\Omega)$, *where, for each* p, u_p *is a narrow minimiser of the penalised* L^p *functional* $\mathrm{F}_{p,1,u_\infty}$ *(given by* (1-19)*) in the space* $\mathrm{W}^{2,p}_{u_0}(\Omega)$ *such that*

$$\begin{cases} \mathrm{F}_{p,1,u_\infty}(u_p) \to \mathrm{F}_\infty(u_\infty), \\ u_p \rightharpoonup u_\infty \quad \text{in } \mathrm{W}^{2,q}(\Omega) \text{ for all } q \in (1,\infty), \\ u_p \to u_\infty \quad \text{in } \mathrm{C}^{1,\gamma}(\overline{\Omega}) \text{ for all } \gamma \in (0,1) \end{cases} \tag{1-20}$$

along a sequence of indices $(p_j)_1^\infty$ *satisfying* $p_j \to \infty$ *as* $j \to \infty$.

(ii) *For any* $p \in (n, \infty)$, *the narrow minimiser* u_p *is a weak solution to the Euler–Lagrange equations of* $\mathrm{F}_{p,1,u_\infty}$, *namely*

$$e_p^{1-p} \int_\Omega [(|F|^{p-2} F \partial_* F)(\mathrm{J}^2 u_p) \cdot \mathrm{J}^2_* \phi] \, \mathrm{d}\mathcal{L}^n + \int_\Omega (u_p - u_\infty)\phi \, \mathrm{d}\mathcal{L}^n = 0 \tag{1-21}$$

for all $\phi \in \mathrm{W}^{2,p}_0(\Omega)$, *where*

$$e_p := \left(\fint_\Omega |F(\mathrm{J}^2 u_p)|^p \, \mathrm{d}\mathcal{L}^n \right)^{\frac{1}{p}}. \tag{1-22}$$

Further, by setting

$$f_p := e_p^{1-p}(|F|^{p-2}F\partial_\xi F)(\mathrm{J}^2 u_p), \tag{1-23}$$

we have

$$\begin{cases} f_p \mathcal{L}^n \llcorner_\Omega \overset{*}{\rightharpoonup} f_\infty \mathcal{L}^n \llcorner_\Omega & \text{in } \mathcal{M}(\overline{\Omega}), \\ f_p \to f_\infty & \text{in } \mathrm{L}^q_{\mathrm{loc}}(\Omega) \text{ for all } q \in (1, n'), \end{cases} \tag{1-24}$$

and also

$$\mathrm{A}{:}\mathrm{D}^2 u_p \to \mathrm{A}{:}\mathrm{D}^2 u_\infty \quad \text{in } \mathrm{L}^q(\Omega \setminus \{f_\infty = 0\}) \text{ for all } q \in (1, \infty)$$

along a sequence of indices $(p_j)_1^\infty$ which satisfies $p_j \to \infty$ as $j \to \infty$.

Our final result concerns an immediate generalisation of all our previous results to more general functionals, which may involve supremands which are perhaps nondifferentiable and with arbitrary growth at infinity.

Corollary 8 (extension to nondifferentiable supremands with arbitrary growth). *Let us consider the supremal functional*

$$\overline{\mathrm{F}}_\infty(u) := \|\mathcal{F}(\mathrm{J}^2 u)\|_{\mathrm{L}^\infty(\Omega)}, \quad u \in \mathcal{W}^{2,\infty}(\Omega),$$

and suppose the supremand $\mathcal{F} : \Omega \times (\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}) \to \mathbb{R}$ can be factorised in a way that

- *there exists a strictly increasing odd function $\Phi : \mathbb{R} \to \mathbb{R}$*
  and a Carathéodory function $F : \Omega \times (\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}) \to \mathbb{R}$ such that $\mathcal{F} = \Phi \circ F$,

where F satisfies (1-6)–(1-7). *Then, all results stated earlier in Theorems 2–5, Corollary 6 and Theorem 7 hold true for $\overline{\mathrm{F}}_\infty$ as well if F satisfies the additional corresponding assumptions* (1-11), (1-12), (1-13), (1-15) *and* (1-16).

Let us now discuss the state of the art for higher-order (scalar-valued) L^∞ variational problems. The relevant higher-order Aronsson equation was first derived formally by Aronsson himself and Barron in [Aronsson and Barron 2012], but noting its peculiarities, it was not explored further. In the paper [Katzourakis and Pryer 2020] by the first author and Pryer, higher-order L^∞ problems were considered systematically for the first time, and several initial results were established for the functional $u \mapsto \|H(\mathrm{D}^2 u)\|_{\mathrm{L}^\infty(\Omega)}$, including existence of generalised solutions to the Dirichlet problem for the corresponding fully nonlinear Aronsson equation. For this task, the general framework of $\mathcal{D}$-solutions was utilised, a systematic approach for fully nonlinear PDE systems, introduced by the first author in [Katzourakis 2017]. Papamikos and Pryer [2019] studied special solutions to the same higher-order Aronsson equation, classifying them by using Lie symmetries. In [Katzourakis and Parini 2017], the first author with Parini studied a higher-order ∞-eigenvalue problem, generalised subsequently to the fully nonlinear setting by the first author with Clark [Clark and Katzourakis 2025]. In [Clark et al. 2022], Clark and Muha jointly with the first author, inspired by questions arising in data assimilation, studied a second-order PDE constrained L^∞ problem involving the Navier–Stokes equations. Higher-order L^∞ variational problems arising in Riemannian geometry were previously studied by the second author with Schwetlick and also by Sakellaris in [Moser and Schwetlick 2012; Sakellaris 2017], and more recently

again by the second author in [Moser 2022]. In [Katzourakis and Moser 2019], the authors considered the special case of (1-1) given by $u \mapsto \|\mathcal{F}(\,\cdot\,, \Delta u)\|_{\mathrm{L}^\infty(\Omega)}$ which, unsurprisingly, is immensely simpler than the case of F_∞ which involves explicit $(u, \mathrm{D}u)$-dependence. Finally, we note that in the most recent paper [Carano et al. 2025], the authors jointly with Carano are discussing extensions of the results from [Katzourakis and Moser 2019] without lower-order terms to the vector-valued case but for a more general second-order linear differential operator.

This paper is organised as follows. This rather long introduction is followed by Section 2, in which we discuss in some detail the preliminaries, our functional space setup, our assumptions, and also where we provide explicit classes of integrands (Section 2.2). Therein we also establish some auxiliary results of independent interest. In particular, we identify $\mathcal{W}^{2,\infty}_0(\Omega)$ as a dual Banach space with separable predual, and we also prove a general lemma regarding the existence of predual spaces for weak*-closed subspaces in dual Banach spaces (Section 2.3). We also establish certain auxiliary results of independent interest, including characterisations and examples of compact sets in $\mathcal{W}^{2,\infty}_0(\Omega)$ (Section 2.4), approximability of constrained minimisers in L^∞ via L^p-constrained minimisers as $p \to \infty$ (Section 2.5), a lemma about strong compactness in $\mathrm{L}^1_{\mathrm{loc}}(\Omega)$ of solutions to the divergence PDE with L^1 bounded coefficients (Section 2.6), and some results regarding the size of nodal sets and other properties of solutions to equations of the form (1-9) (Section 2.7). Finally, in Section 3 we establish all our main results, whose statements are given earlier in this introduction.

We close this introduction by noting that, in our recent paper [Katzourakis and Moser 2023], we consider the case of the functional $u \mapsto \|\mathrm{A}(\,\cdot\,):\mathrm{D}^2u + b(\,\cdot\,, u, \mathrm{D}u)\|_{\mathrm{L}^\infty(\Omega)}$, which involves a semilinear differential operator. Even though there is some overlap of [loc. cit.] with the results obtained herein, the phenomena are significantly different. In particular, in [loc. cit.] the natural minimality notion involves *almost minimisers* with a first-order error term and not global/narrow/local minimisers as herein. Additionally, unlike the present work, uniqueness in [loc. cit.] is an open question for almost minimisers. Therefore, for endogenous mathematical reasons, as well as for practicality purposes of space and size, we treat these related problems separately, as an attempt for a unified approach would considerably complicate and impair the exposition rather than enhance it.

2. Preliminaries and auxiliary results

We begin by discussing in some detail our notation and our functional space setup, and we also establish some auxiliary results of independent interest.

2.1. ***Preliminaries and notational conventions.*** Let ":" denote the Euclidean (Frobenius) inner product in $\mathbb{R}^{n\times n}$, namely $A:B = \mathrm{tr}(A^\top B)$. For $a, b \in \mathbb{R}^n \equiv \mathbb{R}^{n\times 1}$, we define $a \otimes b := ab^\top$. Further, for any open $\Omega \subseteq \mathbb{R}^n$, for a twice (weakly) differentiable function $u : \Omega \to \mathbb{R}$, the notation $\mathrm{D}u = (\mathrm{D}_1u, \dots, \mathrm{D}_nu) : \Omega \to \mathbb{R}^n$ will denote its gradient, and $\mathrm{D}^2u = (\mathrm{D}^2_{ij}u)^n_{i,j=1} : \Omega \to \mathbb{R}^{n\times n}$ will denote its Hessian, where obviously $\mathrm{D}_i \equiv \partial/\partial x_i$. For any function $u : \mathbb{R}^n \to \mathbb{R}^n$ and $z \in \mathbb{R}^n$, we define the translation $\tau_z u$ by setting $\tau_z u(x) := u(x + z)$. The closed ball, open ball and sphere of radius ρ centred at a point $x \in \mathfrak{X}$ of a normed space $(\mathfrak{X}, \|\cdot\|)$ will be denoted by $\overline{\mathbb{B}}^{\mathfrak{X}}_\rho(x)$, $\mathbb{B}^{\mathfrak{X}}_\rho(x)$ and $\mathbb{S}^{\mathfrak{X}}_\rho(x)$, respectively. If $\rho = 1$ and/or

$x = 0$, we will not explicitly display these variables, and if $\mathfrak{X}$ is clear from the context, it may be omitted as well. The functional space notation we employ is either self-explanatory, or is otherwise a convex combination of standard notation (as, e.g., in [Adams and Fournier 2003; Dacorogna 2008; Fonseca and Leoni 2007; Katzourakis and Vărvărucă 2018; Rudin 1991; Toland 2020]). For example, $[\,\cdot\,]_{\mathrm{C}^\alpha(\bar{\Omega})}$ denotes the Hölder seminorm over Ω, $\mathrm{C}^{\nearrow}[0,\infty)$ denotes the set of continuous increasing functions on $[0,\infty)$, and $\mathrm{C}^{\nearrow}_{0^+}[0,\infty)$ denotes the subset of those that vanish at the origin (i.e., moduli of continuity). The n-dimensional Lebesgue measure will be denoted by $\mathcal{L}^n$ and the k-dimensional Hausdorff measure by $\mathcal{H}^k$. Further, all L^p norms appearing in this paper with respect to the Lebesgue measure $\mathcal{L}^n$ over Ω will always be the normalised ones, namely with the average over Ω instead of the integral, as, e.g., in (1-19). Let now $(\mathfrak{X}, \|\cdot\|)$ be a normed space with dual Banach space $(\mathfrak{X}^*, \|\cdot\|_*)$. Given a subset $\mathcal{A} \subseteq \mathfrak{X}$ and a subset $\mathcal{B} \subseteq \mathfrak{X}^*$, the annihilator $\mathcal{A}^\perp \subseteq \mathfrak{X}^*$ and the pre-annihilator ${}^\perp\mathcal{B} \subseteq \mathfrak{X}$ are defined as

$$\mathcal{A}^\perp := \{f \in \mathfrak{X}^* : f(x) = 0, \forall\, x \in \mathcal{A}\} = \{f \in \mathfrak{X}^* : f^{-1}(\{0\}) \supseteq \mathcal{A}\},$$
$${}^\perp\mathcal{B} := \{x \in \mathfrak{X} : f(x) = 0, \forall\, f \in \mathcal{B}\} = \bigcap \{f^{-1}(\{0\}) : f \in \mathcal{B}\}.$$

(Details can be found, e.g., in [Rudin 1991, Chapters 1 and 4].) Then, $\mathcal{A}^\perp$ is a weak*-closed subspace in $\mathfrak{X}^*$, and ${}^\perp\mathcal{B}$ is a norm-closed subspace in $\mathfrak{X}$. If $\mathcal{N} \subseteq \mathfrak{X}$ is any norm-closed subspace, the equivalence relation $x' \sim x'' \Leftrightarrow x'' - x' \in \mathcal{N}$ partitions $\mathfrak{X}$ into cosets, the collection of which comprises the quotient vector space $\mathfrak{X}/_{\mathcal{N}} := \{x + \mathcal{N} : x \in \mathfrak{X}\}$. Let $\pi : \mathfrak{X} \to \mathfrak{X}/_{\mathcal{N}}$ denote the quotient map, given by $\pi(x) := x + \mathcal{N}$. Then, $\mathfrak{X}/_{\mathcal{N}}$ becomes a normed space when endowed with the quotient norm

$$\|\pi(x)\|_{\mathfrak{X}/_{\mathcal{N}}} := \inf\{\|x - n\| : n \in \mathcal{N}\} = \operatorname{dist}(x, \mathcal{N}), \quad x \in \mathfrak{X}.$$

2.2. ***Discussion of assumptions and examples.*** We now focus on assumptions (1-11)–(1-13), which involve convexity of $|F(x,\cdot)|^{\bar{p}}$ on convex subsets of $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$ for large $\bar{p} \geq 2$, and also on assumption (1-17), which requires additional regularity of the restriction of the boundary condition $u_0|_{\partial\Omega}$ to $\partial\Omega$.

Before delving into that, let us briefly discuss the implications of our assumption that the matrix A is symmetric and positive. Whilst the requirement to not be allowed to vary with x as a diffusion coefficient $\mathrm{A}(x)$ is not essential and the choice to be constant was only as a technical simplification (in fact, as commented earlier at the end of the introduction, we discuss this case in our companion paper [Katzourakis and Moser 2023]), the symmetry and positivity are both essential for our methods to work, as we are largely based on elliptic estimates, all of which seem to fail otherwise.

2.2.1. *The restricted convexity assumptions* (1-11)–(1-13). A necessary and sufficient condition for the convexity of $|F(x,\cdot)|^{\bar{p}}$ on an open convex subset $L \times M \times R \subseteq \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$ is the (semiconvexity-type) matrix inequality

$$F\partial^2_{**}F + (\bar{p} - 1)\partial_* F \otimes \partial_* F \geq 0, \tag{2-1}$$

which must hold a.e. on $\Omega \times L \times M \times R$. The equivalence follows from the identity

$$\partial^2_{**}(|F|^p) = p|F|^{p-2}(F\partial^2_{**}F + (p-1)\partial_* F \otimes \partial_* F), \tag{2-2}$$

which is valid for any $p \geq 2$ a.e. on $\Omega \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$. The equivalence is clear on $\{F(x, \cdot) \neq 0\}$ for a.e. $x \in \Omega$, as the (partial) Hessian of F is nonnegative if and only if (2-1) holds true, whilst on $\{F(x, \cdot) = 0\}$ inequality (2-1) is automatically always true, and (2-2) reduces to the (partial) Hessian vanishing for the C^2 function $F(x, \cdot)$. As an additional consequence, if $|F(x, \cdot)|^p$ is convex on $L \times M \times R$, so is $|F(x, \cdot)|^q$ for any $q \geq p$. In the case that there is no explicit gradient dependence (and hence no p-variable), a sufficient condition for (2-1) is the following: Let $\star \in \mathbb{R}^{2\times 2}$ be the Hodge matrix satisfying $\star^2 = -\mathrm{I}$ (rotation by $\pi/2$ in $\mathbb{R}^2$). Then, (2-1) follows from the condition

- $(\partial_* F, \star\, \partial_* F)$ is an orthogonal eigenframe of $\partial^2_{**}(|F|^2)$ with eigenvalues(σ_1, σ_2)
  satisfying $\sigma_2 \geq 0$, $\sigma_1 \geq -m$ for $m \in \mathbb{N}$ a.e. on $\Omega \times L \times R$. (2-3)

Condition (2-3) is a stronger requirement than (2-1) but might be easier to verify. The implication "(2-3) $\Rightarrow$ (2-1)" follows by the spectral theorem applied to $\partial^2_{**}(|F|^2)$,

$$\partial^2_{**}(|F|^2) = \frac{\sigma_1}{|\partial_* F|^2}\partial_* F \otimes \partial_* F + \frac{\sigma_2}{|\partial_* F|^2}(\star\, \partial_* F) \otimes (\star\, \partial_* F) \geq \frac{\sigma_1}{|\partial_* F|^2}\partial_* F \otimes \partial_* F \geq -\frac{m}{c}\partial_* F \otimes \partial_* F,$$

which holds a.e. on $\Omega \times L \times R$. This implies (2-1) when $\bar{p} \geq (m/c) + 1$. In the above inequalities we have used that $|\partial_* F| \geq c > 0$, which follows from (1-7).

2.2.2. *The Whitney differentiability assumption* (1-17). By the (Lipschitz) regularity of $\partial\Omega$, we have the Sobolev embedding $\mathrm{W}^{2,\infty}(\Omega) \subseteq \mathrm{C}^{1,\gamma}(\bar{\Omega})$ for all $\gamma \in (0, 1)$ in the usual sense, namely the precise representative from each a.e.-equivalence class has an extension to a function in $\mathrm{C}^{1,\gamma}(\mathbb{R}^n)$. In particular, $u_0|_{\partial\Omega}$ is at the outset $\mathrm{C}^{1,\gamma}$ on $\partial\Omega$ for all $\alpha \in (0, 1)$. The assumption that $u_0|_{\partial\Omega}$ is C^2 on the compact set $\partial\Omega$ is understood in the standard sense of Whitney [1934, Section 3, p. 64], meaning that it possesses (uniform) second-order Taylor expansions. More precisely, we assume that there exists a symmetric matrix-valued map $X \in \mathrm{C}(\partial\Omega; \mathbb{R}^{n\times n})$ and a modulus of continuity $\omega \in \mathrm{C}^{\uparrow}_0[0, \infty)$ such that

$$\begin{aligned} \left|u_0(x + z) - u_0(x) - \mathrm{D}u_0(x)z - \tfrac{1}{2}X(x) : z \otimes z\right| &\leq \omega(|z|)|z|^2, \\ |\mathrm{D}u_0(x + z) - \mathrm{D}u_0(x) - X(x)z| &\leq \omega(|z|)|z| \end{aligned}$$

for all $x, z \in \mathbb{R}^n$ such that $x, x + z \in \partial\Omega$. It can be easily seen that, due to the smoothness of $\partial\Omega$, C^1 differentiability for a scalar-valued function in the above sense is equivalent to tangential differentiability on $\partial\Omega$ (regarded as a C^2 compact manifold with the induced Riemannian metric). However, requiring a function to be C^2 on $\partial\Omega$, or requiring its Euclidean gradient vector field be tangentially C^1 on $\partial\Omega$, is in general a slightly stronger requirement than being tangentially C^2 on $\partial\Omega$.

2.2.3. *Explicit classes of supremands.* We demonstrate some classes of functions F which satisfy our assumptions, noting that the results proved herein apply to supremands of the form $\mathcal{F} = \Phi \circ F$ for *any* increasing odd function $\Phi : \mathbb{R} \to \mathbb{R}$ (Corollary 8). For the sake of clarity, we assume no explicit gradient dependence in these examples (and hence no p-variable). A similar (but more complicated) analysis can be performed by assuming explicit dependence on the gradient as well. We note that within this subsection, for the sake of brevity, we will simplify the notation of derivatives for coefficients by denoting them merely as subscripts.

Additive ansatz. Let $A : \mathbb{R} \to \mathbb{R}$ and $a : \Omega \times \mathbb{R} \to \mathbb{R}$ be given. We set

$$F(x, \eta, \xi) := a(x, \eta) + A(\xi).$$

If we have $A \in C^2(\mathbb{R})$ with $A_\xi \geq c > 0$ and for all $x \in \Omega$ we have $a(x, \cdot) \in C^2(\mathbb{R})$ with $a, a_\eta, a_{\eta\eta} \in C(\overline{\Omega} \times \mathbb{R})$, then (1-6)–(1-7) are readily satisfied. In particular, since A is strictly increasing, $F(x, \eta, \xi) = 0$ if and only if $\xi = A^{-1}(-a(x, \eta))$, and therefore

$$F(x, \eta, \cdot)^{-1}(\{0\}) \subseteq \big[A^{-1}\big(-\max_{x \in \overline{\Omega}} a(x, \eta)\big), A^{-1}\big(-\min_{x \in \overline{\Omega}} a(x, \eta)\big)\big].$$

Regarding the fulfilment of (1-11)–(1-12)–(1-13), a direct computation gives that, if we set

$$\mathfrak{F} := F\partial^2_{**}F + K\partial_* F \otimes \partial_* F$$

where $K \geq 1$, then we have $\mathfrak{F} \geq 0$ if and only if (suppressing the arguments (x, η, ξ) for brevity)

$$\operatorname{tr}(\mathfrak{F}) = (a + A)(A_{\xi\xi} + a_{\eta\eta}) + K\big[(A_\xi)^2 + (a_\eta)^2\big] \geq 0, \tag{2-4}$$

$$\det(\mathfrak{F}) = (a + A)^2 A_{\xi\xi} a_{\eta\eta} + K(a + A)\big[a_{\eta\eta}(A_\xi)^2 + A_{\xi\xi}(a_\eta)^2\big] \geq 0. \tag{2-5}$$

Let us also consider the *hypergraph* of the function $A^{-1} \circ (-a)$, namely

$$\mathcal{H} := \{(x, \eta, \xi) \in \Omega \times \mathbb{R} \times \mathbb{R} : \xi \geq A^{-1}(-a(x, \eta))\},$$

which has the property that $F(x, \eta, \xi) \geq 0$ if and only if $(x, \eta, \xi) \in \mathcal{H}$. Then:

(a) Let $L, R \subseteq \mathbb{R}$ be given open intervals such that $\Omega \times L \times R \subseteq \mathcal{H}$ (the hypergraph). If, for all $x \in \Omega$, $a(x, \cdot)$ is convex on L and A is convex on R, then (2-4)–(2-5) are satisfied on $\Omega \times L \times R$ for all intervals $M \subseteq \mathbb{R}$. Hence (1-11) is satisfied.

(b) If a is bounded below by $-m$ for some $m \in \mathbb{R}$, then $\mathcal{M} := \Omega \times \mathbb{R} \times (A^{-1}(m), \infty) \subseteq \mathcal{H}$. If, for all $x \in \Omega$, $a(x, \cdot)$ is convex on $\mathbb{R}$ and A is convex on $(A^{-1}(m), \infty)$, then (2-4)–(2-5) are satisfied on $\mathcal{M}$ for all intervals $M \subseteq \mathbb{R}$. Hence (1-11) is satisfied.

(c) If, for all $x \in \Omega$, $a(x, \cdot)$ is convex on $\mathbb{R}$ and A is also convex on $\mathbb{R}$, then (2-4)–(2-5) are satisfied on the entire hypergraph $\mathcal{H}$ for all intervals $M \subseteq \mathbb{R}$. Hence (1-11) is satisfied.

(d) If, for all $x \in \Omega$, $a(x, \cdot)$ is affine on $\mathbb{R}$ and A is also affine on $\mathbb{R}$, then (2-4)–(2-5) are satisfied on $\Omega \times \mathbb{R} \times \mathbb{R}$ for all intervals $M \subseteq \mathbb{R}$. Hence (1-13) is satisfied.

Regarding the fulfilment of (1-15)–(1-16), they are satisfied when A is semiconvex on $\mathbb{R}$ ($A_{\xi\xi}$ is bounded below) and A_ξ is bounded above, with the growth of $a(x, \cdot)$ smaller than the growth of A at infinity, namely when we have $|A(\xi)| \leq C(1 + |\xi|^p)$ and $|a(x, \eta)| \leq C(1 + |\eta|^q)$ for some $C > 0$ and $0 < q < p$.

Multiplicative ansatz. Let $A : \mathbb{R} \to \mathbb{R}$ and $a : \Omega \times \mathbb{R} \to \mathbb{R}$ be given. We set

$$F(x, \eta, \xi) := a(x, \eta)A(\xi).$$

If $A \in C^2(\mathbb{R})$ satisfies $A_\xi \geq c > 0$ and also for all $x \in \Omega$ we have $a(x, \cdot) \in C^2(\mathbb{R})$, with $a, a_\eta, a_{\eta\eta} \in C(\overline{\Omega} \times \mathbb{R})$ and $a \geq c > 0$, then (1-6)–(1-7) are satisfied with $F(x, \eta, \cdot)^{-1}(\{0\}) = \{0\}$. Regarding the fulfilment of

(1-11)–(1-12)–(1-13), let again $K \geq 1$. Suppressing the arguments (x, η, ξ) for brevity, a computation gives that, if $\mathfrak{F} := F\partial^2_{**}F + K\,\partial_* F \otimes \partial_* F$, then $\mathfrak{F} \geq 0$ if and only if

$$\mathrm{tr}(\mathfrak{F}) = A^2\big[aa_{\eta\eta} + K(a_\eta)^2\big] + a^2\big[AA_{\xi\xi} + K(A_\xi)^2\big] \geq 0, \tag{2-6}$$

$$\begin{aligned}\det(\mathfrak{F}) = (aA)^2\big\{&(AA_{\xi\xi})(aa_{\eta\eta}) - (A_\xi)^2(a_\eta)^2 \\ &+ K(A_\xi)^2\big[aa_{\eta\eta} - (a_\eta)^2\big] + K(a_\eta)^2\big[AA_{\xi\xi} - (A_\xi)^2\big]\big\} \geq 0.\end{aligned} \tag{2-7}$$

(a) Let $L, R \subseteq \mathbb{R}$ be open intervals, and let us note the following elementary identities, which are valid for any $h \in \mathrm{C}^2(\mathbb{R})$: $(|h|^p)'' = p|h|^{p-2}[hh'' + (p-1)(h')^2]$ on $\mathbb{R}$, and $(\ln|h|)'' = h^{-2}[hh'' - (h')^2]$ on $\{h \neq 0\}$. It follows that (2-6)–(2-7) are satisfied on $\Omega \times L \times R$ if $\ln a(x, \cdot\,)$ is convex on L for all $x \in \Omega$, and $\ln|A|$ is convex on R. Therefore, we may select C^2 functions $A := e^B$, $a := e^b$, where $B : R \to \mathbb{R}$ is convex and strictly increasing and $b : \overline{\Omega} \times L \to \mathbb{R}$ is convex in the second variable. We extend A as a strictly increasing function in $\mathrm{C}^2(\mathbb{R})$ which is linear outside an interval $[-m, m]$ for some $m \in \mathbb{N}$ large, and extend also a as an arbitrary C^2 function on $\Omega \times \mathbb{R}$ bounded below. It follows that (1-11) is satisfied.

(b) Let A be affine. It can be easily confirmed that (2-6)–(2-7) can be deduced from

$$(\ln a)_{\eta\eta} \geq \frac{K+1}{K}((\ln a)_\eta)^2.$$

Then, for any interval $L \subseteq \mathbb{R}$ and for any $b \in \mathrm{C}^2(\overline{\Omega} \times \mathbb{R})$ which is bounded below and is strictly convex in η, we can find $\varepsilon > 0$ small such that $a := e^{\varepsilon b}$ satisfies the above inequality on $\Omega \times L$. Hence (2-6)–(2-7) hold true on $\Omega \times L \times \mathbb{R}$, and therefore (1-12) is satisfied. Finally, assumptions (1-15)–(1-16) are also trivially satisfied.

2.3. ***The functional space setup and duality.*** Now we expound on the functional spaces we utilise herein and study some of their properties. We set

$$\mathbb{W}^{2,\infty}_{\mathrm{loc}}(\Omega) := \bigcap_{1<p<\infty} \mathrm{W}^{2,p}_{\mathrm{loc}}(\Omega), \tag{2-8}$$

$$\mathbb{W}^{2,\infty}(\Omega) := \bigcap_{1<p<\infty} \mathrm{W}^{2,p}(\Omega), \tag{2-9}$$

$$\mathbb{W}^{2,\infty}_0(\Omega) := \bigcap_{1<p<\infty} \mathrm{W}^{2,p}_0(\Omega). \tag{2-10}$$

Then, we have $\mathbb{W}^{2,\infty}_{\mathrm{loc}}(\Omega) \supset \mathbb{W}^{2,\infty}(\Omega) \supset \mathbb{W}^{2,\infty}_0(\Omega)$, and all three spaces are Fréchet spaces, topologised by the family of seminorms $\{\|\cdot\|_{\mathrm{W}^{2,p}(\Omega')} : \Omega' \Subset \Omega,\ p \in (1,\infty)\}$. If now $\mathbb{A} : \mathrm{W}^{2,1}(\Omega) \to \mathrm{L}^1(\Omega)$ is the (strong) linear differential operator defined by $\mathbb{A}u := \mathrm{A}{:}\mathrm{D}^2 u$, the spaces defined in (1-2) and (1-5) can be written as

$$\mathcal{W}^{2,\infty}(\Omega) = \mathbb{W}^{2,\infty}(\Omega) \cap \mathbb{A}^{-1}(\mathrm{L}^\infty(\Omega)), \tag{2-11}$$

$$\mathcal{W}^{2,\infty}_0(\Omega) = \mathbb{W}^{2,\infty}_0(\Omega) \cap \mathbb{A}^{-1}(\mathrm{L}^\infty(\Omega)). \tag{2-12}$$

As we will show in Sections 2.3.1–2.3.2, $\mathcal{W}_0^{2,\infty}(\Omega)$ becomes a dual Banach space with (separable) predual $\mathbb{A}(L^1(\Omega))$ and convenient weak* density and compactness properties, when endowed with either of the equivalent norms

$$\begin{aligned} \|u\|_{\mathcal{W}_0^{2,\infty}(\Omega)} &:= \|A{:}D^2u\|_{L^\infty(\Omega)}, \\ \|u\|'_{\mathcal{W}_0^{2,\infty}(\Omega)} &:= \|J_*^2 u\|_{L^\infty(\Omega)}. \end{aligned} \tag{2-13}$$

The use of the spaces "$\mathcal{W}^{2,\infty}, \dots$" instead of the more standard Banach spaces "$W^{2,\infty}, \dots$" is dictated by the failure of the Calderón–Zygmund L^p elliptic estimates for $p = \infty$, which implies that our functional is not coercive in the spaces "$W^{2,\infty}, \dots$".

2.3.1. *Completeness and a priori estimates.* We now study some important properties of the space $\mathcal{W}_0^{2,\infty}(\Omega)$ which have independent interest but are also utilised in the proofs of our main results in Section 3.

Lemma 9. *The functional space* (2-12) *is a Banach space when endowed with either of the Lipschitz-equivalent norms defined in* (2-13). *Moreover, for any* $p \in (1, \infty)$, *there exists a constant* $K = K(p) > 0$ (*depending also on* Ω, A) *such that*

$$\|u\|_{W^{2,p}(\Omega)} \le K \|A{:}D^2u\|_{L^p(\Omega)} \le K \|A{:}D^2u\|_{L^\infty(\Omega)}$$

for all $u \in W_0^{2,p}(\Omega)$.

Proof of Lemma 9. We begin by noting that the Lipschitz equivalence between the two norms in (2-13) is an immediate consequence of Morrey's embedding theorem, since $\|u\|_{L^\infty(\Omega)} \le K \|u\|_{W^{2,n}(\Omega)}$ for some $K = K(\Omega) > 0$. Next, by the L^p estimates for the Laplacian on a bounded open set $U \Subset \mathbb{R}^n$ (see, e.g., [Gilbarg and Trudinger 1998, Corollary 9.10]), for any $p \in (1, \infty)$, there exists a $K = K(p, U) > 0$ such that

$$\|v\|_{W^{2,p}(U)} \le K \|\Delta v\|_{L^p(U)}$$

for all $v \in W_0^{2,p}(U)$. Let $A = O^\top \Lambda^2 O$ be the spectral representation of A, namely $O \in O(n, \mathbb{R})$ and $\Lambda = \operatorname{diag}(\sigma(\sqrt{A}))$ is a diagonal matrix with entries the eigenvalues of $\sqrt{A}$. Then, by setting for brevity $\Gamma := \Lambda O$ and selecting $v(x) := u(\Gamma^\top x)$ and also $U := \Gamma^{-\top}\Omega$, we have $A = \Gamma^\top \Gamma$ and the identities $\Delta v(x) = A{:}D^2u(\Gamma^\top x)$ and $Dv(x) = \Gamma Du(\Gamma^\top x)$ for all $x \in \Omega$. By the change-of-variables formula and Hölder's inequality, there exists a new constant $K = K(p, \Omega, A) > 0$ such that

$$\|u\|_{W^{2,p}(\Omega)} \le K \|A{:}D^2u\|_{L^p(\Omega)} \le K \|u\|_{\mathcal{W}_0^{2,\infty}(\Omega)}$$

for all $u \in W_0^{2,p}(\Omega)$. If $(u_m)_1^\infty \subseteq \mathcal{W}_0^{2,\infty}(\Omega)$ is a Cauchy sequence, the above implies

$$\|u_m - u_n\|_{W^{2,p}(\Omega)} \le K \|A{:}D^2u_m - A{:}D^2u_n\|_{L^\infty(\Omega)} \to 0$$

as $m, n \to \infty$. By a diagonal argument and by the completeness of $L^\infty(\Omega)$ and of $\mathbb{W}_0^{2,\infty}(\Omega)$, there exist functions $u \in \mathbb{W}_0^{2,\infty}(\Omega)$ and $f \in L^\infty(\Omega)$ such that $u_m \to u$ in $W^{2,p}(\Omega)$ for any fixed $p \in (1, \infty)$, and

also $\mathrm{A}{:}\mathrm{D}^2 u_m \to f$ in $\mathrm{L}^\infty(\Omega)$, both as $m \to \infty$. Further, for any fixed $\phi \in \mathrm{C}_c^\infty(\Omega)$, we have

$$\int_\Omega \phi(\mathrm{A}{:}\mathrm{D}^2 u_m)\,\mathrm{d}\mathcal{L}^n = \int_\Omega u_m(\mathrm{A}{:}\mathrm{D}^2\phi)\,\mathrm{d}\mathcal{L}^n \to \int_\Omega u(\mathrm{A}{:}\mathrm{D}^2\phi)\,\mathrm{d}\mathcal{L}^n = \int_\Omega \phi(\mathrm{A}{:}\mathrm{D}^2 u)\,\mathrm{d}\mathcal{L}^n$$

as $m \to \infty$, which establishes that $\mathrm{A}:\mathrm{D}^2 u = f \in \mathrm{L}^\infty(\Omega)$. This in turn yields that $u \in \mathcal{W}_0^{2,\infty}(\Omega)$, as claimed. □

Remark 10. The space $\mathcal{W}_0^{2,\infty}(\Omega)$ is not closed in the relative Fréchet topology of $\mathbb{W}_0^{2,\infty}(\Omega)$. Indeed, fix any function $u \in \mathbb{W}_c^{2,\infty}(\Omega) \setminus \mathcal{W}^{2,\infty}(\Omega)$, and let $u^\varepsilon := u * \eta^\varepsilon \in \mathrm{C}_c^\infty(\Omega) \subseteq \mathcal{W}_0^{2,\infty}(\Omega)$ be the standard mollifier for $\varepsilon > 0$ small; see, e.g., [Katzourakis and Vărvărucă 2018, Chapter 9.6]. Then, $u^\varepsilon \to u$ in $\mathrm{W}_0^{2,p}(\Omega)$ as $\varepsilon \to 0$ for any fixed $p \in (1,\infty)$, but $u \notin \mathcal{W}_0^{2,\infty}(\Omega)$.

2.3.2. *Identification of the predual of $\mathcal{W}_0^{2,\infty}(\Omega)$ and weak* approximation.* The next important functional properties have their own independent interest but will also be utilised in Section 3.

Proposition 11. (i) *The space $\mathcal{W}_0^{2,\infty}(\Omega)$ is a dual Banach space, and in particular*

$$\mathcal{W}_0^{2,\infty}(\Omega) \cong \big(\mathbb{A}(\mathrm{L}^1(\Omega))\big)^*,$$

with the duality pairing given by

$$\langle \psi, \mathbb{A}u\rangle := \int_\Omega u(\mathrm{A}{:}\mathrm{D}^2\psi)\,\mathrm{d}\mathcal{L}^n, \quad \psi \in \mathcal{W}_0^{2,\infty}(\Omega), \ \ u \in \mathrm{L}^1(\Omega).$$

In the above, $\mathbb{A} : \mathrm{L}^1(\Omega) \to (\mathcal{W}_0^{2,\infty}(\Omega))^*$ denotes the bounded linear operator given by $u \mapsto \langle\,\cdot\,, \mathbb{A}u\rangle$, and we endow $\mathbb{A}(\mathrm{L}^1(\Omega))$ with the natural norm induced by inclusion into $(\mathcal{W}_0^{2,\infty}(\Omega))^*$, for which it becomes a separable Banach space, namely

$$\|\mathbb{A}u\| := \sup\left\{ \int_\Omega u(\mathrm{A}{:}\mathrm{D}^2\phi)\,\mathrm{d}\mathcal{L}^n \;\middle|\; \phi \in \mathcal{W}_0^{2,\infty}(\Omega),\ \|\phi\|_{\mathcal{W}_0^{2,\infty}(\Omega)} \le 1 \right\}.$$

(ii) *The space $\mathrm{C}_c^\infty(\Omega)$ is sequentially weak* dense in $\mathcal{W}_0^{2,\infty}(\Omega)$. Namely, for any $\psi \in \mathcal{W}_0^{2,\infty}(\Omega)$, there exists $(\psi_m)_{m=1}^\infty \subseteq \mathrm{C}_c^\infty(\Omega)$ such that, as $m \to \infty$, we have*

$$\mathrm{A}{:}\mathrm{D}^2\psi_m \overset{*}{\rightharpoonup} \mathrm{A}{:}\mathrm{D}^2\psi \quad \text{in } \mathrm{L}^\infty(\Omega), \tag{2-14}$$

$$\psi_m \rightharpoonup \psi \quad \text{in } \mathrm{W}_0^{2,p}(\Omega) \text{ for all } p \in (1,\infty). \tag{2-15}$$

Remark 12. Using similar techniques to those used to establish Proposition 11, one can also show that

$$(\mathcal{W}_0^{2,\infty}(\Omega))^* \cong \mathbb{A}(\mathrm{ba}(\Omega)),$$

where $\mathrm{ba}(\Omega) \cong (\mathrm{L}^\infty(\Omega))^* \cong (\mathrm{L}^1(\Omega))^{**}$ is the space of real-valued bounded finitely additive measures, which are absolutely continuous with respect to the Lebesgue measure on Ω; see, e.g., [Fonseca and Leoni 2007; Toland 2020]. However, the identification of the dual space of $\mathcal{W}_0^{2,\infty}(\Omega)$ is not required for the main results, therefore we refrain from this task.

The proof of Proposition 11 is based on the following general result of independent interest about subspaces of dual Banach spaces (Lemma 13). This is perhaps well known to the experts of duality theory, but since we could not locate a specific reference of this fact, we do provide a proof for the sake of completeness.

Lemma 13. *Let* $(\mathfrak{X}, \|\cdot\|)$ *be a Banach space, and let* $(\mathfrak{X}^*, \|\cdot\|_*)$ *be its dual Banach space. If* $\mathfrak{M} \subseteq \mathfrak{X}^*$ *is any weak*-closed subspace, then* $\mathfrak{M}$ *is a dual Banach space itself, and in particular*

$$(\mathfrak{X}/_{\perp\mathfrak{M}})^* \cong \mathfrak{M}.$$

The isometric isomorphism above is given by $\phi \mapsto \phi \circ \pi$, *where* $\pi : \mathfrak{X} \to \mathfrak{X}/_{\perp\mathfrak{M}}$ *is the quotient map, and* ${}^{\perp}\mathfrak{M}$ *denotes the pre-annihilator of* $\mathfrak{M}$ *in* $\mathfrak{X}$ *(Section 2.1).*

Proof of Lemma 13. We endow $\mathfrak{X}/_{\perp\mathfrak{M}}$ with the quotient norm and the spaces $\mathfrak{X}^*$ and $(\mathfrak{X}/_{\perp\mathfrak{M}})^*$ with their natural dual norms, and we define the mapping

$$\pi^* : (\mathfrak{X}/_{\perp\mathfrak{M}})^* \to \mathfrak{X}^*, \quad \pi^*(\phi) := \phi \circ \pi.$$

Then, π^* is linear and well defined, as for any $x \in \mathfrak{X}$ we have

$$|(\pi^*(\phi))(x)| = |\phi(\pi(x))| \le \|\phi\|_{(\mathfrak{X}/_{\perp\mathfrak{M}})^*} \|\pi(x)\|_{\mathfrak{X}/_{\perp\mathfrak{M}}} \le \|\phi\|_{(\mathfrak{X}/_{\perp\mathfrak{M}})^*} \|x\|.$$

Hence $\pi^*(\phi) \in \mathfrak{X}^*$. We will now show that

$$\pi(\mathbb{B}^{\mathfrak{X}}) = \mathbb{B}^{\mathfrak{X}/_{\perp\mathfrak{M}}}. \tag{2-16}$$

(This fact is claimed without proof in [Rudin 1991, p. 97].) On the one hand, the inclusion $\pi(\mathbb{B}^{\mathfrak{X}}) \subseteq \mathbb{B}^{\mathfrak{X}/_{\perp\mathfrak{M}}}$ is an immediate consequence of the inequality $\|\pi(x)\|_{\mathfrak{X}/_{\perp\mathfrak{M}}} \le \|x\|$, which is valid for all $x \in \mathfrak{X}$. Conversely, if we have any $\pi(x) \in \mathfrak{X}/_{\perp\mathfrak{M}}$ which satisfies $\|\pi(x)\|_{\mathfrak{X}/_{\perp\mathfrak{M}}} < 1$, for any $\varepsilon > 0$ small enough there exists $m_\varepsilon \in {}^{\perp}\mathfrak{M}$ such that $\|x - m_\varepsilon\| < \|\pi(x)\|_{\mathfrak{X}/_{\perp\mathfrak{M}}} + \varepsilon < 1$. Since $\pi(x - m_\varepsilon) = \pi(x)$, it follows that there exists a representative $x' := x - m_\varepsilon$ such that $\pi(x) = \pi(x') \in \pi(\mathbb{B}^{\mathfrak{X}})$ and $\|x'\| < 1$. This proves the reverse inclusion $\mathbb{B}^{\mathfrak{X}/_{\perp\mathfrak{M}}} \subseteq \pi(\mathbb{B}^{\mathfrak{X}})$, therefore establishing (2-16). We will now show that π^* is an isometry. In view of (2-16), we have

$$\|\pi^*(\phi)\|_* = \sup_{x \in \mathfrak{X}, \|x\| < 1} |\phi(\pi(x))| = \sup_{\pi(x) \in \mathfrak{X}/_{\perp M}, \|\pi(x)\|_{\mathfrak{X}/_{\perp\mathfrak{M}}} < 1} |\phi(\pi(x))| = \|\phi\|_{(\mathfrak{X}/_{\perp\mathfrak{M}})^*}.$$

We will now show that the image of π^* lies in $\mathfrak{M}$, namely

$$\pi^*((\mathfrak{X}/_{\perp\mathfrak{M}})^*) \subseteq \mathfrak{M}.$$

For any fixed $\phi \in (\mathfrak{X}/_{\perp\mathfrak{M}})^*$, we have that $(\pi^*(\phi))(x) = \phi(\pi(x)) = 0$ for all $x \in {}^{\perp}\mathfrak{M}$, because $\pi(x) = 0$ when $x \in {}^{\perp}\mathfrak{M}$. This implies that $\pi^*(\phi) \in ({}^{\perp}\mathfrak{M})^{\perp}$. Since $\mathfrak{M}$ is by assumption a weak*-closed subspace of $\mathfrak{X}^*$, by [Rudin 1991, Theorem 4.7, p. 96] we have that $({}^{\perp}\mathfrak{M})^{\perp} = \mathfrak{M}$. This implies the claimed inclusion. We finally show that π^* is onto $\mathfrak{M}$. With this aim, note that any functional $f \in \mathfrak{M}$ induces a functional $\bar{f} \in (\mathfrak{X}/_{\perp\mathfrak{M}})^*$ in a natural way by setting $\bar{f}(\pi(x)) := f(x),\ x \in \mathfrak{X}$. This is well defined regardless of the

choice of representative, because if we have $\pi(x'') = \pi(x')$, then $x'' - x' \in {}^{\perp}\mathfrak{M}$ and hence $f(x'') = f(x')$, because $f|_{{}^{\perp}\mathfrak{M}} \equiv 0$ when $f \in \mathfrak{M}$. Further, by the identity

$$|\bar{f}(\pi(x))| = |f(x)| = |f(x-z)| \quad \text{for all } z \in {}^{\perp}\mathfrak{M},$$

for any $x \in \mathfrak{X}$ we estimate

$$|\bar{f}(\pi(x))| \leq \|f\|_* \inf_{z \in {}^{\perp}\mathfrak{M}} \|x - z\| = \|f\|_* \|\pi(x)\|_{\mathfrak{X}/{}_{\perp\mathfrak{M}}}.$$

This shows that $\bar{f} \in (\mathfrak{X}/{}_{\perp\mathfrak{M}})^*$ and $\pi^*(\bar{f}) = f$, completing the proof. □

Remark 14. The assumption that the subspace $\mathfrak{M} \subseteq \mathfrak{X}^*$ be weak*-closed is essential, and in general it cannot be replaced by norm-closedness. For example, the space of sequences $c_0(\mathbb{N})$ vanishing at infinity is norm-closed (and weak*-dense) in $\ell^\infty(\mathbb{N})$, but as is well known, it possesses no predual space.

Having procured the necessary tools, we may now establish Proposition 11.

Proof of Proposition 11. (i) We begin by noting that $\mathrm{L}^1(\Omega)$ is a separable Banach space and $\mathbb{A}$ is a bounded linear operator, because for any $u \in \mathrm{L}^1(\Omega)$ we have

$$\|\mathbb{A}u\| = \sup_{\psi \in \mathcal{W}_0^{2,\infty}(\Omega), \|\psi\|_{\mathcal{W}_0^{2,\infty}(\Omega)} \leq 1} \int_\Omega u(\mathrm{A}{:}\mathrm{D}^2\psi)\, \mathrm{d}\mathcal{L}^n \leq \|u\|_{\mathrm{L}^1(\Omega)}.$$

Hence the image $\mathbb{A}(\mathrm{L}^1(\Omega))$ is a separable subset of $(\mathcal{W}_0^{2,\infty}(\Omega))^*$. Thus, $\mathbb{A}(\mathrm{L}^1(\Omega))$ becomes a separable normed space when endowed with the dual norm induced by $(\mathcal{W}_0^{2,\infty}(\Omega))^*$. Now we show that it is complete, and hence a Banach space, by showing that it is isometrically isomorphic to a quotient of $\mathrm{L}^1(\Omega)$ by a closed subspace. We define

$$\mathcal{W} := \{f \in \mathrm{L}^\infty(\Omega) \mid \exists\, \phi \in \mathcal{W}_0^{2,\infty}(\Omega) \text{ s.t. } f = \mathrm{A}{:}\mathrm{D}^2\phi \text{ a.e. on } \Omega\} = \mathbb{A}(\mathcal{W}_0^{2,\infty}(\Omega)).$$

Then, by arguing as in the proof of Lemma 9, we have that $\mathcal{W}$ is sequentially weak* closed in $\mathrm{L}^\infty(\Omega)$, and therefore it is also weak* closed because its predual $\mathrm{L}^1(\Omega)$ is separable and complete, as a consequence of the Krein–Smulian theorem; we refer, e.g., to [Conway 1990, Section 12, p. 159] or [Ostrovskii 2001]. By applying Lemma 13, it follows that $\pi^* : \mathcal{W} \to (\mathrm{L}^1(\Omega)/{}_{\perp\mathcal{W}})^*$ is an isometric isomorphism, and we also have that

$${}^{\perp}\mathcal{W} = \left\{ u \in \mathrm{L}^1(\Omega) \;\middle|\; \int_\Omega u(\mathrm{A}{:}\mathrm{D}^2\phi)\, \mathrm{d}\mathcal{L}^n = 0, \forall\, \phi \in \mathcal{W}_0^{2,\infty}(\Omega) \right\} = \mathrm{N}(\mathbb{A}).$$

Hence $\mathcal{W} \cong (\mathrm{L}^1(\Omega)/_{\mathrm{N}(\mathbb{A})})^*$. Next, note that the induced map on the quotient space

$$\mathbb{A} : \mathrm{L}^1(\Omega)/_{\mathrm{N}(\mathbb{A})} \to \mathbb{A}(\mathrm{L}^1(\Omega)), \quad \mathbb{A}(u + \mathrm{N}(\mathbb{A})) := \mathbb{A}u,$$

is a well-defined linear bijection, independent of the representative used in the coset. Further, for any fixed $u \in \mathrm{L}^1(\Omega)$, by standard duality results (see, e.g., [Katzourakis and Vărvărucă 2018, Corollary 11.46]) and by the expression of the quotient norm on $\mathrm{L}^1(\Omega)/_{\mathrm{N}(\mathbb{A})}$, it follows that

$$\|\mathbb{A}u\| = \sup_{\psi \in \mathcal{W}_0^{2,\infty}(\Omega), \|\mathrm{A}:\mathrm{D}^2\psi\|_{\mathrm{L}^\infty(\Omega)} \leq 1} \int_\Omega u(\mathrm{A}{:}\mathrm{D}^2\psi)\, \mathrm{d}\mathcal{L}^n = \sup_{f \in \mathcal{W}, \|f\|_{\mathrm{L}^\infty(\Omega)} \leq 1} \int_\Omega uf\, \mathrm{d}\mathcal{L}^n.$$

Hence

$$\|\mathbb{A}u\| = \sup_{f \in (L^1(\Omega)/N(\mathbb{A}))^*, \|f\|_{(L^1(\Omega)/N(\mathbb{A}))^*} \le 1} \int_\Omega uf \, d\mathcal{L}^n = \|u + N(\mathbb{A})\|_{L^1(\Omega)/N(\mathbb{A})}.$$

Thus, the mapping $\mathbb{A} : L^1(\Omega)/N(\mathbb{A}) \to \mathbb{A}(L^1(\Omega))$ is an isometric isomorphism, and therefore $\mathbb{A}(L^1(\Omega))$ is a Banach space. It follows that the induced dual map

$$\mathbb{A}^* : \big(\mathbb{A}(L^1(\Omega))\big)^* \to (L^1(\Omega)/N(\mathbb{A}))^*,$$

given by

$$\langle \mathbb{A}^*(\psi), u\rangle := \langle \psi, \mathbb{A}u\rangle, \quad u \in L^1(\Omega),$$

is also an isometric isomorphism between the corresponding dual spaces. Further, the linear map

$$\mathbb{A} : \mathcal{W}_0^{2,\infty}(\Omega) \to \mathcal{W}, \quad \phi \mapsto \mathbb{A}\phi,$$

is an isometric isomorphism between Banach spaces, since $\|\mathbb{A}\phi\|_{L^\infty(\Omega)} = \|\phi\|_{\mathcal{W}_0^{2,\infty}(\Omega)}$. In conclusion, we have the isometric isomorphism

$$(\mathbb{A}^*)^{-1} \circ \pi^* \circ \mathbb{A} : \mathcal{W}_0^{2,\infty}(\Omega) \to \big(\mathbb{A}(L^1(\Omega))\big)^*,$$

which, in view of the expression that the induced duality pairing takes, establishes the desired conclusion.

(ii) We will use the regularisation scheme introduced and studied in [Katzourakis 2022, Section 5], whose definition we recall for the convenience of the reader. Since $\partial\Omega$ is assumed to be Lipschitz, there exists a smooth vector field $\xi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$ which is transversal to the boundary $\partial\Omega$, namely $\xi \cdot n \ge \delta_0$ which holds $\mathcal{H}^{n-1}$-a.e. on $\partial\Omega$ for some $\delta_0 > 0$, where $n : \partial\Omega \to \mathbb{S}$ is the unit outer normal vector field, and additionally we have $|\xi| \equiv 1$ in an open collar around the boundary $\partial\Omega$. Further, there exist $\ell, \varepsilon_0 > 0$ such that, for all $\varepsilon \in (0, \varepsilon_0)$, we have $\operatorname{dist}(x + \varepsilon\ell\xi(x), \partial\Omega) \ge 2\varepsilon$ for $x \in \partial\Omega$. Then, for any $\psi \in \mathcal{W}_0^{2,\infty}(\Omega)$ extended by zero on $\mathbb{R}^n \setminus \Omega$, we set

$$K_\varepsilon\psi(x) := \int_{\mathbb{R}^n} \psi(x + \varepsilon\ell\xi(x) - \varepsilon y)\eta(y)\, dy,$$

where $\varepsilon \in (0, \varepsilon_0)$ and $\eta \in C_c^\infty(\mathbb{B})$ is a fixed mollifying kernel which satisfies $\eta \ge 0$ and $\|\eta\|_{L^1(\mathbb{R}^n)} = 1$. This is a variant of the standard mollifier, the difference being that it "compresses" ψ to a compactly supported function in Ω before mollifying. This method gives rise to compactly supported smooth approximations, which converge globally on Ω. In [Katzourakis 2022, Proposition 12] it is shown that $K_\varepsilon v \in C_c^\infty(\Omega)$ and that $K_\varepsilon\psi \to \psi$ in $W_0^{1,q}(\Omega)$ as $\varepsilon \to 0$, for any $q \in (1,\infty)$. By iterating this result, we also have that $K_\varepsilon\psi \to \psi$ in $W_0^{2,q}(\Omega)$ as $\varepsilon \to 0$, for any $q \in (1,\infty)$. Hence it remains to show that $\{A{:}D^2(K_\varepsilon\psi)\}_{\varepsilon\in(0,\varepsilon_0)}$ is bounded in $L^\infty(\Omega)$. With this aim, we compute

$$\begin{aligned} A{:}D^2(K_\varepsilon\psi)(x) = (I + \varepsilon\ell D\xi(x))^\top A(I + \varepsilon\ell D\xi(x)){:}\int_{\mathbb{R}^n} D^2\psi(x + \varepsilon\ell\xi(x) - \varepsilon y)\eta(y)\, dy, \\ + \varepsilon\ell(A{:}D^2\xi(x)) \cdot \int_{\mathbb{R}^n} D\psi(x + \varepsilon\ell\xi(x) - \varepsilon y)\eta(y)\, dy \end{aligned}$$

for any $x \in \Omega$. Hence

$$\begin{aligned}\mathrm{A}{:}\mathrm{D}^2(\mathrm{K}_\varepsilon\psi)(x) = \int_{\mathbb{R}^n} \mathrm{A}{:}\mathrm{D}^2\psi(x+\varepsilon\ell\xi(x)-\varepsilon y)\eta(y)\,\mathrm{d}y \\ + \varepsilon\ell(\mathrm{D}\xi(x)^\top\mathrm{A} + \mathrm{AD}\xi(x) + \varepsilon\ell\mathrm{D}\xi(x)^\top\mathrm{AD}\xi){:}\int_{\mathbb{R}^n}\mathrm{D}^2\psi(x+\varepsilon\ell\xi(x)-\varepsilon y)\eta(y)\,\mathrm{d}y \\ + \varepsilon\ell(\mathrm{A}{:}\mathrm{D}^2\xi(x))\cdot\int_{\mathbb{R}^n}\mathrm{D}\psi(x+\varepsilon\ell\xi(x)-\varepsilon y)\eta(y)\,\mathrm{d}y\end{aligned}$$

for any fixed $x \in \Omega$. Since $\eta \in \mathrm{C}^\infty_c(\mathbb{B})$, an integration by parts yields

$$\begin{aligned}\mathrm{A}{:}\mathrm{D}^2(\mathrm{K}_\varepsilon\psi)(x) = \int_{\mathbb{R}^n} \mathrm{A}{:}\mathrm{D}^2\psi(x+\varepsilon\ell\xi(x)-\varepsilon y)\eta(y)\,\mathrm{d}y \\ - \ell(\mathrm{D}\xi(x)^\top\mathrm{A} + \mathrm{AD}\xi(x) + \varepsilon\ell\mathrm{D}\xi(x)^\top A\mathrm{D}\xi){:}\int_{\mathbb{B}}\mathrm{D}\psi(x+\varepsilon\ell\xi(x)-\varepsilon y)\cdot\mathrm{D}\eta(y)\,\mathrm{d}y \\ + \varepsilon\ell(\mathrm{A}{:}\mathrm{D}^2\xi(x))\cdot\int_{\mathbb{R}^n}\mathrm{D}\psi(x+\varepsilon\ell\xi(x)-\varepsilon y)\eta(y)\,\mathrm{d}y\end{aligned}$$

for any $x \in \Omega$. Since η, $|\mathrm{D}\eta|$ are in $\mathrm{L}^1(\mathbb{B})$, by Young's inequality we get the estimate

$$\begin{aligned}|\mathrm{A}{:}\mathrm{D}^2(\mathrm{K}_\varepsilon\psi)(x)| \le \|\mathrm{A}{:}\mathrm{D}^2\psi\|_{\mathrm{L}^\infty(\Omega)} + \varepsilon_0\ell|\mathrm{A}|\|\mathrm{D}^2\xi\|_{\mathrm{L}^\infty(\mathbb{R}^n)}\|\mathrm{D}\psi\|_{\mathrm{L}^\infty(\Omega)} \\ + \ell|\mathrm{A}|\|\mathrm{D}\xi\|_{\mathrm{L}^\infty(\mathbb{R}^n)}(2+\varepsilon_0\ell\|\mathrm{D}\xi\|_{\mathrm{L}^\infty(\mathbb{R}^n)})\|\mathrm{D}\psi\|_{\mathrm{L}^\infty(\Omega)}\|\mathrm{D}\eta\|_{\mathrm{L}^1(\mathbb{R}^n)}\end{aligned}$$

for any $x \in \Omega$. This implies that $\mathrm{A}{:}\mathrm{D}^2(\mathrm{K}_\varepsilon\psi)$ is indeed bounded in $\mathrm{L}^\infty(\Omega)$ uniformly in $\varepsilon \in (0, \varepsilon_0)$. The conclusion follows. □

2.4. ***Compact sets in $\mathcal{W}^{2,\infty}_0(\Omega)$ and examples.*** By using the analogue of the Ascoli–Arzela theorem in $\mathrm{L}^\infty(\Omega)$ (see [Cherkas 1970]) and the results established in Section 2.3, we can explicitly characterise compact sets $\mathcal{K} \subseteq \mathcal{W}^{2,\infty}_0(\Omega)$ (arising in Theorem 2). To state this concisely, we need some additional notation. For any $f \in \mathrm{L}^\infty(\Omega)$, the *essential oscillation* of f over a Lebesgue measurable set $E \subseteq \Omega$ is defined as

$$\operatorname*{ess\,osc}_E[f] := \operatorname*{ess\,sup}_{x,y\in E}|f(x)-f(y)|.$$

Then, a closed subset $\mathcal{K} \subseteq \mathcal{W}^{2,\infty}_0(\Omega)$ is (strongly) compact if and only if

$$\inf_{\{\{\Omega_i\}_{i\in I}\text{ finite Borel partition of }\Omega\}}\sup_{f\in\mathcal{K}}\max_{i\in I}\operatorname*{ess\,osc}_{\Omega_i}[\mathrm{A}{:}\mathrm{D}^2 f] = 0. \tag{2-17}$$

The seemingly complicated condition (2-17) is equivalent to the following statement: for any $\varepsilon > 0$, there exists a finite partition of Ω into Borel sets $\{\Omega_i\}_{i\in I}$ such that

$$|\mathrm{A}{:}\mathrm{D}^2 f(x) - \mathrm{A}{:}\mathrm{D}^2 f(y)| \le \varepsilon \quad \text{a.e. } x, y \in \Omega_i$$

for all functions $f \in \mathcal{K}$ and all indices $i \in I$. By utilising the above, we have the following immediate consequence of Theorem 2 regarding specific interesting compact sets of variations.

Corollary 15 (special interesting classes of directions of test functions in $\mathcal{W}_0^{2,\infty}(\Omega)$). *In the setting of Theorem 2 and under the same assumptions, the following are true*:

(i) *For any $\theta > 0$ and $p > n$, there exists $\rho > 0$ such that*

$$F_\infty(u_\infty) < F_\infty(u_\infty + \psi)$$

among all $\psi \in (W^{3,p} \cap \mathcal{W}_0^{2,\infty})(\Omega) \setminus \{0\}$ satisfying

$$\|\psi\|_{\mathcal{W}_0^{2,\infty}(\Omega)} < \rho, \quad \|D^3\psi\|_{L^p(\Omega)} < \theta\rho.$$

(ii) *Let $\{\Omega_i : i \in I\}$ be a finite family of disjoint open subsets of Ω with Lipschitz boundaries $\partial\Omega_i$ whose closures cover Ω. Then, for any $\theta > 0$, there exists $\rho > 0$ such that*

$$F_\infty(u_\infty) < F_\infty(u_\infty + \psi)$$

among all $\psi \in \mathcal{W}_0^{2,\infty}(\Omega) \setminus \{0\}$ having piecewise Hölder continuous Hessians (namely $\psi|_{\Omega_i} \in C^{2,\alpha}(\overline{\Omega}_i)$ for all $i \in I$) satisfying

$$\|\psi\|_{\mathcal{W}_0^{2,\infty}(\Omega)} < \rho, \quad \max_{i\in I}[D^2\psi]_{C^\alpha(\overline{\Omega}_i)} < \theta\rho.$$

Proof of Corollary 15. (i) Given $\theta > 0$ and $p > n$, we define

$$\mathcal{K}' := \{\psi \in \mathbb{S}^{\mathcal{W}_0^{2,\infty}(\Omega)} \cap W^{3,p}(\Omega) : \|D^3\psi\|_{L^q(\Omega)} \le \theta\}.$$

In view of Theorem 2 (that will be established in Section 3), it suffices to show that $\mathcal{K}'$ is a (strongly) compact set in $\mathcal{W}_0^{2,\infty}(\Omega)$. Let $(\psi_m)_1^\infty$ be a sequence in $\mathcal{K}'$. Then, we have $\|A : D^2\psi_m\|_{L^\infty(\Omega)} = 1$ and $\|D^3\psi_m\|_{L^q(\Omega)} \le \theta$ for all $m \in \mathbb{N}$. By the compactness of the embedding $W^{3,p}(\Omega) \Subset W^{2,\infty}(\Omega)$, there exists $\psi \in W^{3,p}(\Omega)$ such that $\psi_{m_k} \to \psi$ in $W^{2,\infty}(\Omega)$ along a subsequence as $k \to \infty$. In particular, $\psi_{m_k} \to \psi$ in $\mathcal{W}_0^{2,\infty}(\Omega)$ as $k \to \infty$. Since $D^3\psi_{m_k} \rightharpoonup D^3\psi$ in $L^q(\Omega; \mathbb{R}^{n^3})$ as well, the weak lower semicontinuity of $\|\cdot\|_{L^q(\Omega)}$ implies $\psi \in \mathcal{K}'$.

(ii) Given $\theta > 0$ and finitely many disjoint open sets $\{\Omega_i : i \in I\}$ which have Lipschitz boundaries $\partial\Omega_i \in C^{0,1}$ and whose closures cover Ω, we define the set

$$\mathcal{K}'' := \mathbb{S}^{\mathcal{W}_0^{2,\infty}(\Omega)} \bigcap \left\{ \psi \in \bigoplus_{i\in I} C^{2,\alpha}(\overline{\Omega}_i) : \max_{i\in I}[D^2\psi]_{C^\alpha(\overline{\Omega}_i)} \le \theta \right\}.$$

It is then easy to see that $\mathcal{K}''$ is a (strongly) compact set in $\mathcal{W}_0^{2,\infty}(\Omega)$ (the proof follows similar lines to those of part (i), and therefore the details are omitted). □

2.5. L^p-limits of L^∞ constrained minimisers. The proofs of Theorems 4 and 7 are based on the following standalone result of independent interest, which establishes the existence and L^p-approximation of narrow (and global) minimisers of (1-1) through penalised functionals, treated in a unified fashion as constrained minimisers.

Theorem 16 (existence and L^p-approximation of L^∞ constrained minimisers). *Let F satisfy* (1-6)–(1-7) *and* (1-15), *and let $\bar{u} \in \mathcal{W}^{2,\infty}(\Omega)$ be fixed. Suppose that $\{\mathcal{E}^p_{\bar{u}}(\Omega)\}_{n\leq p\leq\infty}$ are classes of functions which satisfy*

$$\begin{cases} \mathcal{E}^p_{\bar{u}}(\Omega) \subseteq \mathrm{W}^{2,p}_{\bar{u}}(\Omega) & \text{when } n \leq p < \infty, \\ \mathcal{E}^p_{\bar{u}}(\Omega) \text{ is weakly closed} & \text{when } p < \infty, \\ \mathcal{E}^p_{\bar{u}}(\Omega) \subseteq \mathcal{E}^q_{\bar{u}}(\Omega) & \text{for } n \leq q \leq p \leq \infty, \\ \mathcal{E}^\infty_{\bar{u}}(\Omega) \subseteq \mathcal{W}^{2,\infty}_{\bar{u}}(\Omega) & \text{when } p = \infty, \\ \mathcal{E}^\infty_{\bar{u}}(\Omega) = \bigcap_{n<p<\infty} \mathcal{E}^p_{\bar{u}}(\Omega), & \\ \bar{u} \in \mathcal{E}^\infty_{\bar{u}}(\Omega). & \end{cases} \tag{2-18}$$

Then, the following are true:

(i) *For any $\epsilon \geq 0$ and $p \geq n + 1/c^3$, the functional $\mathrm{F}_{p,\epsilon,\bar{u}}$ (defined in* (1-19)) *has a minimiser u_p over the admissible class $\mathcal{E}^p_{\bar{u}}(\Omega)$.*

(ii) *For any $\epsilon \geq 0$, the functional $\mathrm{F}_\infty + \frac{\epsilon}{2}\|\cdot - \bar{u}\|^2_{\mathrm{L}^2(\Omega)}$ has a minimiser u_∞ over the admissible class $\mathcal{E}^\infty_{\bar{u}}(\Omega)$, where F_∞ is as in* (1-1). *Additionally, there exists a sequence of indices $(p_j)_1^\infty$ satisfying $p_j \to \infty$ as $j \to \infty$ along which we have*

$$\begin{cases} \mathrm{F}_{p,\epsilon,\bar{u}}(u_p) \to \mathrm{F}_\infty(u_\infty) + \frac{\epsilon}{2}\|u_\infty - \bar{u}\|^2_{\mathrm{L}^2(\Omega)}, \\ u_p \rightharpoonup u_\infty \quad \text{in } W^{2,q}(\Omega) \text{ for all } q \in (1,\infty), \\ u_p \to u_\infty \quad \text{in } \mathrm{C}^{1,\gamma}(\overline{\Omega}) \text{ for all } \gamma \in (0,1). \end{cases} \tag{2-19}$$

(iii) *If finally $\bar{u}$ minimises F_∞ over the admissible class $\mathcal{E}^\infty_{\bar{u}}(\Omega)$ and also $\epsilon > 0$, then we have that $\bar{u} = u_\infty$. In this case, we also have that the convergences in* (2-19) *are in fact full as $p \to \infty$ (not only along a subsequence).*

The main technical step for the proof of Theorem 16 is the following coercivity estimate for (1-19).

Lemma 17 (coercivity). *In the context of Theorem 16, let $\delta > 0$ and let $w \in \mathcal{E}^p_{\bar{u}}(\Omega)$ be a δ-approximate minimiser of $\mathrm{F}_{p,\epsilon,\bar{u}}$ in $\mathcal{E}^p_{\bar{u}}(\Omega)$, namely*

$$\mathrm{F}_{p,\epsilon,\bar{u}}(w) \leq \delta + \inf_{\mathcal{E}^p_{\bar{u}}(\Omega)} \mathrm{F}_{p,\epsilon,\bar{u}}.$$

Then, for any $q \in [n, p]$, we have the estimates

$$\begin{cases} \|w\|_{\mathrm{W}^{2,q}(\Omega)} \leq \dfrac{K(\Omega,q)}{c(1-\alpha)}(M + \delta + \mathrm{F}_\infty(\bar{u})), & \text{(i)} \\ \|\mathrm{A}{:}\mathrm{D}^2 w\|_{\mathrm{L}^q(\Omega)} \leq \dfrac{1}{c}(1 + \delta + \mathrm{F}_\infty(\bar{u}) + \|w\|^\alpha_{\mathrm{W}^{1,q}(\Omega)}), & \text{(ii)} \end{cases} \tag{2-20}$$

where $K(\Omega, q) > 0$ is the constant appearing in Lemma 9 and $M > 0$ is a constant depending only on $\bar{u}, \alpha, c, K(q, \Omega)$, A.

Proof of Lemma 17. Note first that by (1-1), (1-19), minimality and Young's inequality, we have

$$\inf_{\mathcal{E}^p_{\bar{u}}(\Omega)} \mathrm{F}_{p,\epsilon,\bar{u}} \leq \mathrm{F}_{p,\epsilon,\bar{u}}(\bar{u}) \leq \mathrm{F}_\infty(\bar{u}).$$

Further, by (1-7) we have that $F(x,\eta,\mathrm{p},\cdot)^{-1}(\{0\})$ is a singleton set. Hence, if we write

$$F(x,\eta,\mathrm{p},\cdot)^{-1}(\{0\}) = \{\bar{\xi}(x,\eta,\mathrm{p})\}, \tag{2-21}$$

again by (1-7), we can easily deduce the inequality

$$|F(x,\eta,\mathrm{p},\xi)| \geq c|\xi - \bar{\xi}(x,\eta,\mathrm{p})|. \tag{2-22}$$

By Lemma 9 and (1-15), since $w-\bar{u} \in \mathrm{W}_0^{2,q}(\Omega)$, we can estimate

$$\begin{aligned}
\mathrm{F}_{p,\epsilon,\bar{u}}(w) &\geq \|F(\mathrm{J}^2 w)\|_{\mathrm{L}^q(\Omega)} \\
&\geq c\|\mathrm{A}{:}\mathrm{D}^2 w - \bar{\xi}(\,\cdot\,, w, \mathrm{D}w)\|_{\mathrm{L}^q(\Omega)} \\
&\geq c\|\mathrm{A}{:}\mathrm{D}^2(w-\bar{u})\|_{\mathrm{L}^q(\Omega)} - c\|\mathrm{A}{:}\mathrm{D}^2\bar{u}\|_{\mathrm{L}^q(\Omega)} - c\|\bar{\xi}(\,\cdot\,, w, \mathrm{D}w)\|_{\mathrm{L}^q(\Omega)} \\
&\geq c\frac{\|w-\bar{u}\|_{\mathrm{W}^{2,q}(\Omega)}}{K(q,\Omega)} - c\|\mathrm{A}{:}\mathrm{D}^2\bar{u}\|_{\mathrm{L}^\infty(\Omega)} - \|(|w|+|\mathrm{D}w|)^\alpha\|_{\mathrm{L}^q(\Omega)} - 1 \\
&\geq c\frac{\|w-\bar{u}\|_{\mathrm{W}^{2,q}(\Omega)}}{K(q,\Omega)} - c\|\mathrm{A}{:}\mathrm{D}^2\bar{u}\|_{\mathrm{L}^\infty(\Omega)} - \|w\|^\alpha_{\mathrm{W}^{1,\alpha q}(\Omega)} - 1 \\
&\geq c\frac{\|w-\bar{u}\|_{\mathrm{W}^{2,q}(\Omega)}}{K(q,\Omega)} - c\|\mathrm{A}{:}\mathrm{D}^2\bar{u}\|_{\mathrm{L}^\infty(\Omega)} - \|w\|^\alpha_{\mathrm{W}^{1,q}(\Omega)} - 1.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{F}_{p,\epsilon,\bar{u}}(w) &\geq c\frac{\|w-\bar{u}\|_{\mathrm{W}^{2,q}(\Omega)}}{K(q,\Omega)} - c\|\mathrm{A}{:}\mathrm{D}^2\bar{u}\|_{\mathrm{L}^\infty(\Omega)} - \|w-\bar{u}\|^\alpha_{\mathrm{W}^{1,q}(\Omega)} - \|\bar{u}\|^\alpha_{\mathrm{W}^{1,\infty}(\Omega)} - 1 \\
&\geq c\frac{\|w-\bar{u}\|_{\mathrm{W}^{2,q}(\Omega)}}{K(q,\Omega)} - c\|\mathrm{A}{:}\mathrm{D}^2\bar{u}\|_{\mathrm{L}^\infty(\Omega)} - \|w-\bar{u}\|^\alpha_{\mathrm{W}^{1,q}(\Omega)} - \|\bar{u}\|^\alpha_{\mathrm{W}^{1,\infty}(\Omega)} - 1 \\
&\geq \frac{c}{K(q,\Omega)}\|w-\bar{u}\|_{\mathrm{W}^{2,q}(\Omega)} - (\|w-\bar{u}\|_{\mathrm{W}^{2,q}(\Omega)})^\alpha - (c\|\mathrm{A}{:}\mathrm{D}^2\bar{u}\|_{\mathrm{L}^\infty(\Omega)} + \|\bar{u}\|^\alpha_{\mathrm{W}^{1,\infty}(\Omega)} + 1).
\end{aligned}$$

By using the elementary inequality

$$Nt - t^\alpha \geq N(1-\alpha)[t - N^{1/(\alpha-1)}], \quad t \geq 0, \quad N > 0,$$

the previous estimate yields

$$\mathrm{F}_{p,\epsilon,\bar{u}}(w) \geq \frac{c(1-\alpha)}{K(q,\Omega)}\left[\|w-\bar{u}\|_{\mathrm{W}^{2,q}(\Omega)} - \left(\frac{c}{K(q,\Omega)}\right)^{1/(\alpha-1)}\right] - (c\|\mathrm{A}{:}\mathrm{D}^2\bar{u}\|_{\mathrm{L}^\infty(\Omega)} + \|\bar{u}\|^\alpha_{\mathrm{W}^{1,\infty}(\Omega)} + 1).$$

This implies that

$$\begin{aligned}
\mathrm{F}_{p,\epsilon,\bar{u}}(w) &\geq \frac{c(1-\alpha)}{K(q,\Omega)}\|w\|_{\mathrm{W}^{2,q}(\Omega)} - \left\{\frac{c(1-\alpha)}{K(q,\Omega)}\left[\|\bar{u}\|_{\mathrm{W}^{2,q}(\Omega)} + \left(\frac{c}{K(q,\Omega)}\right)^{1/(\alpha-1)}\right]\right. \\
&\qquad \left. + c\|\mathrm{A}{:}\mathrm{D}^2\bar{u}\|_{\mathrm{L}^\infty(\Omega)} + \|\bar{u}\|^\alpha_{\mathrm{W}^{1,\infty}(\Omega)} + 1\right\} \\
&=: \frac{c(1-\alpha)}{K(q,\Omega)}\|w\|_{\mathrm{W}^{2,q}(\Omega)} - M,
\end{aligned}$$

from which (2-20)(i) follows. Similarly, to obtain (2-20)(ii), we have

$$\begin{aligned}\mathrm{F}_{p,\epsilon,\bar{u}}(w) \geq \|F(\mathrm{J}^2 w)\|_{\mathrm{L}^q(\Omega)} &\geq c\|\mathrm{A}{:}\mathrm{D}^2 w - \bar{\xi}(\,\cdot\,, w, \mathrm{D}w)\|_{\mathrm{L}^q(\Omega)} \\ &\geq c\|\mathrm{A}{:}\mathrm{D}^2 w\|_{\mathrm{L}^q(\Omega)} - \|(|w| + |\mathrm{D}w|)^\alpha\|_{\mathrm{W}^{1,q}(\Omega)} - 1 \\ &\geq c\|\mathrm{A}{:}\mathrm{D}^2 w\|_{\mathrm{L}^q(\Omega)} - \|w\|^\alpha_{\mathrm{W}^{1,q}(\Omega)} - 1.\end{aligned}$$

The lemma has been established. □

Now we may continue with the proof the theorem.

Proof of Theorem 16. (i) This is a simple application of the direct method. Fix an exponent $p > n + 1/c^3$, and let $(w_j)_1^\infty \subseteq \mathcal{E}^p_{\bar{u}}(\Omega) \subseteq W^{2,p}_{\bar{u}}(\Omega)$ be a minimising sequence of $\mathrm{F}_{p,\epsilon,\bar{u}}$ in $\mathcal{E}^p_{\bar{u}}(\Omega)$. By Lemma 17, $(w_j)_1^\infty$ is bounded in $\mathrm{W}^{2,p}_{\bar{u}}(\Omega)$, hence by the sequential weak compactness of bounded sets in $\mathrm{W}^{2,p}(\Omega)$ there exists $u_p \in \mathcal{E}^p_{\bar{u}}(\Omega)$ and a subsequence of indices $(j_k)_1^\infty$ such that $w_{j_k} \rightharpoonup u_p$ in $\mathrm{W}^{2,p}_{\bar{u}}(\Omega)$ as $k \to \infty$. By assumption (2-18), $\mathcal{E}^p_{\bar{u}}(\Omega)$ is weakly closed, therefore $u_p \in \mathcal{E}^p_{\bar{u}}(\Omega)$. Since we have assumed $p > n + 1/c^3$, by (1-15) and (2-2) tested against rank-1 matrices of the form $(0, 0, 0, \xi) \otimes (0, 0, 0, \xi)$, it follows that $\partial^2_{\xi\xi}(|F(x, \eta, \mathrm{p}, \cdot)|^p) \geq 0$, which implies that $|F(x, \eta, \mathrm{p}, \cdot)|^p$ is convex. By (1-7) and (1-15), we have the lower bound

$$F(x, \eta, \mathrm{p}, \xi) \geq c(\xi - \bar{\xi}(x, \eta, \mathrm{p})) \geq c\xi - (|\eta| + |\mathrm{p}|)^\alpha - 1,$$

and hence the functional (1-19) is sequentially weakly lower semicontinuous (see, e.g., [Dacorogna 2008, Theorem 3.23, p. 96]) on the weak-closed closed set $\mathcal{E}^p_{\bar{u}}(\Omega)$. Hence u_p minimises the functional $\mathrm{F}_{p,\epsilon,\bar{u}}$ over $\mathcal{E}^p_{\bar{u}}(\Omega)$.

(ii) Let $\{u_p\}_{n<p<\infty} \subseteq \mathrm{W}^{2,n}_{\bar{u}}(\Omega)$ be the family of minimisers from part (i). By applying Lemma 17, for each fixed $q \in (1, \infty)$, $\{u_p\}_{n<p<\infty}$ is bounded in $\mathrm{W}^{2,q}(\Omega)$ as a result of (2-20)(i). By sequential weak compactness and a standard diagonal argument, there exists a sequence of indices $(p_j)_1^\infty$ and $u_\infty \in \mathbb{W}^{2,\infty}(\Omega)$ (recall (2-9)) such that $u_{p_j} \rightharpoonup u_\infty$ in $\mathrm{W}^{2,q}(\Omega)$ as $j \to \infty$. By Morrey's theorem and the Lipschitz regularity of $\partial\Omega$, we also have that $u_{p_j} \to u_\infty$ in $\mathrm{C}^{1,\gamma}(\overline{\Omega})$ as $j \to \infty$. Again by Lemma 17, for any $q \in (1, \infty)$ fixed and $p \geq q$, by (2-20)(ii) we have that

$$\|\mathrm{A}{:}\mathrm{D}^2 u_p\|_{\mathrm{L}^q(\Omega)} \leq \frac{1 + \delta + \mathrm{F}_\infty(\bar{u}) + \|u_p\|^\alpha_{\mathrm{W}^{1,\infty}(\Omega)}}{c}.$$

By the sequential weak lower semicontinuity of the L^q norm and the strong convergence of $u_{p_j} \to u_\infty$ in $\mathrm{C}^{1,\gamma}(\overline{\Omega})$ as $j \to \infty$, we have

$$\|\mathrm{A}{:}\mathrm{D}^2 u_\infty\|_{\mathrm{L}^q(\Omega)} \leq \liminf_{j\to\infty} \|\mathrm{A}{:}\mathrm{D}^2 u_{p_j}\|_{\mathrm{L}^q(\Omega)} \leq \frac{1 + \delta + \mathrm{F}_\infty(\bar{u}) + \|u_\infty\|^\alpha_{\mathrm{W}^{1,\infty}(\Omega)}}{c}.$$

Then, by letting $q \to \infty$ we infer that $\mathrm{A}{:}\mathrm{D}^2 u_\infty \in \mathrm{L}^\infty(\Omega)$, and therefore $u_\infty \in \mathcal{W}^{2,\infty}_{\bar{u}}(\Omega)$. Additionally, by (2-18), for any fixed $q \in (1, \infty)$, the family $(u_p)_{p\geq q}$ is contained in $\mathcal{E}^q_{\bar{u}}(\Omega)$. Since $\mathcal{E}^q_{\bar{u}}(\Omega)$ is weakly closed, by letting $p_j \to \infty$ we infer that $u_\infty \in \mathcal{E}^q_{\bar{u}}(\Omega)$ for all $q \in (1, \infty)$. By invoking (2-18) once again, we deduce that $u_\infty \in \mathcal{E}^\infty_{\bar{u}}(\Omega)$. Now we show that u_∞ minimises $\mathrm{F}_\infty + \frac{\epsilon}{2}\|\,\cdot - \bar{u}\|^2_{\mathrm{L}^2(\Omega)}$ and also show the

convergence of the energies. For any $v \in \mathcal{E}^\infty_{\bar u}(\Omega)$ fixed and $p \geq q$, we have

$$\begin{aligned}\|F(\mathrm{J}^2 u_p)\|_{\mathrm{L}^q(\Omega)} + \frac{\epsilon}{2}\|u_p - \bar u\|^2_{\mathrm{L}^2(\Omega)} &\leq \|F(\mathrm{J}^2 u_p)\|_{\mathrm{L}^p(\Omega)} + \frac{\epsilon}{2}\|u_p - \bar u\|^2_{\mathrm{L}^2(\Omega)} \\ &= \mathrm{F}_{p,\epsilon,\bar u}(u_p) \leq \mathrm{F}_{p,\epsilon,\bar u}(v) \leq \mathrm{F}_\infty(v) + \frac{\epsilon}{2}\|v - \bar u\|^2_{\mathrm{L}^2(\Omega)}.\end{aligned}$$

By the sequential weak lower semicontinuity of the functional, we deduce

$$\begin{aligned}\|F(\mathrm{J}^2 u_\infty)\|_{\mathrm{L}^q(\Omega)} + \frac{\epsilon}{2}\|u_\infty - \bar u\|^2_{\mathrm{L}^2(\Omega)} &\leq \liminf_{p_j \to \infty} \mathrm{F}_{p,\epsilon,\bar u}(u_p) \\ &\leq \limsup_{p_j \to \infty} \mathrm{F}_{p,\epsilon,\bar u}(u_p) \leq \mathrm{F}_\infty(v) + \frac{\epsilon}{2}\|v - \bar u\|^2_{\mathrm{L}^2(\Omega)} \qquad (2\text{-}23)\end{aligned}$$

for any $v \in \mathcal{E}^\infty_{\bar u}(\Omega)$. By letting $q \to \infty$ in inequality (2-23), the choice $v := u_\infty$ completes the proof of (2-19).

(iii) If we have $\epsilon > 0$ and $\bar u$ minimises F_∞ over $\mathcal{E}^\infty_{\bar u}(\Omega)$, choosing $v := \bar u$ in inequality (2-23) and letting $q \to \infty$ yields $\bar u = u_\infty$. □

2.6. ***A lemma of strong compactness in*** $\mathbf{L^1}$**.** The following simple result — which is an extension of a lemma proved in the appendix of [Katzourakis and Parini 2017] — will be used in an essential fashion in Section 3. (Let us recall here our notation τ_z for the translation operator.)

Lemma 18. *Let* $\mathrm{A} \in \mathbb{R}^{n\times n}$ *be a* (*strictly*) *positive matrix,* $\Omega \Subset \mathbb{R}^n$, $\sigma \in \mathrm{L}^1(\Omega)$ *and* $\Sigma \in \mathrm{L}^1(\Omega;\mathbb{R}^n)$. *Suppose* $f \in \mathrm{L}^1(\Omega)$ *solves the PDE*

$$\mathrm{div}(\mathrm{A}\mathrm{D}f - \Sigma) + \sigma = 0 \quad \text{in } (\mathrm{C}^\infty_c(\Omega))^*,$$

namely in the sense of distributions. Then, for any $\Omega' \Subset \Omega$, *there exists a modulus of continuity* $\omega \in \mathrm{C}^\uparrow_{0^+}[0,\infty)$ (*depending only on* $n, \Omega', \Omega, \mathrm{A}$) *such that*

$$\|\tau_z f - f\|_{\mathrm{L}^1(\Omega')} \leq \omega(|z|)(\|f\|_{\mathrm{L}^1(\Omega)} + \|\sigma\|_{\mathrm{L}^1(\Omega)} + \|\Sigma\|_{\mathrm{L}^1(\Omega)})$$

for all $z \in \mathbb{R}^n$ *with* $|z| < \mathrm{dist}(\Omega', \partial\Omega)$. *Further, for any* $q \in (1, n')$, *where* $n' = \frac{n}{n-1}$, *there exists a constant* $K = K(q, n, \Omega', \Omega, A) > 0$ *such that*

$$\|f\|_{\mathrm{L}^q(\Omega')} \leq K(\|f\|_{\mathrm{L}^1(\Omega)} + \|\sigma\|_{\mathrm{L}^1(\Omega)} + \|\Sigma\|_{\mathrm{L}^1(\Omega)}).$$

As a consequence, any set of distributional solutions which is bounded in $\mathrm{L}^1(\Omega)$ *is also strongly precompact in* $\mathrm{L}^q_{\mathrm{loc}}(\Omega)$ *for all* $q \in (1, n')$.

Proof of Lemma 18. Without loss of generality, by a change-of-variables argument as in Lemma 9, we may assume that A is the identity matrix, and therefore we have the Poisson equation $\Delta f = \mathrm{div}(\Sigma) - \sigma$ in Ω. Fix $\Omega' \Subset \Omega$, and fix also Ω'', Ω''' such that $\Omega' \Subset \Omega'' \Subset \Omega''' \Subset \Omega$. By the interior nature of the desired estimate and the linearity of the PDE, without loss of generality, a mollification argument allows us to assume that $u, \sigma \in \mathrm{C}^\infty(\overline{\Omega'''})$, $\Sigma \in \mathrm{C}^\infty(\overline{\Omega'''};\mathbb{R}^n)$, $\partial\Omega'''$ is piecewise C^∞, and the PDE holds classically in Ω'''. Let $\zeta \in \mathrm{C}^\infty_c(\mathbb{R}^n)$ be a cut-off function which satisfies $\chi_{\Omega''} \leq \zeta \leq \chi_{\Omega'''}$. By setting $\bar\sigma := \zeta\sigma$ and $\overline{\Sigma} := \zeta\Sigma$ the PDE implies that $\Delta f = \mathrm{div}(\overline{\Sigma}) - \bar\sigma$ in Ω'', and we have that $\bar\sigma \in \mathrm{C}^\infty_c(\Omega)$, $\overline{\Sigma} \in \mathrm{C}^\infty_c(\Omega;\mathbb{R}^n)$, and also $\bar\sigma|_{\Omega''} \equiv \sigma$ and $\overline{\Sigma}|_{\Omega''} \equiv \Sigma$. Let now $R = R_\Omega > 0$ be the radius of the smallest open ball such that $\Omega - \Omega + \mathbb{B}_1 \subseteq \mathbb{B}_R$. This means that $x - y + z \in \mathbb{B}_R$ for all $x, y \in \Omega$ and all $z \in \mathbb{B}_1$. Let $\eta \in \mathrm{C}^\infty_c(\mathbb{R}^n)$ be

a cut-off function such that $\eta \equiv 1$ on $\mathbb{B}_R$. By Green's formula (see, e.g., [Gilbarg and Trudinger 1998, Chapter 2]), we can represent the solution f on Ω'' as

$$f = h - \Phi * \bar{\sigma} + \Phi * \mathrm{div}(\overline{\Sigma})$$

for some harmonic function h on Ω'', where Φ is the fundamental solution of the Laplace equation. Then, for any fixed $z \in \mathbb{R}^n$ with $|z| < \mathrm{dist}(\Omega', \partial\Omega'')$, we have the estimate

$$\|\tau_z f - f\|_{\mathrm{L}^1(\Omega')} \leq \big\|\tau_z(\Phi * \mathrm{div}(\overline{\Sigma})) - \Phi * \mathrm{div}(\overline{\Sigma})\big\|_{\mathrm{L}^1(\Omega')} + \|\tau_z(\Phi * \bar{\sigma}) - \Phi * \bar{\sigma}\|_{\mathrm{L}^1(\Omega')} + \|\tau_z h - h\|_{\mathrm{L}^1(\Omega')}.$$

We will obtain the desired estimate by estimating each of the three terms in the above inequality separately. By the identities

$$\tau_z(\Phi * \mathrm{div}(\overline{\Sigma})) = \tau_z(\mathrm{D}\Phi * \overline{\Sigma}) = (\tau_z \mathrm{D}\Phi) * \overline{\Sigma},$$

(where we interpret the convoluted inner product of vector functions in the above identities as $\mathrm{D}\Phi * \overline{\Sigma} = \sum_{i=1}^n \mathrm{D}_i\Phi * \overline{\Sigma}_i$), we estimate by using Young's inequality

$$\begin{aligned}
\big\|\tau_z(\Phi * \mathrm{div}(\overline{\Sigma})) - \Phi * \mathrm{div}(\overline{\Sigma})\big\|_{\mathrm{L}^1(\Omega')} &= \big\|(\tau_z(\mathrm{D}\Phi) - \mathrm{D}\Phi) * \overline{\Sigma}\big\|_{\mathrm{L}^1(\Omega')} \\
&= \big\|\big[(\tau_z(\eta\mathrm{D}\Phi) - \eta\mathrm{D}\Phi)\chi_{\mathbb{B}_R}\big] * \overline{\Sigma}\big\|_{\mathrm{L}^1(\mathbb{R}^n)} \\
&\leq \|\tau_z(\eta\mathrm{D}\Phi) - \eta\mathrm{D}\Phi\|_{\mathrm{L}^1(\mathbb{B}_R)}\|\overline{\Sigma}\|_{\mathrm{L}^1(\mathbb{R}^n)} \\
&\leq \|\tau_z(\eta\mathrm{D}\Phi) - \eta\mathrm{D}\Phi\|_{\mathrm{L}^1(\mathbb{R}^n)}\|\Sigma\|_{\mathrm{L}^1(\Omega)}.
\end{aligned}$$

Similarly, we have the estimate

$$\|\tau_z(\Phi * \bar{\sigma}) - \Phi * \bar{\sigma}\|_{\mathrm{L}^1(\Omega')} \leq \|\tau_z(\eta\Phi) - \eta\Phi\|_{\mathrm{L}^1(\mathbb{R}^n)}\|\sigma\|_{\mathrm{L}^1(\Omega)}.$$

Further, by the mean value theorem for harmonic functions and interior derivative estimates, we have

$$\|\tau_z h - h\|_{\mathrm{L}^1(\Omega')} \leq |z|\|\mathrm{D}h\|_{\mathrm{L}^\infty(\Omega'')} \leq |z|\|h\|_{\mathrm{L}^1(\Omega)}.$$

Further, by using that $h = f + \Phi * \bar{\sigma} - \mathrm{D}\Phi * \overline{\Sigma}$, Young's inequality implies

$$\begin{aligned}
\|h\|_{\mathrm{L}^1(\Omega)} &\leq \|f\|_{\mathrm{L}^1(\Omega)} + \|\Phi * \bar{\sigma}\|_{\mathrm{L}^1(\Omega)} + \|\mathrm{D}\Phi * \overline{\Sigma}\|_{\mathrm{L}^1(\Omega)} \\
&= \|f\|_{\mathrm{L}^1(\Omega)} + \|(\Phi\chi_{\mathbb{B}_R}) * \bar{\sigma}\|_{\mathrm{L}^1(\Omega)} + \|(\mathrm{D}\Phi\chi_{\mathbb{B}_R}) * \overline{\Sigma}\|_{\mathrm{L}^1(\Omega)} \\
&\leq \|f\|_{\mathrm{L}^1(\Omega)} + \|\Phi\|_{\mathrm{L}^1(\mathbb{B}_R)}\|\sigma\|_{\mathrm{L}^1(\Omega)} + \|\mathrm{D}\Phi\|_{\mathrm{L}^1(\mathbb{B}_R)}\|\Sigma\|_{\mathrm{L}^1(\Omega)}.
\end{aligned}$$

Therefore, we have

$$\|\tau_z h - h\|_{\mathrm{L}^1(\Omega')} \leq |z|(1 + \|\Phi\|_{\mathrm{W}^{1,1}(\mathbb{B}_R)})(\|f\|_{\mathrm{L}^1(\Omega)} + \|\sigma\|_{\mathrm{L}^1(\Omega)} + \|\Sigma\|_{\mathrm{L}^1(\Omega)}).$$

Finally, we conclude by putting the pieces together. Since $\eta\Phi \in \mathrm{L}^1(\mathbb{R}^n)$ and also $\eta\mathrm{D}\Phi \in \mathrm{L}^1(\mathbb{R}^n; \mathbb{R}^n)$, by the continuity of the translation operation, there exists $\omega \in C_0^{\uparrow}[0, \infty)$ such that

$$\|\tau_z(\eta\Phi) - \eta\Phi\|_{\mathrm{L}^1(\mathbb{R}^n)} + \|\tau_z(\eta\mathrm{D}\Phi) - \eta\mathrm{D}\Phi\|_{\mathrm{L}^1(\mathbb{R}^n)} \leq \omega(|z|).$$

By the above estimates, the desired L^1-equicontinuity estimate ensues. Now we establish the $\mathrm{L}^q_{\mathrm{loc}}$ bound. Fix $q \in (1, n')$, and fix also $\zeta \in \mathrm{C}_c^\infty(\Omega)$ with $\chi_{\Omega'} \leq \zeta \leq \chi_{\Omega'''}$. Let $\phi \in \mathrm{C}^\infty(\Omega)$ be arbitrary, and test the

PDE against $\zeta\phi \in C_c^\infty(\Omega)$ to obtain

$$\int_\Omega \zeta f\,(A{:}D^2\phi)\,d\mathcal{L}^n = -\int_{\Omega'''} f\,A{:}(D\zeta\otimes D\phi + D\phi\otimes D\zeta + \phi D^2\zeta)\,d\mathcal{L}^n \\ - \int_{\Omega'''} \Sigma\cdot(\phi D\zeta + \zeta D\phi)\,d\mathcal{L}^n - \int_{\Omega'''}\sigma\zeta\phi\,d\mathcal{L}^n.$$

This identity implies the estimate

$$\left|\int_\Omega \zeta f\,(A{:}D^2\phi)\,d\mathcal{L}^n\right| \le C(n)(|A|+1)\|\zeta\|_{W^{2,\infty}(\Omega)}(\|f\|_{L^1(\Omega)} + \|\sigma\|_{L^1(\Omega)} + \|\Sigma\|_{L^1(\Omega)})\|\phi\|_{W^{1,\infty}(\Omega''')}$$

for some universal $C = C(n) > 0$. Fix now any $\psi \in C_c^\infty(\Omega)$. By standard L^p theory for elliptic equations and bootstrap regularity (see, e.g., [Gilbarg and Trudinger 1998, Chapters 6 and 9]), there exists a unique $\phi \in C^\infty(\Omega)\cap C_0^1(\Omega)$ such that $A{:}D^2\phi = \psi$ on Ω and also

$$\|\phi\|_{W^{1,\infty}(\Omega''')} \le C\|A{:}D^2\phi\|_{L^{q'}(\Omega''')}$$

for some $C = C(q, n, \Omega) > 0$, which is possible by Morrey's embedding theorem because $q' > n$ (since by assumption $q \in (1, n')$), and also $\Omega''' \Subset \Omega$ has piecewise smooth boundary. Therefore, there exists a $K = K(q, n, \Omega', \Omega, A) > 0$ such that

$$\left|\int_\Omega (\zeta f)\psi\,d\mathcal{L}^n\right| \le K(\|f\|_{L^1(\Omega)} + \|\sigma\|_{L^1(\Omega)} + \|\Sigma\|_{L^1(\Omega)})\|\psi\|_{L^{q'}(\Omega)}$$

for any $\psi \in C_c^\infty(\Omega)$. By standard L^p theory (see, e.g., [Katzourakis and Vărvărucă 2018, Chapter 14]), this implies

$$\|f\|_{L^q(\Omega')} \le \|\zeta f\|_{L^q(\Omega)} \le K(\|f\|_{L^1(\Omega)} + \|\sigma\|_{L^1(\Omega)} + \|\Sigma\|_{L^1(\Omega)}),$$

as claimed. The last claim of the statement regarding strong compactness follows from the Fréchet–Kolmogorov and the Vitali convergence theorems. □

2.7. ***Nodal sets and nontrivial solutions.*** The following result lists some properties of solutions to equations of the form (1-9) but with general L^∞ coefficients. It holds true much more generally, but we specialise the statement only to the case of interest in this paper.

Proposition 19. *Let* $A \in \mathbb{R}^{n\times n}$ *be a* (*strictly*) *positive matrix*, $\Omega \Subset \mathbb{R}^n$, $K \in L^\infty(\Omega)$ *and* $L \in L^\infty(\Omega;\mathbb{R}^n)$. *Suppose that* $f \in L^1(\Omega)$ *solves the PDE*

$$\operatorname{div}(AD f - Lf) + Kf = 0 \quad \text{in } \Omega \tag{2-24}$$

in the sense of distributions, *namely in* $(C_c^\infty(\Omega))^*$. *Then*:

(i) *Any solution* $f \in L^1(\Omega)$ *to* (2-24) *satisfies* $f \in W^{1,p}_{\mathrm{loc}}(\Omega)$ *for all* $p \in (1,\infty)$, *and thus* (*the precise representative of*) f *is in* $C^{0,\alpha}(\Omega)$ *for all* $\alpha \in (0,1)$.

(ii) *There exists a nontrivial solution* $f^* \not\equiv 0$ *to* (2-24).

(iii) *For any nonconstant solution* f *to* (2-24), *the nodal set* $\{f = 0\}$ *is closed and countably* $(n-1)$*-rectifiable. Thus*, *the* $\mathcal{H}^{n-1}$*-Hausdorff measure of* $\{f = 0\}$ *is locally finite*, *and in particular it is an* $\mathcal{L}^n$*-nullset in* Ω.

Proof of Proposition 19. (i) Since the coefficients K, L of (2-24) are essentially bounded and A is a constant positive matrix, by the result [Bogachev et al. 2009, Corollary 1.2.8, p. 10] of Bogachev, Krylov and Röckner, it follows that $f \in \mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega)$ for all $p \in (1,\infty)$. Therefore, upon identifying f with its precise representative, we have $f \in \mathrm{C}^{0,\alpha}(\Omega)$ for all $\alpha \in (0,1)$ as a consequence of Morrey's embedding theorem. In particular, $\{f = 0\}$ is a relatively closed set in Ω.

(ii) For any $\mu \in \mathbb{R}$, we define

$$\mathbb{B}_\mu[f,g] := \int_\Omega ((\mathrm{AD}f - \mathrm{L}f)\cdot \mathrm{D}g + (\mu - \mathrm{K})fg)\,\mathrm{d}\mathcal{L}^n.$$

Then, $\mathbb{B}_\mu : \mathrm{W}^{1,2}_0(\Omega)\times \mathrm{W}^{1,2}_0(\Omega) \to \mathbb{R}$ is a continuous bilinear form, which is also coercive for all $\mu \geq \mu^*$, where the value $\mu^* > 0$ depends only on the coefficients. By the Lax–Milgram theorem, the operator $\mathfrak{L}_\mu : \mathrm{W}^{1,2}_0(\Omega) \to \mathrm{W}^{-1,2}(\Omega)$ given by

$$\mathfrak{L}_\mu f := \mathrm{div}(\mathrm{AD}f - \mathrm{L}f) + \mathrm{K}f - \mu f$$

is a linear isomorphism. Therefore, the inverse $(\mathfrak{L}_\mu)^{-1} : \mathrm{W}^{-1,2}(\Omega) \to \mathrm{W}^{-1,2}(\Omega)$ is a compact operator as a consequence of the compactness of the canonical inclusions $\mathrm{W}^{1,2}_0(\Omega) \Subset \mathrm{L}^2(\Omega) \Subset \mathrm{W}^{-1,2}(\Omega)$. By the Fredholm alternative, the equation

$$f + (\mu(\mathfrak{L}_\mu)^{-1})f = (\mathfrak{L}_\mu)^{-1}\phi \tag{2-25}$$

either has a solution $f \not\equiv 0$ when $\phi \equiv 0$, or else is solvable for all $\phi \in \mathrm{W}^{-1,2}(\Omega)$. By noting that (2-25) can be rewritten as $\mathfrak{L}_\mu f = \phi - \mu f$, which is equivalent to $\mathfrak{L}_0 f = \phi$, it follows that either $\mathfrak{L}_0 f = 0$ has a solution $f^* \in \mathrm{W}^{1,2}_0(\Omega)\setminus\{0\}$, or else the problem $\mathfrak{L}_0 h = \mathrm{K} - \mathrm{div}\,\mathrm{L}$ has a solution $h \in \mathrm{W}^{1,2}_0(\Omega)$. The latter is equivalent to the solvability of the boundary value problem $\mathfrak{L}_0 f = 0$ in Ω with $f = 1$ on $\partial\Omega$ by some $f^* \in \mathrm{W}^{1,2}(\Omega)\setminus\{0\}$. In either case, it follows that (2-24) has a nontrivial solution $f^* \not\equiv 0$.

(iii) We begin by considering the auxiliary boundary value problem

$$\mathrm{div}(\mathrm{AD}\psi) = \mathrm{div}\,\mathrm{L} \ \text{ on } \Omega, \quad \psi = 0 \ \text{ on } \partial\Omega.$$

By the Lax–Milgram theorem, there exists a unique solution $\psi \in \mathrm{W}^{1,2}_0(\Omega)$, which by [Di Fazio 1996, Theorem 3.1] satisfies $\psi \in \mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega)$ for all $p \in (1,\infty)$. By Morrey's theorem, (the precise representative of) ψ is in $\mathrm{C}^{0,\alpha}(\Omega)$ for all $\alpha \in (0,1)$. We now set $\phi := e^{-\psi} f$, where $f \in \mathrm{L}^1(\Omega)$ is any given nonconstant solution to (2-24). By part (i), all three functions ϕ, ψ, f lie in $\mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega)$ for all $p \in (1,\infty)$, whilst $\mathrm{L} \in \mathrm{L}^\infty(\Omega;\mathbb{R}^n)$. Therefore, by using that $\mathrm{div}(\mathrm{AD}\psi - \mathrm{L}) = 0$ in Ω, we may rigorously compute

$$\begin{aligned} 0 = \mathrm{div}(\mathrm{AD}f - \mathrm{L}f) + \mathrm{K}f &= \mathrm{div}(e^\psi \mathrm{AD}\phi + e^\psi \phi(\mathrm{AD}\psi - \mathrm{L})) + \mathrm{K}e^\psi\phi \\ &= \mathrm{div}((e^\psi \mathrm{A})\mathrm{D}\phi) + (\mathrm{AD}\psi - \mathrm{L})\cdot[e^\psi \mathrm{D}\phi + e^\psi \phi \mathrm{D}\psi] + (e^\psi \mathrm{K})\phi \end{aligned}$$

in Ω. Hence the function ϕ is a weak solution to the divergence PDE

$$\mathrm{div}((e^\psi \mathrm{A})\mathrm{D}\phi) + [e^\psi(\mathrm{AD}\psi - \mathrm{L})]\cdot \mathrm{D}\phi + \big[e^\psi((\mathrm{AD}\psi - \mathrm{L})\cdot \mathrm{D}\psi + \mathrm{K})\big]\phi = 0 \quad \text{in } \Omega.$$

By the previous, the coefficients satisfy

$$e^{\psi}(\mathrm{AD}\psi - \mathrm{L}),\ e^{\psi}((\mathrm{AD}\psi - \mathrm{L})\cdot \mathrm{D}\psi + \mathrm{K}) \in \mathrm{L}^p_{\mathrm{loc}}(\Omega), \quad e^{\psi}\mathrm{A} \in \mathrm{C}^{0,\alpha}(\Omega; \mathbb{R}^{n\times n})$$

for all $p \in (1,\infty)$ and all $\alpha \in (0,1)$. By [Hardt and Simon 1989, Theorem (1.7) and Remark (1.8)(3)], it follows that, for any point $\bar{x} \in \Omega$ at which ϕ has finite order of vanishing, the set $\{\phi = 0\} \cap \mathbb{B}_\rho(\bar{x})$ is countably $(n-1)$-rectifiable for all sufficiently small $\rho > 0$. Since $f = e^{\psi}\phi$, we have $\{f = 0\} = \{\phi = 0\}$, and since $e^{\pm\psi} \in \mathrm{L}^\infty_{\mathrm{loc}}(\Omega)$, it follows that f and ϕ have the same sets of points with finite order of vanishing. On the other hand, by the strong unique continuation principle applied to (2-24), it follows that f has finite order of vanishing everywhere on Ω; see, e.g., [Koch and Tataru 2001a; 2001b]. Hence $\{f = 0\}$ is countably $(n-1)$-rectifiable and in particular it is a Lebesgue nullset. The proposition follows. □

3. Proofs of the main results

We may now establish our first main result.

Proof of Theorem 2. The idea here is based on the use of second-order Taylor expansions of $|F|^2$ at J^2u_∞, showing that the remainder can be made very small and of the right sign for small enough variations, as a consequence of the compactness of the set of directions and of the fulfilment of the PDEs.

Step 1: Let $u_\infty \in \mathcal{W}^{2,\infty}(\Omega)$ and $f_\infty \in \mathrm{L}^1(\Omega)$ be given, and suppose they satisfy the system of equations (1-8)–(1-9). We assume first that

$$f_\infty \not\equiv 0.$$

Step 2: We define the functions

$$\mathrm{K}_\infty := \frac{\partial_\eta F(\mathrm{J}^2u_\infty)}{\partial_\xi F(\mathrm{J}^2u_\infty)}, \quad \mathrm{L}_\infty := \frac{\partial_{\mathrm{p}} F(\mathrm{J}^2u_\infty)}{\partial_\xi F(\mathrm{J}^2u_\infty)}, \tag{3-1}$$

and note that K_∞, $|\mathrm{L}_\infty|$ are in $\mathrm{L}^\infty(\Omega)$: indeed, by (1-7) we have that $\partial_\xi F(\mathrm{J}^2u_\infty) \geq c$ a.e. on Ω, and also

$$|\partial_\eta F(\mathrm{J}^2u_\infty)| + |\partial_{\mathrm{p}} F(\mathrm{J}^2u_\infty)| \leq C(\|u_\infty\|_{\mathrm{W}^{1,\infty}(\Omega)} + \|\mathrm{A}{:}\mathrm{D}^2u_\infty\|_{\mathrm{L}^\infty(\Omega)})$$

a.e. on Ω.

Step 3: By (1-9) and (3-1), $f_\infty \in \mathrm{L}^1(\Omega)$ is a distributional solution to

$$\mathrm{div}(\mathrm{AD}f_\infty - \mathrm{L}_\infty f_\infty) + \mathrm{K}_\infty f_\infty = 0 \quad \text{in } \Omega. \tag{3-2}$$

By Proposition 19, we have $\mathcal{L}^n(\{f_\infty = 0\}) = 0$ and $f_\infty \in \mathrm{W}^{1,p}_{\mathrm{loc}}(\Omega)$ for all $p \in (1,\infty)$. Additionally, upon identifying f_∞ with its precise representative, the nodal set $\{f_\infty = 0\}$ is countably $(n-1)$-rectifiable and closed.

Step 4: We will now establish that there exists some $C \in \mathrm{C}^\nearrow[0,\infty)$ such that

$$|F(x,\eta,\mathrm{p},\xi)| \leq C(|\eta| + |\mathrm{p}| + |\xi|)$$

for a.e. $x \in \Omega$ and all $(\eta, \mathrm{p}, \xi) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$. With this aim, recalling (2-21) and assumption (1-7), it follows that

$$|F(x,\eta,\mathrm{p},\xi)| = |F(x,\eta,\mathrm{p},\xi) - F(x,\eta,\mathrm{p},\bar{\xi}(x,\eta,\mathrm{p}))| = \left| \int_{\bar{\xi}(x,\eta,\mathrm{p})}^{\xi} \partial_\xi F(x,\eta,\mathrm{p},\cdot)\, \mathrm{d}\mathcal{L}^1 \right|,$$

which yields

$$\begin{aligned} |F(x,\eta,\mathrm{p},\xi)| &\leq |\bar{\xi}(x,\eta,\mathrm{p}) - \xi| C(|\bar{\xi}(x,\eta,\mathrm{p})| + |\xi| + |\eta| + |\mathrm{p}|) \\ &\leq (C(|\eta| + |\mathrm{p}|) + |\xi|) C(C(|\eta| + |\mathrm{p}|) + |\xi| + |\eta| + |\mathrm{p}|). \end{aligned}$$

Hence the desired estimate follows for the new function in $\mathrm{C}^{\nearrow}[0,\infty)$ given by $t \mapsto [C(t)+t]C(C(t)+t)$.

Step 5: By assumption (1-7), we may fix $x \in \Omega$ (this is possible on a subset of full measure) and apply Taylor's theorem to $|F(x,\,\cdot\,,\,\cdot\,,\cdot)|^2 \in \mathrm{C}^2(\mathbb{R} \times \mathbb{R}^n \times \mathbb{R})$ to deduce that

$$\begin{aligned} |F(\mathrm{J}^2 u_\infty + \mathrm{J}^2\psi)|^2 = |F(\mathrm{J}^2 u_\infty)|^2 &+ 2F(\mathrm{J}^2 u_\infty)\{\partial_* F(\mathrm{J}^2 u_\infty) \cdot \mathrm{J}^2_* \psi\} \\ &+ \left(\int_0^1 (1-\lambda) \partial^2_{**}(|F|^2)(\mathrm{J}^2 u_\infty + \lambda \mathrm{J}^2 \psi)\, \mathrm{d}\lambda \right) : \mathrm{J}^2_* \psi \otimes \mathrm{J}^2_* \psi \end{aligned} \tag{3-3}$$

a.e. on Ω. By the identity (2-2) for $p = 2$, assumption (1-7) and Step 4, we have the bound

$$|\partial^2_{**}(|F|^2)(x,\eta,\xi)| \leq C(|\eta| + |\mathrm{p}| + |\xi|)$$

for some function $C \in \mathrm{C}^{\nearrow}[0,\infty)$. By using this bound and (1-8) (which implies $|F(\mathrm{J}^2 u_\infty)| = \mathrm{F}_\infty(u_\infty)$ a.e. on Ω), (3-3) yields

$$\begin{aligned} |F(\mathrm{J}^2 u_\infty + \mathrm{J}^2\psi)|^2 \geq |\mathrm{F}_\infty(u_\infty)|^2 &+ 2\mathrm{F}_\infty(u_\infty)\, \mathrm{sgn}(f_\infty)\{\partial_* F(\mathrm{J}^2 u_\infty) \cdot \mathrm{J}^2_* \psi\} \\ &- C(\|\mathrm{J}^2_* u_\infty\|_{\mathrm{L}^\infty(\Omega)} + \|\mathrm{J}^2_* \psi\|_{\mathrm{L}^\infty(\Omega)}) \|\mathrm{J}^2_* \psi\|^2_{\mathrm{L}^\infty(\Omega)} \end{aligned} \tag{3-4}$$

a.e. on Ω.

Step 6: Let us now define the functional $\Theta_\infty : \mathcal{W}^{2,\infty}_0(\Omega) \to \mathbb{R}$ by setting

$$\Theta_\infty(\psi) := \operatorname*{ess\,sup}_\Omega \big[\mathrm{sgn}(f_\infty)\{\partial_* F(\mathrm{J}^2 u_\infty) \cdot \mathrm{J}^2_* \psi\} \big], \tag{3-5}$$

and set

$$N_\infty := C(1 + \|\mathrm{J}^2_* u_\infty\|_{\mathrm{L}^\infty(\Omega)}). \tag{3-6}$$

In view of (3-5)–(3-6), we may deduce from (3-4) that

$$(\mathrm{F}_\infty(u_\infty + \psi))^2 \geq (\mathrm{F}_\infty(u_\infty))^2 + \{2\mathrm{F}_\infty(u_\infty)\Theta_\infty(\psi) - N_\infty \|\mathrm{J}^2_* \psi\|^2_{\mathrm{L}^\infty(\Omega)}\} \tag{3-7}$$

for all $\psi \in \mathcal{W}^{2,\infty}_0(\Omega)$ satisfying $\|\mathrm{J}^2_* \psi\|_{\mathrm{L}^\infty(\Omega)} \leq 1$ (namely in the closed unit ball of $\mathcal{W}^{2,\infty}_0(\Omega)$ with respect to the $\|\cdot\|'_{\mathcal{W}^{2,\infty}_0(\Omega)}$-norm in (2-13)).

Step 7: Let $\mathcal{K} \subseteq \mathbb{S}^{\mathcal{W}^{2,\infty}_0(\Omega)}$ be a compact set. We will regard the unit sphere of $\mathcal{W}^{2,\infty}_0(\Omega)$ to be with respect to the $\|\cdot\|'_{\mathcal{W}^{2,\infty}_0(\Omega)}$-norm in (2-13). We aim to show that there exists $\theta_\infty > 0$ such that

$$\Theta_\infty(\psi) \geq \theta_\infty \quad \text{for all } \psi \in \mathcal{K}. \tag{3-8}$$

Temporarily assuming that (3-8) holds, we will now show how we can complete the proof of the theorem. From (3-7)–(3-8), for any $\psi \in \mathcal{K}$ and $0 < t < 1$, by noting that Θ_∞ is 1-homogeneous, we have

$$(\mathrm{F}_\infty(u_\infty + t\psi))^2 \geq (\mathrm{F}_\infty(u_\infty))^2 + t\{2\mathrm{F}_\infty(u_\infty)\theta_\infty - tN_\infty\}$$

for all $\psi \in \mathcal{K}$. By restricting $t \in (0, \rho)$, where

$$\rho := \frac{2\mathrm{F}_\infty(u_\infty)\theta_\infty}{N_\infty},$$

the desired conclusion (1-10) follows (having assumed that (3-8) holds true).

Step 8: The remainder of the proof is devoted to establishing the lower bound (3-8) for the functional $\Theta_\infty : \mathcal{W}_0^{2,\infty}(\Omega) \to \mathbb{R}$ defined in (3-5). We will first show

$$\Theta_\infty(\psi) > 0, \quad \psi \in \mathcal{W}_0^{2,\infty}(\Omega) \setminus \{0\}. \tag{3-9}$$

With this aim, fix $\psi \in \mathcal{W}_0^{2,\infty}(\Omega) \setminus \{0\}$. By Proposition 11(ii), the divergence equation (1-9) is satisfied also weakly in $(\mathcal{W}_0^{2,\infty}(\Omega))^*$, namely when testing against functions in the space $\mathcal{W}_0^{2,\infty}(\Omega)$. Thus, (1-9) implies

$$\int_\Omega f_\infty(\mathrm{A}{:}\mathrm{D}^2\psi + \mathrm{L}_\infty \cdot \mathrm{D}\psi + \mathrm{K}_\infty\psi)\,\mathrm{d}\mathcal{L}^n = 0. \tag{3-10}$$

We now claim that

$$\mathcal{L}^n(\{f_\infty(\mathrm{A}{:}\mathrm{D}^2\psi + \mathrm{L}_\infty \cdot \mathrm{D}\psi + \mathrm{K}_\infty\psi) \neq 0\}) > 0. \tag{3-11}$$

For the sake of contradiction we suppose that

$$\mathrm{A}{:}\mathrm{D}^2\psi + \mathrm{L}_\infty \cdot \mathrm{D}\psi + \mathrm{K}_\infty\psi = 0 \quad \text{a.e. on } \Omega.$$

Since $|\psi| = |\mathrm{D}\psi| = 0$ on $\partial\Omega$ and $\psi \in W_0^{2,q}(\Omega)$ for all $q \in (1, \infty)$, the extension $\bar{\psi}$ of ψ by zero to $\mathbb{R}^n \setminus \Omega$ is compactly supported in $\mathbb{R}^n$ and satisfies $\bar{\psi} \in \mathrm{W}_c^{2,q}(\mathbb{R}^n)$, solving strongly the PDE

$$\mathrm{A}{:}\mathrm{D}^2\bar{\psi} + \bar{\mathrm{L}}_\infty \cdot \mathrm{D}\bar{\psi} + \overline{\mathrm{K}}_\infty\bar{\psi} = 0 \quad \text{a.e. on } \mathbb{R}^n,$$

where $\overline{\mathrm{K}}_\infty$, $|\bar{\mathrm{L}}_\infty| \in \mathrm{L}_c^\infty(\mathbb{R}^n)$ are the extensions of K_∞, L_∞ by zero on $\mathbb{R}^n \setminus \Omega$. In that case, by [Hardt and Simon 1989, Theorem (1.7) and Remark (1.8)] we infer that $\mathcal{L}^n(\{\bar{\psi} = 0\}) = 0$, contradicting that we have selected $\bar{\psi} \equiv 0$ on $\mathbb{R}^n \setminus \{0\}$. Since we already know from Step 3 that $f_\infty = 0$ at most on an $\mathcal{L}^n$-nullset, (3-11) follows. Further, we already know from (3-10) that the average must vanish over Ω. As a result, there must exist Borel sets $\Omega^\pm \subseteq \Omega$ with $\mathcal{L}^n(\Omega^\pm) > 0$ such that

$$\frac{f_\infty}{|f_\infty|}(\mathrm{A}{:}\mathrm{D}^2\psi + \mathrm{L}_\infty \cdot \mathrm{D}\psi + \mathrm{K}_\infty\psi) \gtrless 0 \quad \text{a.e. on } \Omega^\pm.$$

In view of (3-1), the above yields

$$\frac{f_\infty}{|f_\infty|}(\partial_\xi F(\mathrm{J}^2u_\infty)\mathrm{A}{:}\mathrm{D}^2\psi + \partial_\mathrm{p} F(\mathrm{J}^2u_\infty) \cdot \mathrm{D}\psi + \partial_\eta F(\mathrm{J}^2u_\infty)\psi) > 0 \quad \text{a.e. on } \Omega^+.$$

Using our notation for jets, we can rewrite the above as

$$\mathrm{sgn}(f_\infty)\{\partial_* F(\mathrm{J}^2u_\infty) \cdot \mathrm{J}_*^2\psi\} > 0 \quad \text{a.e. on } \Omega^+,$$

and $\mathcal{L}^n(\Omega^+) > 0$. In conclusion, we obtain (in view of (3-5)) that

$$\Theta_\infty(\psi) = \operatorname*{ess\,sup}_\Omega \big[\operatorname{sgn}(f_\infty)\{\partial_* F(\mathrm{J}^2 u_\infty)\cdot \mathrm{J}^2_* \psi\}\big] > 0$$

for any $\psi \in \mathcal{W}^{2,\infty}_0(\Omega)\setminus\{0\}$.

Step 9: Now we establish the bound (3-8) for the functional Θ_∞. In view of Step 8, it suffices to show that Θ_∞ is continuous. Indeed, on the one hand, for any ϕ, $\psi \in \mathcal{W}^{2,\infty}_0(\Omega)$ we have the estimate

$$\Theta_\infty(\phi) = \Theta_\infty(\phi+\psi-\psi) \le \Theta_\infty(\phi+\psi) + \|\mathrm{J}^2_* u_\infty\|_{\mathrm{L}^\infty(\Omega)}\|\mathrm{J}^2_*\psi\|_{\mathrm{L}^\infty(\Omega)}.$$

On the other hand, by the convexity of Θ_∞, we have that

$$\begin{aligned}\Theta_\infty(\phi+\psi) - \Theta_\infty(\phi) &\le \left(\Theta_\infty\left(\phi + \frac{\psi}{\|\mathrm{J}^2_*\psi\|_{\mathrm{L}^\infty(\Omega)}}\right) - \Theta_\infty(\phi)\right)\|\mathrm{J}^2_*\psi\|_{\mathrm{L}^\infty(\Omega)}\\ &\le 2\|\mathrm{J}^2_* u_\infty\|_{\mathrm{L}^\infty(\Omega)}(1+\|\mathrm{J}^2_*\phi\|_{\mathrm{L}^\infty(\Omega)})\|\mathrm{J}^2_*\psi\|_{\mathrm{L}^\infty(\Omega)}.\end{aligned}$$

Step 10: In conclusion, Θ_∞ is continuous and positive on the compact set $\mathcal{K}$, yielding that

$$\theta_\infty := \inf_{\mathcal{K}} \Theta_\infty > 0.$$

The proof of Theorem 2 is now complete. □

We may now establish our second main result.

Proof of Theorem 3. The main idea here is the same as in Theorem 2, building on the use of Taylor expansions of $|F|^2$ at $\mathrm{J}^2 u_\infty$. The new ingredient though is that the additional assumptions together with the fulfilment of the equations can guarantee the remainder having the right sign when the classes of variations are restricted in a manner compatible with the type of minimiser sought to obtain the local/narrow/global version of the concept without any compactness assumption on the set of directions.

Let $p \ge 2$ and fix $\psi \in \mathcal{W}^{2,\infty}_0(\Omega)$. We continue our arguments from the proof of Step 5 of Theorem 2. By Taylor's theorem for $|F(x,\,\cdot\,,\cdot\,)|^p$, we have

$$\begin{aligned}|F(\mathrm{J}^2 u_\infty + \mathrm{J}^2\psi)|^p = |F(\mathrm{J}^2 u_\infty)|^p + p|F(\mathrm{J}^2 u_\infty)|^{p-2}F(\mathrm{J}^2 u_\infty)\{\partial_* F(\mathrm{J}^2 u_\infty)\cdot \mathrm{J}^2_*\psi\}\\ + \left(\int_0^1 (1-\lambda)\partial^2_{**}(|F|^p)(\mathrm{J}^2 u_\infty + \lambda \mathrm{J}^2\psi)\,\mathrm{d}\lambda\right) : \mathrm{J}^2_*\psi\otimes\mathrm{J}^2_*\psi \quad (3\text{-}12)\end{aligned}$$

a.e. on Ω. We will now show that, under (1-11)–(1-13), the last term of (3-12) above is nonnegative for all sufficiently large $p \ge 2$ when ψ is restricted appropriately. More precisely:

Case 1: Suppose that (1-11) is satisfied. Then, the essential range of $\mathrm{J}^2 u_\infty$ satisfies $\operatorname{ess}((\mathrm{J}^2 u_\infty)(\Omega)) \subseteq \Omega\times L\times R$. Since L, R are open intervals and the essential range is a compact set, for any $\psi \in \mathcal{W}^{2,\infty}_0(\Omega)$ and any $\lambda \in (0,1)$, we may select $\rho > 0$ small enough such that $\|\mathrm{J}^2_*\psi\|_{\mathrm{L}^\infty(\Omega)} < \rho$ in order to obtain

$$\operatorname{ess}((\mathrm{J}^2 u_\infty + \lambda \mathrm{J}^2\psi)(\Omega)) \subseteq \Omega\times L\times R.$$

By choosing the corresponding $\bar p \ge 2$, it follows that $\partial^2_{**}(|F|^p)(\mathrm{J}^2 u_\infty + \lambda \mathrm{J}^2\psi) \ge 0$ for all $p \ge \bar p$ when $\psi \in \mathbb{B}^{\mathcal{W}^{2,\infty}_0(\Omega)}_\rho$.

Case 2: Suppose (1-12) is satisfied. Then, the essential range satisfies $\operatorname{ess}(u_\infty(\Omega)) \subseteq L$. Since L is an open interval and the essential range is a compact set, for any $\psi \in \mathcal{W}_0^{2,\infty}(\Omega)$ and any $\lambda \in (0,1)$, we may select $\rho > 0$ small enough that $\|\psi\|_{L^\infty(\Omega)} < \rho$ and any sufficiently large interval $R \subseteq \mathbb{R}$ in order to obtain

$$\operatorname{ess}((J^2 u_\infty + \lambda J^2 \psi)(\Omega)) \subseteq \Omega \times L \times R.$$

Then, for the corresponding $\bar{p} \geq 2$, it follows that $\partial^2_{**}(|F|^p)(J^2 u_\infty + \lambda J^2 \psi) \geq 0$ for all $p \geq \bar{p}$ when $\psi \in \mathbb{B}_\rho^{L^\infty(\Omega)} \cap \mathcal{W}_0^{2,\infty}(\Omega)$.

Case 3: Suppose (1-13) is satisfied. Fix any $\psi \in \mathcal{W}_0^{2,\infty}(\Omega)$ and choose $R, L \subseteq \mathbb{R}$ large enough and the corresponding $\bar{p} \geq 2$, so that the essential range satisfies

$$\operatorname{ess}((J^2 u_\infty + \lambda J^2 \psi)(\Omega)) \subseteq \Omega \times L \times R.$$

It follows that $\partial^2_{**}(|F|^p)(J^2 u_\infty + \lambda J^2 \psi) \geq 0$ for all $p \geq \bar{p}$.

In conclusion, for any fixed $\psi \in \mathcal{W}_0^{2,\infty}(\Omega)$ appropriately restricted depending on the assumptions, we have

$$|F(J^2 u_\infty + J^2 \psi)|^p - |F(J^2 u_\infty)|^p \geq p|F(J^2 u_\infty)|^{p-2} F(J^2 u_\infty)\{\partial_* F(J^2 u_\infty) \cdot J^2_* \psi\}$$

a.e. on Ω. In view of (1-8) and (1-1), the above yields

$$\operatorname*{ess\,sup}_\Omega |F(J^2 u_\infty + J^2 \psi)|^p - (F_\infty(u_\infty))^p \geq p(F_\infty(u_\infty))^{p-1} \operatorname*{ess\,sup}_\Omega \big[\operatorname{sgn}(f_\infty)\{\partial_* F(J^2 u_\infty) \cdot J^2_* \psi\}\big].$$

By (3-5) and (1-1), since powers commute with the essential supremum of nonnegative functions, the above inequality can be rewritten as

$$(F_\infty(u_\infty + \psi))^p - (F_\infty(u_\infty))^p \geq p(F_\infty(u_\infty))^{p-1} \Theta_\infty(\psi).$$

By Step 8 in the proof of Theorem 2, we know that $\Theta_\infty(\psi) > 0$ for all functions $\psi \in \mathcal{W}_0^{2,\infty}(\Omega) \setminus \{0\}$. Therefore, if $F_\infty(u_\infty) > 0$, we have obtained the inequality

$$F_\infty(u_\infty) < F_\infty(u_\infty + \psi)$$

when $\psi \neq 0$. If on the other hand $F_\infty(u_\infty) = 0$, the previous arguments do not allow us to conclude that $F_\infty(u_\infty + \psi) > 0 = F_\infty(u_\infty)$ when $\psi \in \mathcal{W}_0^{2,\infty}(\Omega) \setminus \{0\}$. To show that the last inequality is indeed strict, suppose for the sake of contradiction that instead $F(J^2 u_\infty + J^2 \psi) = 0$ a.e. on Ω for some $\psi \in \mathcal{W}_0^{2,\infty}(\Omega) \setminus \{0\}$. In view of (2-21), this can be restated as

$$A{:}D^2 \psi = \bar{\xi}(\cdot\,, u_\infty + \psi, Du_\infty + D\psi) - A{:}D^2 u_\infty \quad \text{a.e. in } \Omega. \tag{3-13}$$

Further, since we also have $F(J^2 u_\infty) = 0$ a.e. on Ω, (3-13) is true when $\psi \equiv 0$ as well, namely

$$A{:}D^2 u_\infty = \bar{\xi}(\cdot\,, u_\infty, Du_\infty) \quad \text{a.e. in } \Omega. \tag{3-14}$$

By (3-13)–(3-14), we deduce

$$A{:}D^2 \psi = \bar{\xi}(\cdot\,, u_\infty + \psi, Du_\infty + D\psi) - \bar{\xi}(\cdot\,, u_\infty, Du_\infty) \quad \text{a.e. in } \Omega. \tag{3-15}$$

By Taylor's theorem, we can write

$$\bar{\xi}(\,\cdot\,, u_\infty + \psi, \mathrm{D}u_\infty + \mathrm{D}\psi) - \bar{\xi}(\,\cdot\,, u_\infty, \mathrm{D}u_\infty)$$
$$= \left[\int_0^1 \partial_\eta \bar{\xi}(\,\cdot\,, u_\infty + \lambda\psi, \mathrm{D}u_\infty + \lambda\mathrm{D}\psi)\,\mathrm{d}\lambda\right]\psi + \left[\int_0^1 \partial_{\mathrm{p}} \bar{\xi}(\,\cdot\,, u_\infty + \lambda\psi, \mathrm{D}u_\infty + \lambda\mathrm{D}\psi)\,\mathrm{d}\lambda\right]\cdot \mathrm{D}\psi \quad (3\text{-}16)$$

a.e. on Ω. Note also that $|\psi| = |\mathrm{D}\psi| = 0$ $\mathcal{H}^{n-1}$-a.e. on $\partial\Omega$ and $\psi \in \mathrm{W}^{2,q}_0(\Omega)$ for all $q \in (1, \infty)$. Let ψ be extended by zero on $\mathbb{R}^n \setminus \Omega$ to a compactly supported function $\psi \in \mathrm{W}^{2,q}_c(\mathbb{R}^n)$. Then, by (3-15)–(3-16), there exist compactly supported essentially bounded functions $\mathfrak{B} \in \mathrm{L}^\infty_c(\mathbb{R}^n; \mathbb{R}^n)$ and $\mathfrak{C} \in \mathrm{L}^\infty_c(\mathbb{R}^n)$ such that ψ satisfies the next PDE in the strong sense:

$$\mathrm{A}{:}\mathrm{D}^2\psi + \mathfrak{B}\cdot\mathrm{D}\psi + \mathfrak{C}\psi = 0 \quad \text{a.e. in } \mathbb{R}^n.$$

By [Hardt and Simon 1989, Theorem (1.7) and Remark (1.8)], we infer that $\mathcal{L}^n(\{\psi = 0\}) = 0$, which contradicts that $\psi \equiv 0$ on $\mathbb{R}^n \setminus \Omega$. Therefore, the proof of Theorem 3 is complete. □

Now we will establish the proofs of Theorems 4 and 7 together by utilising Theorem 16 (established in Section 2.5).

Proofs of Theorems 4 and 7. The main ideas here can be summarised as follows: in order to show that a "localised" minimiser satisfies the equations, the key ingredient is the passage through L^p approximations as $p \to \infty$. Namely, we aim at minimising with the direct method the corresponding L^p functional, obtaining the Euler–Lagrange equations, and then trying to show that the approximate objects converge to the desired ones. There are several issues in trying to implement this approach. Firstly, standard local minimisers are not stable under the weak/weak* modes of convergence we obtain subsequentially as $p \to \infty$ from the L^p approximations, whilst we also have standard weak continuity problems as we involve nonlinear functionals. The lack of stability under weak limits is the main rationale behind the introduction of narrow minimisers, which resolve this problem, and for technical convenience we treat them as constrained minimisers. However, even if we start from a narrow minimiser in L^∞, the L^p approximations will converge to *some* narrow minimiser in L^∞, but not necessarily the preselected one, as we do not have uniqueness. Instead, we wish to use this device to prove uniqueness. (This is radically different from the first-order case, whereat uniqueness is obtained for the ∞-Laplacian through the maximum principle, and then uniqueness of subsequential limits of L^p approximations follows from there.) The penalisation device resolves this issue, as it makes it energetically too expensive for the L^p approximations to converge to a narrow L^∞ minimiser different from the preselected one. Finally, our weak convergence issues are resolved by utilising the PDE estimates established in earlier sections. One further particular difficulty not directly related to weak convergence issues is to show that f_∞ is nonzero on "fat" sets of positive measure. For this (and only this) we utilise the additional C^2 regularity on the boundary condition, and utilise an extension of the unique continuation principle adapted to our setting, which we established earlier for this specific reason.

Step 1: Let $u_\infty \in \mathcal{W}^{2,\infty}_{u_0}(\Omega)$ be a narrow minimiser of (1-1). By Definition 1, this means there exists $\rho > 0$ such that

$$\mathrm{F}_\infty(u_\infty) \leq \mathrm{F}_\infty(u_\infty + \psi), \quad \psi \in \mathbb{B}_\rho^{\mathrm{W}^{1,\infty}(\Omega)} \cap \mathcal{W}^{2,\infty}_0(\Omega).$$

We consider now the functional $\mathrm{F}_{p,1,u_\infty}$ given by (1-19) for $\epsilon = 1$ (and $\bar{u} = u_\infty$). For $p \geq p_0 := \bar{p} + n + 1/c^3$, where c is the constant from assumption (1-15), we set

$$\begin{aligned}\mathcal{E}^\infty_{u_\infty}(\Omega) &:= \overline{\mathbb{B}}^{\mathrm{W}^{1,\infty}(\Omega)}_{\rho/2}(u_\infty) \cap \mathcal{W}^{2,\infty}_{u_0}(\Omega),\\ \mathcal{E}^p_{u_\infty}(\Omega) &:= \overline{\mathbb{B}}^{\mathrm{W}^{1,\infty}(\Omega)}_{\rho/2}(u_\infty) \cap \mathrm{W}^{2,p}_{u_0}(\Omega), \quad p_0 \leq p < \infty.\end{aligned}$$

Then, the class of sets $\{\mathcal{E}^p_{u_\infty}(\Omega)\}_{p_0 \leq p \leq \infty}$ satisfies the properties set out in assumption (2-18). It suffices to confirm only that $\mathcal{E}^p_{u_\infty}(\Omega)$ is weakly closed in $\mathrm{W}^{2,p}_{u_0}(\Omega)$ when $p \geq p_0$: indeed, since $\mathcal{E}^p_{u_\infty}(\Omega)$ is convex and norm closed in $\mathrm{W}^{2,p}(\Omega)$, its weak closedness is a consequence of Mazur's theorem. Therefore, by Theorem 16, $\mathrm{F}_{p,1,u_\infty}$ has a minimiser $u_p \in \mathcal{E}^p_{u_\infty}(\Omega)$, namely

$$\mathrm{F}_{p,1,u_\infty}(u_p) \leq \mathrm{F}_{p,1,u_\infty}(v), \quad v \in \mathcal{E}^p_{u_\infty}(\Omega).$$

By setting $\phi := v - u_\infty \in \mathrm{W}^{2,p}_0(\Omega)$, we can rewrite the above condition as

$$\mathrm{F}_{p,1,u_\infty}(u_p) \leq \mathrm{F}_{p,1,u_\infty}(u_\infty + \phi), \quad \phi \in \overline{\mathbb{B}}^{\mathrm{W}^{1,\infty}(\Omega)}_{\rho/2} \cap \mathrm{W}^{2,p}_0(\Omega).$$

Note further that, again by Theorem 16, there exists a sequence $(p_j)_1^\infty$ along which $\mathrm{F}_{p_j,1,u_\infty}(u_p) \to \mathrm{F}_\infty(u_\infty)$, and also $u_{p_j} \rightharpoonup u_\infty$ in $\mathrm{W}^{2,q}_{u_0}(\Omega)$ for any fixed $q \in (n, \infty)$, both as $j \to \infty$. By Morrey's theorem, we have $u_{p_j} \to u_\infty$ in $\mathrm{W}^{1,\infty}(\Omega)$ as $j \to \infty$. This implies, for $j \geq j_0$ sufficiently large, that

$$u_{p_j} \in \mathbb{B}^{\mathrm{W}^{1,\infty}(\Omega)}_{\rho/4}(u_\infty) \cap \mathrm{W}^{2,p}_{u_0}(\Omega),$$

and therefore u_{p_j} is a point in the strong interior of $\mathcal{E}^p_{u_\infty}(\Omega)$. It follows that

$$\mathrm{F}_{p,1,u_\infty}(u_p) \leq \mathrm{F}_{p,1,u_\infty}(u_p + \psi), \quad \psi \in \mathbb{B}^{\mathrm{W}^{1,\infty}(\Omega)}_{\rho/4} \cap \mathrm{W}^{2,p}_0(\Omega),$$

which establishes that u_p is indeed a narrow minimiser of $\mathrm{F}_{p,1,u_\infty}$ in $\mathrm{W}^{2,p}_{u_0}(\Omega)$.

Step 2: Let u_p be the narrow minimiser of $\mathrm{F}_{p,1,u_\infty}$ in $\mathrm{W}^{2,p}_{u_0}(\Omega)$ from Step 1. Since the corresponding set of variations

$$\mathbb{B}^{\mathrm{W}^{1,\infty}(\Omega)}_{\rho/4}(u_\infty) \cap \mathrm{W}^{2,p}_{u_0}(\Omega)$$

is strongly (norm) open in $\mathrm{W}^{2,p}_{u_0}(\Omega)$ and F satisfies the growth and regularity conditions given in (1-6), (1-7) and (1-16), it follows that $\mathrm{F}_{p,1,u_\infty}$ is Gateaux differentiable in $\mathrm{W}^{2,p}_{u_0}(\Omega)$ at u_p, therefore u_p is a weak solution to the Euler–Lagrange equations, whose weak formulation is given by (1-21)–(1-22). Let f_p be defined by (1-23) and let us also define the functions

$$\begin{aligned}\mathrm{K}_p &:= \frac{\partial_\eta F(\mathrm{J}^2 u_p)}{\partial_\xi F(\mathrm{J}^2 u_p)} : \Omega \to \mathbb{R},\\ \mathrm{L}_p &:= \frac{\partial_\mathrm{p} F(\mathrm{J}^2 u_p)}{\partial_\xi F(\mathrm{J}^2 u_p)} : \Omega \to \mathbb{R}^n.\end{aligned} \tag{3-17}$$

By (1-7) and (1-16), we have the bounds

$$|\mathrm{K}_p| = \frac{|\partial_\eta F(\mathrm{J}^2 u_p)|}{\partial_\xi F(\mathrm{J}^2 u_p)} \leq \frac{C(\|u_p\|_{\mathrm{W}^{1,\infty}(\Omega)})}{c}(1+|\mathrm{A}{:}\mathrm{D}^2 u_p|),$$
$$|\mathrm{L}_p| = \frac{|\partial_{\mathrm{p}} F(\mathrm{J}^2 u_p)|}{\partial_\xi F(\mathrm{J}^2 u_p)} \leq \frac{C(\|u_p\|_{\mathrm{W}^{1,\infty}(\Omega)})}{c}(1+|\mathrm{A}{:}\mathrm{D}^2 u_p|).$$

In view of Lemma 17, the above estimate implies that, for any fixed $q \in (1,\infty)$, $\{\mathrm{K}_p\}_{p\geq p_0}$ and $\{|\mathrm{L}_p|\}_{p\geq p_0}$ are bounded in $\mathrm{L}^q(\Omega)$. By the identities

$$\begin{aligned} \mathrm{K}_p f_p &= e_p^{1-p}(|F|^{p-2}F\partial_\eta F)(\mathrm{J}^2 u_p), \\ \mathrm{L}_p f_p &= e_p^{1-p}(|F|^{p-2}F\partial_{\mathrm{p}} F)(\mathrm{J}^2 u_p), \end{aligned} \tag{3-18}$$

which follow by (1-23) and (3-17), and in view of our notation for reduced jets and reduced derivatives of F, we can rewrite (1-21)–(1-23) as

$$\mathrm{div}(\mathrm{A}\mathrm{D} f_p - \mathrm{L}_p f_p) + \mathrm{K}_p f_p + u_p - u_\infty = 0 \quad \text{weakly in } (\mathrm{W}^{2,p}_0(\Omega))^*. \tag{3-19}$$

We now show that $\{f_p\}_{p\geq p_0}$ and $\{\mathrm{K}_p f_p\}_{p\geq p_0}$ are bounded in $\mathrm{L}^1(\Omega)$ and that $\{\mathrm{L}_p f_p\}_{p\geq p_0}$ is bounded in $\mathrm{L}^1(\Omega;\mathbb{R}^n)$, at least along a subsequence of indices $(p_j)_1^\infty$. By (1-22), (1-23) and (1-16), if $p' = \frac{p}{p-1}$ is the conjugate exponent, we estimate

$$\begin{aligned} \|f_p\|_{\mathrm{L}^{p'}(\Omega)} &= e_p^{1-p}\left(\fint_\Omega ((|F|^{p-1}\partial_\xi F)(\mathrm{J}^2 u_p))^{p'}\,\mathrm{d}\mathcal{L}^n\right)^{\frac{1}{p'}} \\ &\leq e_p^{1-p} C(\|u_\infty\|_{\mathrm{W}^{1,\infty}(\Omega)}+1)\left(\fint_\Omega |F(\mathrm{J}^2 u_p)|^p\,\mathrm{d}\mathcal{L}^n\right)^{\frac{p-1}{p}} = C(\|u_\infty\|_{\mathrm{W}^{1,\infty}(\Omega)}+1) \end{aligned}$$

for all sufficiently large $p \geq p_0$ (made possible by (1-20)). Further, by (2-22) and (1-15), we have the estimate

$$|\mathrm{A}{:}\mathrm{D}^2 w| \leq \frac{|F(\mathrm{J}^2 w)| + (|w|+|\mathrm{D}w|)^\alpha + 1}{c} \quad \text{a.e. on } \Omega \tag{3-20}$$

for any $w \in \mathrm{W}^{2,1}_{\mathrm{loc}}(\Omega)$. Therefore, by (1-16) and (3-18), we may estimate

$$\begin{aligned} \|\mathrm{K}_p f_p\|_{\mathrm{L}^1(\Omega)} &\leq e_p^{1-p}\fint_\Omega \left|(|F|^{p-1}\partial_\eta F)(\mathrm{J}^2 u_p)\right|\mathrm{d}\mathcal{L}^n \\ &\leq e_p^{1-p} C(\|u_\infty\|_{\mathrm{W}^{1,\infty}(\Omega)}+1)\fint_\Omega |F(\mathrm{J}^2 u_p)|^{p-1}(1+|\mathrm{A}{:}\mathrm{D}^2 u_p|)\,\mathrm{d}\mathcal{L}^n. \end{aligned}$$

This implies

$$\begin{aligned} \|\mathrm{K}_p f_p\|_{\mathrm{L}^1(\Omega)} &\leq e_p^{1-p} C(\|u_\infty\|_{\mathrm{W}^{1,\infty}(\Omega)}+1)\left(\fint_\Omega |F(\mathrm{J}^2 u_p)|^p\,\mathrm{d}\mathcal{L}^n\right)^{\frac{p-1}{p}}\left(\fint_\Omega (1+|\mathrm{A}{:}\mathrm{D}^2 u_p|)^p\,\mathrm{d}\mathcal{L}^n\right)^{\frac{1}{p}} \\ &= C(\|u_\infty\|_{\mathrm{W}^{1,\infty}(\Omega)}+1)\left(\fint_\Omega (1+|\mathrm{A}{:}\mathrm{D}^2 u_p|)^p\,\mathrm{d}\mathcal{L}^n\right)^{\frac{1}{p}}. \end{aligned}$$

Hence, in view of (3-20), the above estimate yields (for large enough p)

$$\begin{aligned}\|\mathrm{K}_p f_p\|_{\mathrm{L}^1(\Omega)} &\le C(\|u_\infty\|_{\mathrm{W}^{1,\infty}(\Omega)}+1)\left[\fint_\Omega\left(1+\frac{1}{c}+\frac{(|u_p|+|\mathrm{D}u_p|)^\alpha}{c}+\frac{|F(\mathrm{J}^2u_p)|}{c}\right)^p \mathrm{d}\mathcal{L}^n\right]^{\frac{1}{p}}\\ &\le C(\|u_\infty\|_{\mathrm{W}^{1,\infty}(\Omega)}+1)\left[1+\frac{1}{c}+\frac{\|u_\infty\|^\alpha_{\mathrm{W}^{1,\infty}(\Omega)}+1}{c}+\frac{e_p}{c}\right]\\ &= C(\|u_\infty\|_{\mathrm{W}^{1,\infty}(\Omega)}+1)\frac{1}{c}(2+c+\|u_\infty\|^\alpha_{\mathrm{W}^{1,\infty}(\Omega)}+e_p).\end{aligned}$$

Again by (1-16), since $\partial_\eta F$ and $\partial_{\mathrm{p}} F$ satisfy the same bounds, by (3-18) we have that the same estimate is true also for $\mathrm{L}_p f_p$:

$$\|\mathrm{L}_p f_p\|_{\mathrm{L}^1(\Omega)} \le C(\|u_\infty\|_{\mathrm{W}^{1,\infty}(\Omega)}+1)\frac{1}{c}(2+c+\|u_\infty\|^\alpha_{\mathrm{W}^{1,\infty}(\Omega)}+e_p).$$

The above inequalities imply the desired bound for sufficiently large $p \ge p_0$ upon nothing that, by (1-22), (1-19) and (1-20), we have

$$e_p = \mathrm{F}_{p,1,u_\infty}(u_p) - \tfrac{1}{2}\|u_p-u_\infty\|^2_{\mathrm{L}^2(\Omega)} \to \mathrm{F}_\infty(u_\infty) \tag{3-21}$$

as $p_j \to \infty$. In view of the previous uniform L^1-bounds, there exist Radon measures

$$\varphi_\infty \in \mathcal{M}(\overline{\Omega}), \quad \kappa_\infty \in \mathcal{M}(\overline{\Omega}), \quad \lambda_\infty \in \mathcal{M}(\overline{\Omega};\mathbb{R}^n),$$

and by applying Lemma 18 (Section 2.6), there exists

$$f_\infty \in \mathrm{L}^1(\Omega) \cap \bigcap_{1<q<n'} \mathrm{L}^q_{\mathrm{loc}}(\Omega),$$

where $n' = \frac{n}{n-1}$, and a further subsequence, denoted again by $(p_j)_1^\infty$, along which

$$\begin{cases} f_p \to f_\infty & \text{in } \mathrm{L}^q_{\mathrm{loc}}(\Omega),\ q\in(1,n'),\\ f_p\mathcal{L}^n\llcorner_\Omega \overset{*}{\rightharpoonup} \varphi_\infty & \text{in } \mathcal{M}(\overline{\Omega}),\\ (\mathrm{K}_p f_p + u_p - u_\infty)\mathcal{L}^n\llcorner_\Omega \overset{*}{\rightharpoonup} \kappa_\infty & \text{in } \mathcal{M}(\overline{\Omega}),\\ (\mathrm{L}_p f_p)\mathcal{L}^n\llcorner_\Omega \overset{*}{\rightharpoonup} \lambda_\infty & \text{in } \mathcal{M}(\overline{\Omega};\mathbb{R}^n)\end{cases}$$

as $p_j \to \infty$. Further, since $(\mathrm{K}_{p_j})_1^\infty$ and $(|\mathrm{L}_{p_j}|)_1^\infty$ are both bounded in $\mathrm{L}^q(\Omega)$ for all $q \in (1,\infty)$, there exist

$$\mathrm{K}_\infty \in \bigcap_{1<q<\infty} \mathrm{L}^q(\Omega), \quad \mathrm{L}_\infty \in \bigcap_{1<q<\infty} \mathrm{L}^q(\Omega;\mathbb{R}^n) \tag{3-22}$$

such that $\mathrm{K}_p \rightharpoonup \mathrm{K}_\infty$ in $\mathrm{L}^q(\Omega)$ and also $\mathrm{L}_p \rightharpoonup \mathrm{L}_\infty$ in $\mathrm{L}^q(\Omega;\mathbb{R}^n)$ as $p_j \to \infty$ for all $q \in (1,\infty)$. Since $f_p \to f_\infty$ in $\mathrm{L}^q_{\mathrm{loc}}(\Omega)$ for all $q \in (1,n')$, we have

$$\mathrm{K}_p f_p \rightharpoonup \mathrm{K}_\infty f_\infty \quad \text{in } \mathrm{L}^r_{\mathrm{loc}}(\Omega) \text{ for all } r\in(1,n') \text{ as } p_j\to\infty,$$

and similarly

$$\mathrm{L}_p f_p \rightharpoonup \mathrm{L}_\infty f_\infty \quad \text{in } \mathrm{L}^r_{\mathrm{loc}}(\Omega;\mathbb{R}^n) \text{ for all } r\in(1,n') \text{ as } p_j\to\infty.$$

By (1-20) we also have that $u_p \to u_\infty$ in $\mathrm{C}(\overline{\Omega})$ as $p_j \to \infty$. Hence, by the uniqueness of limits, the measures φ_∞, κ_∞, λ_∞ are absolutely continuous with respect to the Lebesgue measure restricted to Ω, and in addition

$$\varphi_\infty \llcorner_\Omega = f_\infty \mathcal{L}^n \llcorner_\Omega, \quad \kappa_\infty \llcorner_\Omega = (\mathrm{K}_\infty f_\infty)\mathcal{L}^n \llcorner_\Omega, \quad \lambda_\infty \llcorner_\Omega = (\mathrm{L}_\infty f_\infty)\mathcal{L}^n \llcorner_\Omega.$$

We therefore have for any fixed $q \in (1, n')$ the convergences

$$\begin{cases} f_p \to f_\infty & \text{in } \mathrm{L}^q_{\mathrm{loc}}(\Omega), \\ f_p \mathcal{L}^n \llcorner_\Omega \overset{*}{\rightharpoonup} f_\infty \mathcal{L}^n \llcorner_\Omega + \varphi_\infty \llcorner_{\partial\Omega} & \text{in } \mathcal{M}(\overline{\Omega}), \\ (\mathrm{K}_p f_p + u_p - u_\infty)\mathcal{L}^n \llcorner_\Omega \overset{*}{\rightharpoonup} (\mathrm{K}_\infty f_\infty)\mathcal{L}^n \llcorner_\Omega + \kappa_\infty \llcorner_{\partial\Omega} & \text{in } \mathcal{M}(\overline{\Omega}), \\ (\mathrm{L}_p f_p)\mathcal{L}^n \llcorner_\Omega \overset{*}{\rightharpoonup} (\mathrm{L}_\infty f_\infty)\mathcal{L}^n \llcorner_\Omega + \lambda_\infty \llcorner_{\partial\Omega} & \text{in } \mathcal{M}(\overline{\Omega}; \mathbb{R}^n) \end{cases} \tag{3-23}$$

as $p_j \to \infty$. By passing to the limit as $j \to \infty$ in the PDE (3-19) along the sequence $(p_j)_1^\infty$, the previous considerations allow us to deduce

$$\mathrm{div}(\mathrm{A}\mathrm{D} f_\infty - \mathrm{L}_\infty f_\infty) + \mathrm{K}_\infty f_\infty = 0 \quad \text{in } (\mathrm{C}^\infty_c(\Omega))^*, \tag{3-24}$$

namely in the sense of distributions. By applying Proposition 11(ii), it follows that (3-24) is also satisfied weakly in $(\mathcal{W}^{2,\infty}_0(\Omega))^*$.

We now show that, if (1-17) is satisfied, $\varphi_\infty \llcorner_{\partial\Omega} = 0$. (This is actually the only place in the paper that the C^2-regularity assumption on $\partial\Omega$ is utilised.)

Claim 20. *If $\partial\Omega$ is C^2, we have that $\varphi_\infty \llcorner_{\partial\Omega} = 0$.*

Proof of Claim 20. By equations (3-24) and (3-19), we have

$$\mathrm{div}(\mathrm{A}\mathrm{D} f_p - \mathrm{A}\mathrm{D} f_\infty) = \mathrm{K}_\infty f_\infty - \mathrm{K}_p f_p + u_\infty - u_p - \mathrm{div}(\mathrm{L}_\infty f_\infty - \mathrm{L}_p f_p) \tag{3-25}$$

weakly in $(\mathcal{W}^{2,\infty}_0(\Omega))^*$. Since by assumption $\partial\Omega$ is C^2, there exists $r_0 > 0$ such that the distance function $\mathrm{dist}(\,\cdot\,, \partial\Omega)$ is C^2 when restricted to the inner tubular neighbourhood $\overline{\Omega} \cap \{\mathrm{dist}(\,\cdot\,, \partial\Omega) < r_0\}$. Let $d \in \mathrm{C}^2(\overline{\Omega})$ be any extension of the distance function $\mathrm{dist}(\,\cdot\,, \partial\Omega)$ from $\overline{\Omega} \cap \{\mathrm{dist}(\,\cdot\,, \partial\Omega) < r_0\}$ to $\overline{\Omega}$. Let us fix also any $g \in \mathrm{C}^2(\overline{\Omega})$. We set

$$\psi := d^2 g \in \mathrm{C}^2(\overline{\Omega}) \cap \mathcal{W}^{2,\infty}_0(\Omega).$$

By testing ψ in the weak formulation of (3-25), we have

$$\int_\Omega (f_p - f_\infty)\mathrm{A}{:}\mathrm{D}^2\psi \, \mathrm{d}\mathcal{L}^n = \int_\Omega (\mathrm{K}_\infty f_\infty - \mathrm{K}_p f_p + u_\infty - u_p)\psi \, \mathrm{d}\mathcal{L}^n + \int_\Omega (\mathrm{L}_\infty f_\infty - \mathrm{L}_p f_p)\cdot \mathrm{D}\psi \, \mathrm{d}\mathcal{L}^n. \tag{3-26}$$

By a computation, we have

$$\mathrm{D}^2\psi = 2d\Big(\mathrm{D}g \otimes \mathrm{D}d + \mathrm{D}d \otimes \mathrm{D}g + g\mathrm{D}^2 d + \frac{d}{2}\mathrm{D}^2 g\Big) + 2g\mathrm{D}d \otimes \mathrm{D}d \quad \text{on } \overline{\Omega},$$

noting also that

$$\mathrm{D}^2\psi|_{\partial\Omega} = 2g\mathrm{n} \otimes \mathrm{n},$$

where $\mathrm{n} = \mathrm{D}d$ is the outwards unit normal vector field on $\partial\Omega$. By (3-23), since $\psi \in \mathrm{C}_0(\overline{\Omega})$ and $\mathrm{D}\psi \in \mathrm{C}_0(\overline{\Omega}; \mathbb{R}^n)$, we may pass to the limit as $p_j \to \infty$ in (3-26) to obtain

$$\int_{\partial\Omega} (g\mathrm{A}{:}\mathrm{n} \otimes \mathrm{n})\, \mathrm{d}\varphi_\infty = 0$$

for any $g \in \mathrm{C}^2(\overline{\Omega})$. Fix now $\varepsilon > 0$ and an arbitrary $h \in \mathrm{C}(\partial\Omega)$. We may extend h to a function in $\mathrm{C}(\overline{\Omega})$, denoted again by h (e.g., by using its McShane–Whitney extension $x \mapsto \max_{y\in\partial\Omega}\{h(y) - |x - y|\}$). It follows that there exists $g \in \mathrm{C}^2(\overline{\Omega})$ such that

$$\left\| \frac{h}{\mathrm{A}{:}\mathrm{n} \otimes \mathrm{n}} - g \right\|_{\mathrm{C}(\partial\Omega)} < \frac{\varepsilon}{\|A\|}.$$

In conclusion, by letting "$\|\cdot\|$" denote the total variation of a Radon measure, we have the estimate

$$\left| \int_{\partial\Omega} h\, \mathrm{d}\varphi_\infty \right| \le \int_{\partial\Omega} |h - g\mathrm{A}{:}\mathrm{n} \otimes \mathrm{n}|\, \mathrm{d}\|\varphi_\infty\| < \varepsilon \|\varphi_\infty\|(\partial\Omega),$$

which yields that $\varphi_\infty \llcorner_{\partial\Omega} = 0$, as claimed. □

Note now that, by the definition (1-23) and the fact that the inverse function of $t \mapsto |t|^{p-2}t$ on $\mathbb{R}$ is given by $s \mapsto \operatorname{sgn}(s)|s|^{1/(p-1)}$, we have the identity

$$F(\mathrm{J}^2 u_p) = e_p \operatorname{sgn}(f_p)|f_p|^{\frac{1}{p-1}} (\partial_\xi F(\mathrm{J}^2 u_p))^{-\frac{1}{p-1}} \quad \text{a.e. on } \Omega. \tag{3-27}$$

Let us fix $x \in \Omega \setminus \{f_\infty = 0\}$. Since by (3-23) we have that $f_p \to f_\infty$ a.e. on Ω as $p_j \to \infty$, it follows that $f_p(x) \in \left[\frac{1}{2}|f_\infty(x)|, 2|f_\infty(x)|\right]$ for large enough $j \in \mathbb{N}$. By assumptions (1-7) and (1-16), as well as (3-21), we may pass to the a.e. limit in (3-27) to obtain

$$F(\mathrm{J}^2 u_p) \to \operatorname{sgn}(f_\infty)\mathrm{F}(u_\infty) \quad \text{a.e. on } \Omega \setminus \{f_\infty = 0\} \tag{3-28}$$

as $p_j \to \infty$. Again by (1-7), the function F is invertible with respect to the last variable, therefore (3-28) implies

$$\mathrm{A}{:}\mathrm{D}^2 u_p \to (F(\cdot\,, u_\infty, \mathrm{D}u_\infty, _))^{-1}(\operatorname{sgn}(f_\infty)\mathrm{F}(u_\infty)) \quad \text{a.e. on } \Omega \setminus \{f_\infty = 0\}$$

as $p_j \to \infty$. (Note the notation $F(\cdot\,, u_\infty, \mathrm{D}u_\infty, _)$, in which "$\cdot$" is a placeholder for the variable x and "$_$" is a placeholder for the variable ξ.) However, by (1-20) we know that $\mathrm{A}{:}\mathrm{D}^2 u_p \rightharpoonup \mathrm{A}{:}\mathrm{D}^2 u_\infty$ in $\mathrm{L}^q(\Omega)$ as $p_j \to \infty$ for all $q \in (1, \infty)$. By the Vitali convergence theorem and the uniqueness of limits, it follows that

$$\mathrm{A}{:}\mathrm{D}^2 u_p \to \mathrm{A}{:}\mathrm{D}^2 u_\infty = (F(\cdot\,, u_\infty, \mathrm{D}u_\infty, _))^{-1}(\operatorname{sgn}(f_\infty)\mathrm{F}(u_\infty)) \tag{3-29}$$

in $\mathrm{L}^q(\Omega \setminus \{f_\infty = 0\})$ for all $q \in (1, \infty)$ and additionally a.e. on $\Omega \setminus \{f_\infty = 0\}$ as $p_j \to \infty$. The convergence in (3-29) has two consequences. On the one hand, by applying the function $F(\cdot\,, u_\infty, \mathrm{D}u_\infty, _)$ to (3-29), we have

$$F(\mathrm{J}^2 u_\infty) = \operatorname{sgn}(f_\infty)\mathrm{F}(u_\infty) \quad \text{a.e. on } \Omega \setminus \{f_\infty = 0\}. \tag{3-30}$$

The equality (3-30) implies that the pair (u_∞, f_∞) satisfies (1-8) a.e. on $\Omega \setminus \{f_\infty = 0\}$. Further, by (3-17) and (1-20), (3-29) implies

$$\left.\begin{aligned} \mathrm{K}_p &= \frac{\partial_\eta F(\mathrm{J}^2 u_p)}{\partial_\xi F(\mathrm{J}^2 u_p)} \to \frac{\partial_\eta F(\mathrm{J}^2 u_\infty)}{\partial_\xi F(\mathrm{J}^2 u_\infty)} \\ \mathrm{L}_p &= \frac{\partial_\mathrm{p} F(\mathrm{J}^2 u_p)}{\partial_\xi F(\mathrm{J}^2 u_p)} \to \frac{\partial_\mathrm{p} F(\mathrm{J}^2 u_\infty)}{\partial_\xi F(\mathrm{J}^2 u_\infty)} \end{aligned}\right\} \quad \text{a.e. on } \Omega \setminus \{f_\infty = 0\} \text{ as } p_j \to \infty. \tag{3-31}$$

Recalling that $\mathrm{K}_p \rightharpoonup \mathrm{K}_\infty$ in $\mathrm{L}^q(\Omega)$ and also $\mathrm{L}_p \rightharpoonup \mathrm{L}_\infty$ in $\mathrm{L}^q(\Omega; \mathbb{R}^n)$ as $p_j \to \infty$, for all $q \in (1, \infty)$, the Vitali convergence theorem yields that the modes of convergence in (3-31) are valid in $\mathrm{L}^q(\Omega \setminus \{f_\infty = 0\})$ for all $q \in (1, \infty)$. Further, by the uniqueness of limits, we infer that K_∞, L_∞ are characterised as

$$\mathrm{K}_\infty = \frac{\partial_\eta F(\mathrm{J}^2 u_\infty)}{\partial_\xi F(\mathrm{J}^2 u_\infty)}, \quad \mathrm{L}_\infty = \frac{\partial_\mathrm{p} F(\mathrm{J}^2 u_\infty)}{\partial_\xi F(\mathrm{J}^2 u_\infty)} \qquad \text{a.e. on } \Omega \setminus \{f_\infty = 0\}. \tag{3-32}$$

In view of (3-32) and (3-24), it follows that the PDE (1-9) is indeed satisfied by the pair (u_∞, f_∞). By Proposition 19 applied to (1-9), we have that $f_\infty \in \mathrm{W}^{1,q}_{\mathrm{loc}}(\Omega)$ for all $q \in (1, \infty)$, and $\{f_\infty = 0\}$ is a countably $(n-1)$-rectifiable subset of Ω. In particular, $\{f_\infty = 0\}$ is a Lebesgue nullset. We may further assume that $\mathrm{F}_\infty(u_\infty) > 0$. If instead $\mathrm{F}_\infty(u_\infty) = 0$, then (1-8) is trivially satisfied a.e. on Ω for any f_∞, whilst (1-9) is also satisfied by choosing any nontrivial solution $f_\infty \neq 0$, guaranteed to exist by Proposition 19. The proofs of Theorems 4 and 7 will therefore be completed once we prove that $f_\infty \not\equiv 0$. The remainder of the proof is therefore devoted to showing $f_\infty \not\equiv 0$ by using assumption (1-17). (This is the only point where the C^2-regularity assumption for $u_0|_{\partial\Omega}$ is utilised.) As a first step, we establish the estimate

$$F(\mathrm{J}^2 u_p)\partial_\xi F(\mathrm{J}^2 u_p)(\mathrm{A}{:}\mathrm{D}^2 u_p - \bar{\xi}(\,\cdot\,, u_p, \mathrm{D}u_p)) \geq \Big(\frac{c}{C_\infty}\Big)^2 (F(\mathrm{J}^2 u_p))^2 \tag{3-33}$$

a.e. on Ω for all sufficiently large $p \geq p_0$, where

$$C_\infty := C(1 + \|u_\infty\|_{\mathrm{W}^{1,\infty}(\Omega)}).$$

To establish (3-33), we use (1-7) and (1-16) to estimate

$$\begin{aligned} F(x, \eta, \xi, \mathrm{p})(\xi - \bar{\xi}(x, \eta, \mathrm{p})) &= \big[F(x, \eta, \mathrm{p}, \xi) - F(x, \eta, \mathrm{p}, \bar{\xi}(x, \eta, \mathrm{p}))\big](\xi - \bar{\xi}(x, \eta, \mathrm{p})) \\ &= \bigg(\int_{\bar{\xi}(x,\eta,\mathrm{p})}^{\xi} \partial_\xi F(x, \eta, \mathrm{p}, \cdot)\, \mathrm{d}\mathcal{L}^1\bigg)(\xi - \bar{\xi}(x, \eta, \mathrm{p})) \\ &\geq c|\xi - \bar{\xi}(x, \eta, \mathrm{p})|^2 \geq \frac{c}{C(|\eta| + |\mathrm{p}|)^2}(F(x, \eta, \mathrm{p}, \xi))^2, \end{aligned}$$

which, again by (1-7), yields

$$F(x, \eta, \mathrm{p}, \xi)\partial_\xi F(x, \eta, \mathrm{p}, \xi)(\xi - \bar{\xi}(x, \eta, \mathrm{p})) \geq \frac{c^2 (F(x, \eta, \mathrm{p}, \xi))^2}{C(|\eta| + |\mathrm{p}|)^2} \tag{3-34}$$

for a.e. $x \in \Omega$ and all $(\eta, \mathrm{p}, \xi) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$. Inequality (3-34) readily implies (3-33). Next, we establish the inequality

$$\begin{aligned} e_p\Big(\frac{c}{C_\infty}\Big)^2 \leq \fint_\Omega f_p\big[\mathrm{A}{:}\mathrm{D}^2 w - \bar{\xi}(\,\cdot\,, u_p, \mathrm{D}u_p)\big]\,\mathrm{d}\mathcal{L}^n + \fint_\Omega f_p\big[\mathrm{K}_p(w - u_p) - \mathrm{L}_p \cdot \mathrm{D}(w - u_p)\big]\,\mathrm{d}\mathcal{L}^n \\ + \|w - u_p\|_{\mathrm{L}^1(\Omega)}\|u_\infty - u_p\|_{\mathrm{L}^\infty(\Omega)} \end{aligned} \tag{3-35}$$

for any fixed $w \in \mathcal{W}^{2,\infty}_{u_0}(\Omega)$, where e_p is given by (1-22) and satisfies (3-21) as $j \to \infty$ along a subsequence $(p_j)_1^\infty$. With this aim, let $w \in \mathcal{W}^{2,\infty}_{u_0}(\Omega)$ be fixed. Since $w - u_p \in \mathrm{W}^{2,p}_0(\Omega)$, by testing this in the weak formulation of (3-19) and employing (3-33) and (1-23), we have

$$\int_\Omega \big[(\mathrm{K}_p f_p + (u_p - u_\infty))(w - u_p) - (\mathrm{L}_p f_p)\cdot \mathrm{D}(w - u_p)\big]\,\mathrm{d}\mathcal{L}^n + \int_\Omega f_p\big[\mathrm{A}{:}\mathrm{D}^2 w - \bar{\xi}(\,\cdot\,, u_p, \mathrm{D}u_p)\big]\,\mathrm{d}\mathcal{L}^n$$
$$= \int_\Omega f_p\big[\mathrm{A}{:}\mathrm{D}^2 u_p - \bar{\xi}(\,\cdot\,, u_p, \mathrm{D}u_p)\big]\,\mathrm{d}\mathcal{L}^n \geq e_p^{1-p}\Big(\frac{c}{C_\infty}\Big)^2 \int_\Omega (F(\mathrm{J}^2 u_p))^p\,\mathrm{d}\mathcal{L}^n.$$

The above inequality readily implies (3-35).

By noting that assumption (1-17) requires that $u_0|_{\partial\Omega}$ is C^2 on the closed set $\partial\Omega$, the extension theorem of Whitney [1934, Theorem I, p. 65], yields the existence of some $u_* \in \mathrm{C}^2(\mathbb{R}^n)$ such that $u_0 = u_*$ on $\partial\Omega$, and also $\mathrm{D}u_0 = \mathrm{D}u_*$ on $\partial\Omega$. For any $\varepsilon > 0$, we set $\Omega^\varepsilon := \Omega \cap \{\mathrm{dist}(\,\cdot\,, \partial\Omega) < \varepsilon\}$. Since $\|u_* - u_0\|_{\mathrm{C}^1(\Omega^\varepsilon)} \to 0$ as $\varepsilon \to 0$, by using a standard regularisation argument involving partitions of unity, we can construct a function $u_\varepsilon \in \mathrm{C}^2(\overline{\Omega})$ satisfying both that $u_\varepsilon \equiv u_*$ in an inner neighbourhood of the boundary and that u_ε equals the standard mollification of u_∞ (for small enough regularisation parameter) on the complement of such an inner neighbourhood. Then, by construction u_ε satisfies $u_\infty = u_\varepsilon$ on $\partial\Omega$ and also $\mathrm{D}u_\infty = \mathrm{D}u_\varepsilon$ on $\partial\Omega$, with the additional property $\|u_\infty - u_\varepsilon\|_{\mathrm{W}^{1,\infty}(\Omega)} \to 0$ as $\varepsilon \to 0$. By recalling that $\{\mathrm{K}_p f_p\}_{p>p_0}$ and $\{|\mathrm{L}_p f_p|\}_{p>p_0}$ are bounded in $\mathrm{L}^1(\Omega)$, the choice $w := u_\varepsilon$ in (3-35) yields

$$e_p\Big(\frac{c}{C_\infty}\Big)^2 \leq \int_\Omega f_p\big[\mathrm{A}{:}\mathrm{D}^2 u_\varepsilon - \bar{\xi}(\,\cdot\,, u_p, \mathrm{D}u_p)\big]\,\mathrm{d}\mathcal{L}^n$$
$$+ \|u_\varepsilon - u_p\|_{\mathrm{W}^{1,\infty}(\Omega)}\big[\|u_\infty - u_p\|_{\mathrm{L}^\infty(\Omega)} + \|\mathrm{K}_p f_p\|_{\mathrm{L}^1(\Omega)} + \|\mathrm{L}_p f_p\|_{\mathrm{L}^1(\Omega)}\big].$$

By letting $p \to \infty$ along a subsequence in the above inequality, we may use that $f_p \mathcal{L}^n \llcorner_\Omega \overset{*}{\rightharpoonup} f_\infty \mathcal{L}^n \llcorner_\Omega$ in $\mathcal{M}(\overline{\Omega}) \cong (\mathrm{C}(\overline{\Omega}))^*$ (recall Claim 20) and the fact that $\mathrm{A}:\mathrm{D}^2 u_\varepsilon \in \mathrm{C}(\overline{\Omega})$ and that $\bar{\xi}(\,\cdot\,, u_p, \mathrm{D}u_p) \to \bar{\xi}(\,\cdot\,, u_\infty, \mathrm{D}u_\infty)$ uniformly on $\overline{\Omega}$ as $p \to \infty$ along a subsequence to deduce the inequality

$$\mathrm{F}_\infty(u_\infty)\Big(\frac{c}{C_\infty}\Big)^2 - o(1) \leq \int_\Omega f_\infty\big[\mathrm{A}{:}\mathrm{D}^2 u_\varepsilon - \bar{\xi}(\,\cdot\,, u_\infty, \mathrm{D}u_\infty)\big]\,\mathrm{d}\mathcal{L}^n \quad \text{as } \varepsilon \to 0.$$

Since by assumption $\mathrm{F}_\infty(u_\infty) > 0$, if $\varepsilon > 0$ is chosen small enough, the left-hand side in the above inequality becomes positive, and this implies $f_\infty \not\equiv 0$. In conclusion, the proofs of Theorems 4 and 7 are complete. □

We now conclude by establishing Theorem 5.

Proof of Theorem 5. (I) Let $u_\infty \in \mathcal{W}^{2,\infty}_{u_0}(\Omega)$ be a narrow minimiser of the functional F_∞. By Definition 1, there exists $\rho > 0$ such that

$$\mathrm{F}_\infty(u_\infty) \leq \mathrm{F}_\infty(u_\infty + \psi), \quad \psi \in \mathbb{B}_\rho^{\mathrm{W}^{1,\infty}(\Omega)} \cap \mathcal{W}^{2,\infty}_0(\Omega).$$

By Theorem 4, it follows that there exists $f_\infty \in \mathrm{L}^1(\Omega) \setminus \{0\}$ such that the pair (u_∞, f_∞) satisfies the system of equations (1-8)–(1-9). By Theorem 3, it follows that u_∞ is actually a *strict* narrow minimiser,

therefore there exists $\rho' > 0$ such that

$$\mathrm{F}_\infty(u_\infty) < \mathrm{F}_\infty(u_\infty + \psi), \quad \psi \in (\mathbb{B}_{\rho'}^{\mathrm{W}^{1,\infty}(\Omega)} \cap \mathcal{W}_0^{2,\infty}(\Omega)) \setminus \{0\}.$$

Hence u_∞ is the unique narrow minimiser of the functional F_∞ (at the same energy level) within the admissible class

$$\mathbb{B}_{\min\{\rho,\rho'\}}^{\mathrm{W}^{1,\infty}(\Omega)}(u_\infty) \cap \mathcal{W}_{u_0}^{2,\infty}(\Omega).$$

(II) Let $u_\infty \in \mathcal{W}_{u_0}^{2,\infty}(\Omega)$ be a global minimiser of the functional F_∞ (Definition 1). By Theorem 4, it follows that there exists $f_\infty \in \mathrm{L}^1(\Omega) \setminus \{0\}$ such that the pair (u_∞, f_∞) satisfies the system of equations (1-8)–(1-9). By Theorem 3, it follows that u_∞ is a strict global minimiser (Definition 1). In conclusion, u_∞ is the unique global minimiser of the functional F_∞ in the space $\mathcal{W}_{u_0}^{2,\infty}(\Omega)$. □

Proof of Corollary 6. Let $\mathrm{J}^1 u(x) := (x, u(x), \mathrm{D}u(x))$ denote the first-order jet operator. Let us also set $\Omega^* := \Omega \setminus \{f_\infty = 0\}$. Since $\{f_\infty = 0\}$ is nowhere dense, Ω^* is an open dense set in Ω. By (1-8) and assumption (1-7), we have

$$\mathrm{A}{:}\mathrm{D}^2 u_\infty = F(\mathrm{J}^1 u_\infty, \cdot)^{-1}(\mathrm{F}_\infty(u_\infty)\,\mathrm{sgn}(f_\infty)) \quad \text{a.e. on } \Omega^*.$$

Since $u_\infty \in \mathrm{C}^{1,\alpha}(\Omega)$ for all $\alpha \in (0,1)$, the right-hand side is in $\mathrm{C}^{0,\alpha}(\Omega^*)$. By the Schauder interior estimates (see, e.g., [Gilbarg and Trudinger 1998, Chapter 6]), we have that $u_\infty \in \mathrm{C}^{2,\alpha}(\Omega^*)$. Hence, since F is by assumption C^1, the right-hand side is in $\mathrm{C}^{1,\alpha}(\Omega^*)$, which implies $u_\infty \in \mathrm{C}^{3,\alpha}(\Omega^*)$. The conclusion for general $k \geq 2$ is a consequence of a standard bootstrap argument and finite induction. Finally, by Proposition 19, the nodal set $\{f_\infty = 0\}$ is countably $(n-1)$-rectifiable. □

Proof of Corollary 8. Since any increasing $\Phi : \mathbb{R} \to \mathbb{R}$ which satisfies $\Phi(-\,\cdot) = -\Phi$ commutes with the absolute value and the (essential) supremum, we have

$$\overline{\mathrm{F}}_\infty(u) = \operatorname*{ess\,sup}_\Omega |\Phi(F(\mathrm{J}^2 u))| = \Phi\big(\operatorname*{ess\,sup}_\Omega |F(\mathrm{J}^2 u)|\big) = \Phi(\mathrm{F}_\infty(u)),$$

yielding that F_∞ and $\overline{\mathrm{F}}_\infty$ have the same sets of global/narrow/local minimisers. □

Acknowledgements

Katzourakis would like to thank Daniel Tataru and Alexander Logunov for their insights on the fulfilment of the unique continuation principle and the size of nodal sets of solutions to the equations arising in this work. Both authors are indebted to the anonymous referees whose constructive comments improved considerably the content and the quality of the exposition in this paper.

References

[Adams and Fournier 2003] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, 2nd ed., Pure and Applied Mathematics **140**, Elsevier, 2003. MR

[Ansini and Prinari 2015] N. Ansini and F. Prinari, "On the lower semicontinuity of supremal functional under differential constraints", *ESAIM Control Optim. Calc. Var.* **21**:4 (2015), 1053–1075. MR

[Aronsson 1965] G. Aronsson, "Minimization problems for the functional $\sup_x F(x, f(x), f'(x))$, I", *Ark. Mat.* **6** (1965), 33–53. MR

[Aronsson 1966] G. Aronsson, "Minimization problems for the functional $\sup_x F(x, f(x), f'(x))$, II", *Ark. Mat.* **6** (1966), 409–431. MR

[Aronsson 1968] G. Aronsson, "On the partial differential equation $u_x{}^2 u_{xx} + 2u_x u_y u_{xy} + u_y{}^2 u_{yy} = 0$", *Ark. Mat.* **7** (1968), 395–425. MR

[Aronsson 1969] G. Aronsson, "Minimization problems for the functional $\sup_x F(x, f(x), f'(x))$, III", *Ark. Mat.* **7** (1969), 509–512. MR

[Aronsson and Barron 2012] G. Aronsson and E. N. Barron, "L^∞ variational problems with running costs and constraints", *Appl. Math. Optim.* **65**:1 (2012), 53–90. MR

[Barron et al. 2001] E. N. Barron, R. R. Jensen, and C. Y. Wang, "Lower semicontinuity of L^∞ functionals", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **18**:4 (2001), 495–517. MR

[Bogachev et al. 2009] V. I. Bogachev, N. V. Krylov, and M. Röckner, "Elliptic and parabolic equations for measures", *Uspekhi Mat. Nauk* **64**:6(390) (2009), 5–116. In Russian; translated in *Russian Math. Surveys* **64** (2009), 973–1078. MR

[Brizzi and De Pascale 2023] C. Brizzi and L. De Pascale, "A property of absolute minimizers in L^∞ calculus of variations and of solutions of the Aronsson–Euler equation", *Adv. Differential Equations* **28**:3-4 (2023), 287–310. MR

[Bungert and Korolev 2022] L. Bungert and Y. Korolev, "Eigenvalue problems in L$^\infty$: optimality conditions, duality, and relations with optimal transport", *Commun. Am. Math. Soc.* **2** (2022), 345–373. MR

[Carano et al. 2025] S. Carano, N. Katzourakis, and R. Moser, "Existence, uniqueness and characterisation of vector-valued absolute minimisers for a second order L^∞-variational problem", preprint, 2025. arXiv 2504.04181

[Champion et al. 2004] T. Champion, L. De Pascale, and F. Prinari, "Γ-convergence and absolute minimizers for supremal functionals", *ESAIM Control Optim. Calc. Var.* **10**:1 (2004), 14–27. MR

[Cherkas 1970] B. M. Cherkas, "Compactness in L_∞ spaces", *Proc. Amer. Math. Soc.* **25** (1970), 347–350. MR

[Clark and Katzourakis 2025] E. Clark and N. Katzourakis, "Generalized second order vectorial ∞-eigenvalue problems", *Proc. Roy. Soc. Edinburgh Sect. A* **155**:6 (2025), 2150–2170. MR

[Clark et al. 2022] E. Clark, N. Katzourakis, and B. Muha, "Vectorial variational problems in L^∞ constrained by the Navier–Stokes equations", *Nonlinearity* **35**:1 (2022), 470–491. MR

[Conway 1990] J. B. Conway, *A course in functional analysis*, 2nd ed., Graduate Texts in Mathematics **96**, Springer, New York, 1990. MR

[Dacorogna 2008] B. Dacorogna, *Direct methods in the calculus of variations*, 2nd ed., Applied Mathematical Sciences **78**, Springer, New York, 2008. MR

[Danskin 1966] J. M. Danskin, "The theory of Max-Min, with applications", *SIAM J. Appl. Math.* **14** (1966), 641–664. MR

[Di Fazio 1996] G. Di Fazio, "L^p estimates for divergence form elliptic equations with discontinuous coefficients", *Boll. Un. Mat. Ital. A* (7) **10**:2 (1996), 409–420. MR

[Fonseca and Leoni 2007] I. Fonseca and G. Leoni, *Modern methods in the calculus of variations: L^p spaces*, Springer, New York, 2007. MR

[Gilbarg and Trudinger 1998] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Grundl. Math. Wissen. **224**, Springer, Berlin, 1998. MR

[Hardt and Simon 1989] R. Hardt and L. Simon, "Nodal sets for solutions of elliptic equations", *J. Differential Geom.* **30**:2 (1989), 505–522. MR

[Jensen et al. 2008] R. Jensen, C. Wang, and Y. Yu, "Uniqueness and nonuniqueness of viscosity solutions to Aronsson's equation", *Arch. Ration. Mech. Anal.* **190**:2 (2008), 347–370. MR

[Katzourakis 2015] N. Katzourakis, *An introduction to viscosity solutions for fully nonlinear PDE with applications to calculus of variations in L^∞*, Springer, Cham, 2015. MR

[Katzourakis 2017] N. Katzourakis, "Generalised solutions for fully nonlinear PDE systems and existence-uniqueness theorems", *J. Differential Equations* **263**:1 (2017), 641–686. MR

[Katzourakis 2022] N. Katzourakis, "Generalised vectorial ∞-eigenvalue nonlinear problems for L^∞ functionals", *Nonlinear Anal.* **219** (2022), art. id. 112806. MR

[Katzourakis and Moser 2019] N. Katzourakis and R. Moser, "Existence, uniqueness and structure of second order absolute minimisers", *Arch. Ration. Mech. Anal.* **231**:3 (2019), 1615–1634. MR

[Katzourakis and Moser 2023] N. Katzourakis and R. Moser, "Variational problems in L^∞ involving semilinear second order differential operators", *ESAIM Control Optim. Calc. Var.* **29** (2023), art. id. 76. MR

[Katzourakis and Parini 2017] N. Katzourakis and E. Parini, "The eigenvalue problem for the ∞-bilaplacian", *NoDEA Nonlinear Differential Equations Appl.* **24**:6 (2017), art. id. 68. MR

[Katzourakis and Pryer 2020] N. Katzourakis and T. Pryer, "Second-order L^∞ variational problems and the ∞-polylaplacian", *Adv. Calc. Var.* **13**:2 (2020), 115–140. MR

[Katzourakis and Vărvărucă 2018] N. Katzourakis and E. Vărvărucă, *An illustrative introduction to modern analysis*, CRC Press, Boca Raton, FL, 2018. MR

[Koch and Tataru 2001a] H. Koch and D. Tataru, "Carleman estimates and unique continuation for second-order elliptic equations with nonsmooth coefficients", *Comm. Pure Appl. Math.* **54**:3 (2001), 339–360. MR

[Koch and Tataru 2001b] H. Koch and D. Tataru, "Recent results on unique continuation for second order elliptic equations", pp. 73–84 in *Carleman estimates and applications to uniqueness and control theory* (Cortona, 1999), edited by F. Colombini and C. Zuily, Progr. Nonlinear Differential Equations Appl. **46**, Birkhäuser, Boston, MA, 2001. MR

[Kreisbeck and Zappale 2020] C. Kreisbeck and E. Zappale, "Lower semicontinuity and relaxation of nonlocal L^∞-functionals", *Calc. Var. Partial Differential Equations* **59**:4 (2020), art. id. 138. MR

[Kristensen and Taheri 2003] J. Kristensen and A. Taheri, "Partial regularity of strong local minimizers in the multi-dimensional calculus of variations", *Arch. Ration. Mech. Anal.* **170**:1 (2003), 63–89. MR

[Miao et al. 2017] Q. Miao, C. Wang, and Y. Zhou, "Uniqueness of absolute minimizers for L^∞-functionals involving Hamiltonians $H(x, p)$", *Arch. Ration. Mech. Anal.* **223**:1 (2017), 141–198. MR

[Moser 2022] R. Moser, "Structure and classification results for the ∞-elastica problem", *Amer. J. Math.* **144**:5 (2022), 1299–1329. MR

[Moser and Schwetlick 2012] R. Moser and H. Schwetlick, "Minimizers of a weighted maximum of the Gauss curvature", *Ann. Global Anal. Geom.* **41**:2 (2012), 199–207. MR

[Ostrovskii 2001] M. I. Ostrovskii, "Weak* sequential closures in Banach space theory and their applications", pp. 21–34 in *General topology in Banach spaces*, edited by T. Banakh, Nova Sci., Huntington, NY, 2001. MR

[Papamikos and Pryer 2019] G. Papamikos and T. Pryer, "A Lie symmetry analysis and explicit solutions of the two-dimensional ∞-polylaplacian", *Stud. Appl. Math.* **142**:1 (2019), 48–64. MR

[Peng et al. 2021] F. Peng, C. Wang, and Y. Zhou, "Regularity of absolute minimizers for continuous convex Hamiltonians", *J. Differential Equations* **274** (2021), 1115–1164. MR

[Prinari and Zappale 2020] F. Prinari and E. Zappale, "A relaxation result in the vectorial setting and power law approximation for supremal functionals", *J. Optim. Theory Appl.* **186**:2 (2020), 412–452. MR

[Ribeiro and Zappale 2014] A. M. Ribeiro and E. Zappale, "Existence of minimizers for nonlevel convex supremal functionals", *SIAM J. Control Optim.* **52**:5 (2014), 3341–3370. MR

[Rudin 1991] W. Rudin, *Functional analysis*, 2nd ed., McGraw-Hill, New York, 1991. MR

[Sakellaris 2017] Z. N. Sakellaris, "Minimization of scalar curvature in conformal geometry", *Ann. Global Anal. Geom.* **51**:1 (2017), 73–89. MR

[Toland 2020] J. Toland, *The dual of $L_\infty(X, \mathcal{L}, \lambda)$, finitely additive measures and weak convergence—a primer*, Springer, Cham, 2020. MR

[Whitney 1934] H. Whitney, "Analytic extensions of differentiable functions defined in closed sets", *Trans. Amer. Math. Soc.* **36**:1 (1934), 63–89. MR

[Yu 2006] Y. Yu, "L^∞ variational problems and Aronsson equations", *Arch. Ration. Mech. Anal.* **182**:1 (2006), 153–180. MR

Received 17 Jun 2024. Revised 16 Jul 2025. Accepted 19 Sep 2025.

NIKOS KATZOURAKIS: n.katzourakis@reading.ac.uk
Department of Mathematics and Statistics, University of Reading, Reading, United Kingdom

ROGER MOSER: r.moser@bath.ac.uk
Department of Mathematical Sciences, University of Bath, Bath, United Kingdom

ANALYSIS AND PDE
Vol. 19 (2026), No. 7, pp. 1295–1338
DOI: 10.2140/apde.2026.19.1295

AN $L^{3/2}$ SL_2 KAKEYA MAXIMAL INEQUALITY

JOHN GREEN, TERENCE L. J. HARRIS AND YUMENG OU

We prove a special case of the Kakeya maximal function conjecture in $\mathbb{R}^3$, with $C_\epsilon \delta^{-\epsilon}$ loss, when the centre lines of the tubes are SL_2 lines and the tubes satisfy a 2-dimensional ball condition (equivalent to the Wolff axioms in this case). We show that the exponent $p = \frac{3}{2}$ is sharp and that some loss (such as $C_\epsilon \delta^{-\epsilon}$) is necessary, even in the SL_2 case where the δ-tubes have δ-separated directions and the cardinality of the tube family is maximal ($\sim \delta^{-2}$).

The SL_2 Kakeya maximal inequality is deduced from an $L^{3/2}$ inequality for restricted families of projections onto planes. A related $L^{3/2-\epsilon}$ inequality is also derived for restricted projections onto lines, and an application is given to generic intersections of sets in $\mathbb{R}^3$ with "light rays" and "light planes".

1. Introduction

The aim of this work is to prove the following special case of the Kakeya maximal function conjecture:

Theorem 1.1. *Let $\mathbb{T}$ be a collection of SL_2 tubes in $\mathbb{R}^3$ of length 1 and radius δ, satisfying the condition*

$$|\{T \in \mathbb{T} : T \subseteq S\}| \leq (r/\delta)^2 \tag{1-1}$$

for any tube S of length 2 and radius r, and any $r \in [\delta, 1]$. Then for any $\epsilon > 0$,

$$\left\| \sum_{T \in \mathbb{T}} \chi_T \right\|^{3/2}_{L^{3/2}(B_3(0,1))} \leq C_\epsilon \delta^{-\epsilon} |\mathbb{T}| \delta^2. \tag{1-2}$$

In Theorem 1.1, a tube is an SL_2 tube if its centre line is an SL_2 line. A line $\ell \subseteq \mathbb{R}^3$ is called an SL_2 line if there exist $a, b, c, d \in \mathbb{R}$ with $ad - bc = 1$ such that

$$\ell = \{(a, b, 0) + \lambda(c, d, 1) : \lambda \in \mathbb{R}\},$$

or if there exists $t_0 \in \mathbb{R}$ and $(c, d) \in \mathbb{R}^2 \setminus \{0\}$ such that

$$\ell = \{(0, 0, t_0) + \lambda(c, d, 0) : \lambda \in \mathbb{R}\}.$$

In the SL_2 case on $B_3(0, 1)$, the Wolff axioms (see [Katz and Zahl 2019, Definition 1.1]) are equivalent [Katz et al. 2023, Remark 2] to the condition (1-1), which is why Theorem 1.1 is a special case of the Kakeya maximal function conjecture in $\mathbb{R}^3$ for families of tubes satisfying the Wolff axioms. We note that the Wolff axioms (and (1-1)) hold whenever the tubes in $\mathbb{T}$ have $\sim \delta$-separated directions.

Theorem 1.1 will be deduced from an $L^{3/2}$ inequality for restricted families of projections onto planes stated below, and the point-line duality principle from [Liu 2022; Fässler and Orponen 2023b].

MSC2020: 28A78, 28A80.
Keywords: Besicovitch set, Kakeya maximal function, restricted projections.

© 2026 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.

Theorem 1.2. *If $\alpha > 2$, then for any Borel measure μ on $B_3(0,1)$ with $c_\alpha(\mu) < \infty$,*

$$\int_0^{2\pi} \|\pi_{\theta\sharp}\mu\|_{L^{3/2}}^{3/2}\, d\theta \le C_\alpha \mu(\mathbb{R}^3) c_\alpha(\mu)^{1/2}.$$

Here π_θ is the orthogonal projection onto the plane $\frac{1}{\sqrt{2}}(\cos\theta, \sin\theta, 1)^\perp$, and

$$c_\alpha(\mu) = \sup_{r>0,\, x\in\mathbb{R}^3} \frac{\mu(B(x,r))}{r^\alpha}.$$

As a counterpart to Theorem 1.2, the following version for projections ρ_θ onto the line defined as the span of $\frac{1}{\sqrt{2}}(\cos\theta, \sin\theta, 1)$ will be proved with a similar method, but using a maximal estimate from [Pramanik et al. 2022] as a crucial input.

Theorem 1.3. *If $\alpha > 1$ and $1 \le p < \frac{3}{2}$, then for any Borel measure μ on $B_3(0,1)$ with $c_\alpha(\mu) < \infty$,*

$$\int_0^{2\pi} \|\rho_{\theta\sharp}\mu\|_{L^p}^{p}\, d\theta \le C_\alpha \mu(\mathbb{R}^3) c_\alpha(\mu)^{p-1}.$$

The proofs of Theorems 1.2 and 1.3, as well as Theorem 1.4 below, can easily be generalised to include C^2 curves $\gamma : [0,1] \to S^2$ with $\det(\gamma, \gamma', \gamma'')$ nonvanishing (i.e., nondegenerate curves), with π_θ and ρ_θ redefined analogously, but for simplicity they are only written here for the model curve $\gamma(\theta) = \frac{1}{\sqrt{2}}(\cos\theta, \sin\theta, 1)$.

In Section 2 it is shown via an explicit construction that there exist closed SL_2 Besicovitch sets of measure zero, using the same point-line duality as that used in [Liu 2022]. This is then used to show the corollary that a uniform L^p bound (i.e., without some loss such as $|\log\delta|$) is not possible in the SL_2 Kakeya maximal inequality for any nontrivial p. In Section 3, we show that the exponent $\frac{3}{2}$ in Theorems 1.1, 1.2, and 1.3 is sharp. This yields a counterexample to [Katz et al. 2023, Remark 1], where it was suggested that $p = 2$ might be possible in the SL_2 case of Theorem 1.1. The version of the SL_2 Kakeya maximal inequality in [Katz et al. 2023] appears slightly different to that used here, but the equivalence of the various forms is discussed in Section 1.2.

By a recent result of Mattila [2023a, Theorem 1.2], a consequence of Theorems 1.2 and 1.3 is the following result about generic intersections of sets in $\mathbb{R}^3$ with "light rays" and "light planes".

Theorem 1.4. *Let $A \subseteq \mathbb{R}^3$ be $\mathcal{H}^s$-measurable with $0 < \mathcal{H}^s(A) < \infty$. If $s > 1$ then for $(\mathcal{H}^s \times \mathcal{H}^1)$-a.e. $(x,\theta) \in A \times [0,2\pi)$,*

$$\dim\big(A \cap \rho_\theta^{-1}(\rho_\theta(x))\big) = s - 1,$$

and for a.e. $\theta \in [0,2\pi)$,

$$\mathcal{H}^1\big\{w \in \mathrm{span}(\gamma(\theta)) : \dim(A \cap \rho_\theta^{-1}(w)) = s-1\big\} > 0.$$

If $s > 2$ then for $(\mathcal{H}^s \times \mathcal{H}^1)$-a.e. $(x,\theta) \in A \times [0,2\pi)$,

$$\dim\big(A \cap \pi_\theta^{-1}(\pi_\theta(x))\big) = s - 2,$$

and for a.e. $\theta \in [0,2\pi)$,

$$\mathcal{H}^2\big\{w \in \gamma(\theta)^\perp : \dim(A \cap \pi_\theta^{-1}(w)) = s-2\big\} > 0.$$

Here and throughout the paper, dim denotes the Hausdorff dimension, and $\mathcal{H}^s$ refers to the s-dimensional Hausdorff measure in Euclidean space. Theorem 1.4 is a refinement of the Marstrand–Mattila slicing theorem in $\mathbb{R}^3$ (see Theorems 10.10 and 10.11 in [Mattila 1995]), which has the same statement as the above, but with lines and planes varying over all lines and planes through a point, rather than just "light rays" and "light planes". By Fubini's theorem and a simple scaling argument, Theorem 1.4 implies the standard Marstrand–Mattila slicing theorem in $\mathbb{R}^3$. Although much work has been done on restricted projection families since they were introduced in [Järvenpää et al. 2008; Fässler and Orponen 2014], the only "restricted intersection" type result in Euclidean space of which we are aware is from [Mattila 2023b], which is different from Theorem 1.4. There also exist results on the dimension of exceptional sets in intersection theorems (see, e.g., [Mattila 2023b] for a survey), but Theorem 1.4 seems to be the first refinement of Marstrand's intersection theorem for (curved) families of subspaces of dimension less than the full Grassmannian.

Note that it is necessary to work with nondegenerate curves γ. Indeed, consider an example of a set $A = [0,1]^2 \times C \subset \mathbb{R}^2 \times \mathbb{R}$, where C is a set with $0 < \mathcal{H}^\epsilon(C) < \infty$. Then $0 < \mathcal{H}^s(A) < \infty$, where $s = 2+\epsilon$. However, if one considers the curve $\gamma(\theta) = (\cos(\theta), \sin(\theta), 0)$, then for $\mathcal{H}^s$-a.e. $x \in A$, $A \cap \pi_\theta^{-1}(\pi_\theta(x))$ contains a line segment, hence has dimension $1 \neq s - 2 = \epsilon$.

Theorem 1.4 also implies that if $A \subseteq \mathbb{R}^3$ is Borel with $\dim A > 1$, then $\mathcal{H}^1(\rho_\theta(A)) > 0$ for a.e. $\theta \in [0, 2\pi)$, which was first proved in [Harris 2024]. This implication follows by applying a theorem of Davies [1952] to reduce to the case where A has positive finite s-dimensional $\mathcal{H}^s$-measure for some $s > 1$, and then observing that

$$\mathcal{H}^1(\rho_\theta(A)) = \mathcal{H}^1\big\{w \in \operatorname{span}(\gamma(\theta)) : A \cap \rho_\theta^{-1}(w) \neq \varnothing\big\},$$

and that clearly the nonempty condition above is weaker than the dimension $s-1$ condition in Theorem 1.4. Similarly, Theorem 1.4 implies that if $A \subseteq \mathbb{R}^3$ is Borel with $\dim A > 2$, then $\mathcal{H}^2(\pi_\theta(A)) > 0$ for a.e. $\theta \in [0, 2\pi)$, which was first proved in [Gan et al. 2022]. These length and area statements in turn imply the Hausdorff dimension versions $\dim(\rho_\theta(A)) = \min\{1, \dim A\}$ and $\dim(\pi_\theta(A)) = \min\{2, \dim A\}$ for a.e. $\theta \in [0, 2\pi)$; Fässler and Orponen proved that for general families of Euclidean projections, the dimension exponents in projection theorems can always be lowered by randomly adding points to lower dimensional sets to obtain higher dimensional sets (see, e.g., [Fässler and Orponen 2023b] for this type of argument), though this method probably loses information about the exceptional set. The a.e. equality $\dim(\rho_\theta(A)) = \min\{1, \dim A\}$ was first obtained by Käenmäki, Orponen and Venieri [Käenmäki et al. 2025], and then for C^2 curves by Pramanik, Yang, and Zahl [Pramanik et al. 2022] and Gan, Guth, and Maldague [Gan et al. 2024] (and all of [Käenmäki et al. 2025; Pramanik et al. 2022; Gan et al. 2024] contain better information about the exceptional set than what follows from Theorem 1.4). The Hausdorff dimension theorem for the projections π_θ was obtained in [Gan et al. 2022] (and by a different method in [Katz et al. 2023]). The a.e. equalities $\dim(\rho_\theta(A)) = \min\{1, \dim A\}$ and $\dim(\pi_\theta(A)) = \min\{2, \dim A\}$ can also be obtained as corollaries of Theorem 1.3 and 1.2 directly (by Hölder's inequality and the definition of Hausdorff dimension).

L^p estimates for projected measures have recently received quite some attention and have proven to be useful in the study of many important problems in geometric measure theory. For instance, an L^p

estimate for radial projections in $\mathbb{R}^d$ was obtained by Orponen [2019], which was used as an essential tool in the proof of the $\frac{5}{4}$ threshold for the two dimensional Falconer distance set problem in [Guth et al. 2020]. In the case of orthogonal projections, an L^p estimate for the orthogonal projected measures over the full Grassmannian was obtained by Dąbrowski, Orponen, and Villa [Dąbrowski et al. 2022], where they considered the case of general orthogonal projections from $\mathbb{R}^d$ to n dimensional subspaces, $1 \leq n < d$. Their result is known to be sharp when the dimension of the measure is at least $d-1$, and implies improved Furstenberg set estimates and discretized sum-product estimates. (Very recently, a similar L^p estimate with more general measures on the Grassmannian was derived by Li and Liu [2026].) Our L^p estimates extend this line of investigation of regularity of projected measures and seem to be the first of their kind concerning curved families of subspaces of dimension less than the full Grassmannian.

The SL_2 example related to the Kakeya problem first appeared in Katz and Zahl's $\frac{5}{2}+\epsilon$ lower bound for the Hausdorff dimension of Besicovitch sets in $\mathbb{R}^3$ [2019], where the proof needed to negotiate a hypothetical SL_2 "almost-counterexample". Later, Wang and Zahl [2026] conjectured that SL_2 Besicovitch sets in $\mathbb{R}^3$ have Hausdorff dimension 3 (as a special case of the Kakeya conjecture in $\mathbb{R}^3$). This was proved by Fässler and Orponen [2023a], and also by Katz, Wu, and Zahl [Katz et al. 2023]. By a standard argument (see, e.g., [Wolff 2003, Proposition 10.2]), Theorem 1.1 also implies that SL_2 Besicovitch sets have Hausdorff dimension 3. A type of SL_2 Kakeya maximal inequality first appeared in [Katz et al. 2023]. The inequality there implies that SL_2 Besicovitch sets have Hausdorff dimension 3, but does not imply the SL_2 maximal inequality in (1-2) for any $p > 1$ $\left(\text{in place of } \frac{3}{2}\right)$. By applying Heisenberg dilations $(x, y, z) \mapsto (\lambda x, \lambda y, \lambda^2 z)$, which preserve SL_2 lines, Theorem 1.1 still holds if $B_3(0, 1)$ is replaced by $B_3(0, R)$ for any $R \geq 1$ in the left-hand side of (1-2), if the constant in the right-hand side of (1-2) is allowed to depend on R. By then running the argument from [Orponen 2017] with the SL_2 Kakeya maximal function (see Section 1.2) instead of the general Kakeya maximal function, Theorem 1.1 also implies that SL_2 Besicovitch sets in $\mathbb{R}^3$ have *conformal* dimension 3. Conformal dimension is a notion of dimension which is less than or equal to the Hausdorff dimension, and for general sets may be smaller; see [Orponen 2017] for the precise definition. Unlike the Hausdorff dimension case, the argument from [Orponen 2017] relies crucially on the value of p in the Kakeya maximal inequality, and a value of $p < \frac{3}{2}$ would give a nonsharp result for the conformal dimension.

Finally, see [Mattila 2019, Sections 6.7 and 6.8] for a general discussion of connections between restricted projection problems and Kakeya-type problems. We also mention that a related maximal function to the SL_2 Kakeya maximal function was considered in [Fässler et al. 2025], where the optimal $L^{3/2}$ inequality was obtained. This maximal function was a function of the angle $\theta \in [0, \pi)$ of the tube, rather than the direction. The tubes considered were also Heisenberg tubes rather than Euclidean tubes, and the problem was related to the projections ρ_θ rather than π_θ.

1.1. ***History and novelty of the method.*** This subsection outlines a brief history of the "good-bad" decomposition of a measure in the context of this problem, as this may give some motivation for the proofs of Theorems 1.2 and 1.3. The decomposition of a measure into "good" and "bad" parts was first developed and applied to the planar distance set problem in [Guth et al. 2020]. This decomposition was modified in [Harris 2022] to work for restricted projections onto planes in $\mathbb{R}^3$. The argument from

[Harris 2022] contained significant loss in the refined decoupling step, arising from defining the bad part of the measure with tubes rather than planks, which are the natural objects to consider in wave packet decompositions adapted to the cone. The reason for using tubes in [Harris 2022] was partly that the argument used a lemma from Orponen and Venieri's previous work [2020] on the restricted projection problem as a black box to handle the "bad" part, and this lemma was formulated in terms of tubes. In [Harris 2023b], a bootstrapping argument for handling the "bad" part was developed, which removed the dependence on the lemma from [Orponen and Venieri 2020], but it was not yet realised in [Harris 2023b] that "tubes" could be replaced with "planks". This improvement was made in the second part of [Gan et al. 2022], and the possibility of the improvement was suggested by the first part of [Gan et al. 2022], which developed a different and independent approach to this problem via the "high-low method". In the second part of [Gan et al. 2022], it was shown that pushforwards under π_θ of measures of dimension greater than 2 are almost surely in L^1, and the method used to control the L^1 norm of the "bad" part of the measure was an incidence bound proved via the "high-low method". In this work, for projections onto planes we use an entirely recursive approach to bound the $L^{3/2}$ norm of the "bad" part, which allows us to make efficient use of L^2 bounds on the "good" part. The proof is mostly self-contained (in particular, the proof does not use the "high-low method"), and the only nontrivial result from [Gan et al. 2022] that we use is the refined decoupling theorem. Note however that we need only L^4 decoupling for the cone to prove Theorem 1.2, which has a much shorter[1] proof than L^6 decoupling.

The novelty in the proof of Theorem 1.2 is a "self-improving" argument for the exponents, where in practice we need to prove Theorem 1.2 for an increasing sequence of exponents $\{p_n\}_{n=0}^{\infty}$, where $p_0 = 1$, $p_1 = \frac{4}{3}$ and $p_\infty = \frac{3}{2}$. For each fixed exponent, we use a separate bootstrapping argument to prove a weaker version of Theorem 1.2 over a union of "bad" rectangles in each plane, where the meaning of "bad" is gradually weakened until we integrate over the whole planes. At each exponent, a crucial input is a bound on the measure of the unions of "bad" rectangles, which we get by applying the projection inequality for the previous exponent p_{n-1}. Because this projection inequality is better for each n than the previous n, the best exponent we can obtain in Theorem 1.2 improves at each step and we calculate that it converges to $\frac{3}{2}$.

The "good-bad" decomposition was modified to work for projections onto lines in [Harris 2024], where it was shown that the projections of measures of dimensions greater than 1 are almost surely in L^1. By modifying the recursive argument from [Harris 2024] to get an L^p bound as in the planar case, the best possible exponent via a purely recursive argument seems to be $\frac{4}{3}$, which is not the sharp exponent in Theorem 1.3. To get the exponent $\frac{3}{2} - \epsilon$, we need to use an $L^{3/2}$ Kakeya bound for curves satisfying a nonconcentration condition of Pramanik, Yang, and Zahl [Pramanik et al. 2022], as a black box, and therefore our only substantial contribution to Theorem 1.3 is to combine this with L^6 decoupling and the $\alpha > 1$ assumption in Theorem 1.3 to get a uniform bound as in Theorem 1.3.

We remark that this work supersedes the results from [Harris 2023a]. For completeness, and since the work from [Harris 2023a] will not be submitted for publication, some of the work from [Harris 2023a] is

[1]We also note that the refined decoupling theorem for the standard light cone has a simpler proof than in [Gan et al. 2022]; because Lorentz rescalings preserve the light cone exactly, the technicalities introduced in [Gan et al. 2022] to track how the curvature of the cone varies under rescaling are not necessary.

reproduced here (especially in Sections 2 and 3). We also want to point out that, even though the L^p bounds in analogues of Theorems 1.2 and 1.3 in [Harris 2023a] are weaker (at only $p = \frac{4}{3}$ and $p = \frac{6}{5}$, respectively), they are sufficient to derive the same restricted intersection results (with light rays and light planes) as in Theorem 1.4.

1.2. ***Different forms of the*** $\mathbf{SL_2}$ ***Kakeya maximal inequality.*** Given a Borel function f on $\mathbb{R}^3$, let $f^*_{\delta,\mathrm{SL}_2} : S^2 \setminus \{(0,0,\pm 1)\} \to \mathbb{R}$ be the SL_2 Kakeya maximal function at scale δ, which is defined in the same way as for the standard Kakeya maximal function except that the centre lines $\ell(T)$ of the δ-tubes T are required to be SL_2 lines:

$$f^*_{\delta,\mathrm{SL}_2}(e) = \sup_{\substack{\mathrm{dir}(T)=e \\ T=\delta\times 1\text{-tube} \\ \ell(T)\in \mathrm{SL}_2}} \frac{1}{m(T)} \int_T |f|,$$

where $m(T)$ is the Lebesgue measure of T. The terms "SL_2" and "horizontal" will be used interchangeably, though strictly speaking they are slightly different.

If $p^* \in (1,\infty)$ is fixed and p^* is the Hölder conjugate to p, then the inequality

$$\|f^*_{\delta,\mathrm{SL}_2}\|_{L^{p^*}(S^2)} \le C_\epsilon \delta^{-\epsilon} \|f\|_{p^*} \tag{1-3}$$

(for any $\epsilon > 0$) is equivalent (by the standard duality argument; see, e.g., [Wolff 2003, Lemma 10.4]) to the inequality

$$\left\| \sum_{T\in\mathbb{T}} \chi_T \right\|_p^p \le C_\epsilon \delta^{-\epsilon} |\mathbb{T}| \delta^2 \tag{1-4}$$

whenever $\mathbb{T}$ is a family of SL_2 δ-tubes of length 1 with δ-separated directions. It is known that (1-4) is equivalent to

$$\mathcal{H}^3\left(\bigcup_{T\in\mathbb{T}} Y(T) \right) \ge C_\epsilon^{-1} \delta^\epsilon \lambda^{p^*} \delta^2 |\mathbb{T}| \tag{1-5}$$

whenever $\mathbb{T}$ is a family of SL_2 δ-tubes of length 1 with δ-separated directions, $\delta \le \lambda \le 1$, and Y is a λ-shading of $\mathbb{T}$ (i.e., a union of δ-cubes such that at least a fraction λ of each tube in $\mathbb{T}$ is covered by the cubes in Y, meaning that $\mathcal{H}^3(Y(T)) \ge \lambda \mathcal{H}^3(T)$ for any $T \in \mathbb{T}$, where $Y(T) = Y \cap T$). The inequality from [Katz et al. 2023] states that (1-4) holds with $p^* = M$ depending on ϵ (and possibly tending to ∞ as $\epsilon \to 0$). The proof of the equivalence of (1-4) and (1-5) is similar to that for the standard Kakeya maximal function, but will be summarised here for convenience. That (1-4) implies (1-5) follows from Hölder's inequality. The reverse implication follows by pigeonholing a set Y of δ-cubes on which the value of $\sum_{T\in\mathbb{T}} \chi_T$ is approximately constant in order to reverse Hölder's inequality, then by pigeonholing again to a subset $\mathbb{T}' \subseteq \mathbb{T}$ (of potentially smaller cardinality) of tubes containing an approximately equal fraction λ of the cubes from Y in each tube, and then by applying (1-5) to this subset.

An a priori weaker inequality than (1-4) is

$$\left\| \sum_{T\in\mathbb{T}} \chi_T \right\|_{L^p(B_3(0,1))}^p \le C_\epsilon \delta^{-\epsilon} \tag{1-6}$$

whenever $\mathbb{T}$ is a family of SL_2 δ-tubes of length 1 with δ-separated directions. Clearly (1-4) implies (1-6), and the implication can be reversed for the standard Kakeya maximal function by rotational averaging (see, e.g., [Mattila 2015, Proposition 22.7]), but rotations in $O(3)$ do not necessarily preserve SL_2 lines, so it does not seem obvious that (1-6) implies (1-4) in the SL_2 case. In Section 3 it is shown that (1-6) cannot hold in the SL_2 case for all $\epsilon > 0$ unless $p \le \frac{3}{2}$ (and therefore neither can (1-5), (1-4) or (1-3)).

2. SL_2 Besicovitch sets of measure zero

The construction of a measure zero SL_2 Besicovitch set in this section is based on Kahane's construction of a measure zero Besicovitch set in the plane [1969]; for an exposition see [Stein and Shakarchi 2005, Chapter 7] or [Mattila 2015, Chapter 11].

Given $(x_1, y_1, t_1), (x_2, y_2, t_2) \in \mathbb{R}^3$, let

$$(x_1, y_1, t_1) * (x_2, y_2, t_2) = \big(x_1 + x_2, y_1 + y_2, t_1 + t_2 + \tfrac{1}{2}(x_1 y_2 - x_2 y_1)\big)$$

be the Heisenberg product. A line $\ell \subseteq \mathbb{R}^3$ is called horizontal if there exists $\theta \in [0, \pi)$ and $p \in \mathbb{R}^3$ such that

$$\ell = p * \mathbb{V}_\theta,$$

where $\mathbb{V}_\theta = \{(\lambda e^{i\theta}, 0) : \lambda \in \mathbb{R}\}$, and we identify $\lambda e^{i\theta} \in \mathbb{C}$ with $(\lambda \cos\theta, \lambda \sin\theta) \in \mathbb{R}^2$. Horizontal line segments and SL_2 line segments are line segments which are subsets of horizontal and SL_2 lines, respectively. A set $K \subseteq \mathbb{R}^3$ will be called a horizontal Besicovitch set if K contains a horizontal unit line segment in every direction in $S^2 \setminus \{(0, 0, \pm 1)\}$. Similarly, K will be called an SL_2 Besicovitch set if K contains an SL_2 unit line segment in every direction in $S^2 \setminus \{(0, 0, \pm 1)\}$. If $F(x, y, z) = (x, y, 2z)$, then F sends horizontal lines to SL_2 lines, and $\ell \mapsto F(\ell)$ is a bijection between horizontal and SL_2 lines with inverse $\ell \mapsto F^{-1}(\ell)$, where $F^{-1}(x, y, z) = (x, y, z/2)$. For any horizontal line segment I, the length of $F(I)$ is comparable to the length of I, so SL_2 lines and horizontal lines are more or less the same.

Every horizontal line can be written in the form $p * \mathbb{V}_\theta$ with (p_1, p_2) in the span of $ie^{i\theta}$, where $p = (p_1, p_2, p_3)$. If $R > 0$ and if such a horizontal line intersects $B(0, R)$, then $|(p_1, p_2)| \le R$ and the direction of $p * \mathbb{V}_\theta$ has distance at least $2/\sqrt{R^2 + 4}$ from $(0, 0, \pm 1)$. It follows that a bounded set cannot contain a horizontal unit line segment in every direction in $S^2 \setminus \{0, 0, \pm 1\}$ (and similarly for SL_2 line segments). For this reason, the compact sets of horizontal and SL_2 line segments in Theorem 2.2 below contain only unit line segments in every direction in $S^2 \setminus \mathcal{N}_\epsilon(\{(0, 0, \pm 1)\})$ for a given $\epsilon > 0$, and to get all directions in $S^2 \setminus \{(0, 0, \pm 1)\}$ it is necessary to replace "compact" with "closed".

The following (straightforward) lemma is used to make the proof of Theorem 2.2 below constructive, though it is possible to give a nonconstructive proof of Theorem 2.2 below by substituting the application of Lemma 2.1 in the proof of Theorem 2.2 with an averaging/Fubini/scaling argument.

Lemma 2.1. *If $E \subseteq (-\pi/2, \pi/2)$ is Lebesgue measurable with $\mathcal{H}^1(E) = 0$ and if F is the union of great circles defined by*

$$F = \{(\lambda e^{i\theta}, \mu) \in S^2 : (\lambda, \mu) \in S^1,\ \theta \in E\},$$

then $\mathcal{H}^1(F \cap C) = 0$ for any circle $C \subseteq S^2$ with the property that either $(0, 0, 1) \notin C$ or $(0, 0, -1) \notin C$.

Proof. Let E, C be given. By stereographic projection on the Riemann sphere (which maps circles to circles or lines) it suffices to show that if $\widetilde{C}$ is either a circle in $\mathbb{R}^2$ or a line in $\mathbb{R}^2$ which does not pass through the origin, then $\mathcal{H}^1(\widetilde{F} \cap \widetilde{C}) = 0$, where

$$\widetilde{F} = \{\lambda e^{i\theta} : \lambda \in \mathbb{R},\ \theta \in E\}.$$

To show this, it suffices to show that for any $(x_0, y_0) \in \widetilde{C} \setminus \{0\}$, there exists an $\epsilon > 0$ such that $\mathcal{H}^1(\widetilde{F} \cap \widetilde{C} \cap B((x_0, y_0), \epsilon)) = 0$. By scaling and rotation it may be assumed that $(x_0, y_0) = (1, 0)$. If $\widetilde{C}$ is a circle, then let (a, b) and $r > 0$ be such that $\widetilde{C}$ is parametrised by $\gamma(t) = (\gamma_1(t), \gamma_2(t)) = (a + r\cos t, b + r\sin t)$, where $\gamma : [0, 2\pi) \to \widetilde{C}$ satisfies $\gamma(t_0) = (1, 0)$ for some $t_0 \in [0, 2\pi)$. Otherwise let $\gamma(t) = (1, 0) + t(v_1, v_2)$ be the parametrisation of $\widetilde{C}$, where $v_2 \neq 0$ and $t \in \mathbb{R}$, and define $t_0 = 0$. It suffices to show that for $\delta > 0$ sufficiently small,

$$\mathcal{H}^1\big\{t \in (t_0 - \delta, t_0 + \delta) : \arctan(\gamma_2(t)/\gamma_1(t)) \in E\big\} = 0.$$

Equivalently, if $G : (t_0 - \delta, t_0 + \delta) \to (-\pi/2, \pi/2)$ is defined by $G(t) = \arctan(\gamma_2(t)/\gamma_1(t))$, then $(G_\sharp \mathcal{H}^1)(E) = 0$ for $\delta > 0$ sufficiently small. Since $\mathcal{H}^1(E) = 0$, it is enough to show that $G_\sharp \mathcal{H}^1 \ll \mathcal{H}^1$ for $\delta > 0$ sufficiently small. If $\widetilde{C}$ is a circle and $(a, b) = 0$, then $G'(t) = \frac{d}{dt} \arctan(\gamma_2(t)/\gamma_1(t)) = 1$ for all t in a neighbourhood of $t = t_0 = 0$, and $G_\sharp \mathcal{H}^1 \ll \mathcal{H}^1$ follows. If $\widetilde{C}$ is a line, then

$$G'(t) = \frac{v_2}{|\gamma(t)|^2} \neq 0$$

for all t in a neighbourhood of $t_0 = 0$, and again $G_\sharp \mathcal{H}^1 \ll \mathcal{H}^1$ follows. The remaining case is where $\widetilde{C}$ is a circle and $(a, b) \neq 0$. Let $H(t) = \det(\gamma(t), \gamma'(t))$. Then

$$G'(t) = \frac{H(t)}{|\gamma(t)|^2}$$

for all t in a neighbourhood of 0, so if $H(t)$ is nonvanishing at $t = t_0$, it follows that $G_\sharp \mathcal{H}^1 \ll \mathcal{H}^1$ for $\delta > 0$ sufficiently small. If $H(t_0) = 0$, then $H(t) \neq 0$ for all $t \in (t_0 - \delta, t_0 + \delta) \setminus \{t_0\}$ for some $\delta > 0$ sufficiently small, which follows from the identity $|H'(t)|^2 + |H''(t)|^2 = r^2(a^2 + b^2)$. Therefore $G_\sharp \mathcal{H}^1 \ll \mathcal{H}^1$ for $\delta > 0$ sufficiently small. This finishes the proof. □

Theorem 2.2. (1) *For any $\epsilon > 0$, there exists a compact set of measure zero containing a horizontal unit line segment in every direction in $S^2 \setminus \mathcal{N}_\epsilon(\{(0, 0, \pm 1)\})$.*

(2) *There exists a closed set of measure zero containing a horizontal unit line segment in every direction in $S^2 \setminus \{(0, 0, \pm 1)\}$.*

(3) *There exists a closed set of measure zero containing a horizontal line in every direction in*

$$S^2 \setminus \{(0, 0, \pm 1)\}.$$

The above claims all hold if "a horizontal" is replaced by "an SL_2*".*

Proof. Let A be any compact subset of $\mathbb{R}^3$ such that $0 < \mathcal{H}^2(A) < \infty$, such that the projection of A down to the (x, y)-plane contains a nonempty open disc, and such that for any nongreat circle $C \subseteq S^2$,

$\mathcal{H}^2(\pi_{v^\perp}(A)) = 0$ for $\mathcal{H}^1$-a.e. $v \in C$, where $\pi_{v^\perp}$ denotes projection to the orthogonal complement of v. Such a set can be constructed as follows. Let $\mathcal{C}$ be the Cantor set in $\mathbb{R}$ of dimension $\frac{1}{2}$ obtained by starting with the unit interval $[0, 1]$ and removing the open middle half interval at each step. It is a standard result that

$$\{2x + y : x, y \in \mathcal{C}\} = [0, 3],$$

as can be seen, e.g., by characterising $\mathcal{C}$ as those elements of $[0, 1]$ of the form $\sum_{j=1}^{\infty} \varepsilon_j/4^j$ with $\varepsilon_j \in \{0, 3\}$ for all j. It follows that if $B = \mathcal{C} \times \mathcal{C} \times [0, 1]$, then the projection of B onto the plane spanned by $\frac{1}{\sqrt{5}}(2, 1, 0)$ and $(0, 0, 1)$ contains a nonempty open disc. Moreover, if $v \in S^2$ is of the form $(\lambda_1 e^{i\theta}, \lambda_2)$, then

$$\pi_{v^\perp}(B) \subseteq \pi_{(ie^{i\theta},0)}(\mathcal{C} \times \mathcal{C} \times \{0\}) + \pi_{(\lambda_2 e^{i\theta},-\lambda_1)}(\mathbb{R}^3).$$

Hence, by Fubini's theorem,

$$\big\{v \in S^2 : \mathcal{H}^2(\pi_{v^\perp}(B)) > 0\big\} \subseteq F := \big\{(\lambda_1 e^{i\theta}, \lambda_2) : \theta \in E,\ (\lambda_1, \lambda_2) \in S^1\big\},$$

where

$$E = \big\{\theta \in [0, \pi) : \mathcal{H}^1(\pi_{ie^{i\theta}}(\mathcal{C} \times \mathcal{C})) > 0\big\}.$$

But $\mathcal{H}^1(E) = 0$ (see, e.g., [Mattila 2015, Theorem 10.1]), and by Lemma 2.1 this implies that $\mathcal{H}^1(F \cap C) = 0$ for any nongreat circle C. It follows that for any nongreat circle C, $\mathcal{H}^2(\pi_{v^\perp}(B)) = 0$ for $\mathcal{H}^1$-a.e. $v \in C$. Since this property is rotation-invariant, replacing B by a rotation of B which sends the plane spanned by $\frac{1}{\sqrt{5}}(2, 1, 0)$ and $(0, 0, 1)$ to the (x, y)-plane in $\mathbb{R}^3$ yields the required set A.

For the first claim in the theorem statement, let

$$K = K(A) = \bigcup_{(a,b,c)\in A} \left\{\big(as + b, s, c + \tfrac{1}{2}bs\big) : |s| \le \frac{1}{\sqrt{4 + 4a^2 + b^2}}\right\}.$$

Then, by Fubini's theorem,

$$\mathcal{H}^3(K) \le \int_{-1}^{1} \mathcal{H}^2\big\{(x, y) \in \mathbb{R}^2 : (x, y) = \big(as + b, c + \tfrac{1}{2}bs\big) : (a, b, c) \in A\big\}\, ds.$$

For each $s \in (-1, 1)$, the set

$$\big\{(x, y) \in \mathbb{R}^2 : (x, y) = \big(as + b, c + \tfrac{1}{2}bs\big) : (a, b, c) \in A\big\}$$

can be written as

$$\big\{(\langle p, (s, 1, 0)\rangle, \langle p, (0, s/2, 1)\rangle) \in \mathbb{R}^2 : p \in A\big\}.$$

The vectors $(s, 1, 0)$ and $(0, s/2, 1)$ are both orthogonal to $(1, -s, s^2/2)$. The vectors $(s, 1, 0)$ and $(0, s/2, 1)$ are not necessarily orthogonal to each other, but by following the Gram–Schmidt process,[2] the above can be written as

$$A_s\left\{\left(\langle p, (s, 1, 0)\rangle, \left\langle p, (0, s/2, 1) - \frac{s}{2(1 + s^2)}(s, 1, 0)\right\rangle\right) \in \mathbb{R}^2 : p \in A\right\},$$

[2]The same observation was made in [Fässler and Orponen 2023a].

where $A_s : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear map

$$A_s(x, y) = \left(x, y + \frac{sx}{2(1+s^2)}\right).$$

Since $\det A_s = 1$,

$$\mathcal{H}^2\big\{\big(\langle p, (s, 1, 0)\rangle, \langle p, (0, s/2, 1)\rangle\big) \in \mathbb{R}^2 : p \in A\big\}$$
$$= \mathcal{H}^2\left\{\left(\langle p, (s, 1, 0)\rangle, \left\langle p, (0, s/2, 1) - \frac{s}{2(1+s^2)}(s, 1, 0)\right\rangle\right) \in \mathbb{R}^2 : p \in A\right\} \sim \mathcal{H}^2(\pi_{(1,-s,s^2/2)^\perp}(A)).$$

The curve $(1, -s, s^2/2)$ is a parabola inside the cone $\eta_2^2 = 2\eta_1\eta_3$, and this cone is a clockwise rotation of the cone $\xi_3^2 = \xi_1^2 + \xi_2^2$ by $\pi/4$ in the (ξ_1, ξ_3)-plane. It follows that the normalised curve $(1/\sqrt{1+s^2+s^4/4})(1, -s, s^2/2)$ lies in the intersection of the cone $\eta_2^2 = 2\eta_1\eta_3$ with the sphere S^2, and this intersection is a nongreat circle in S^2. Hence $\mathcal{H}^2(\pi_{(1,-s,s^2/2)^\perp}(A)) = 0$ for a.e. $s \in \mathbb{R}$, which yields $\mathcal{H}^3(K) = 0$. Since the projection of A down to the (x, y)-plane contains an open disc $U \subseteq \mathbb{R}^2$, for all $(a, b) \in U$, the set K contains a horizontal unit line segment with direction $1/(a^2+1+(b/2)^2)(a, 1, b/2)$.

Any translate $A_{(x_0,y_0)} := A + (x_0, y_0, 0)$ of A with $(x_0, y_0) \in \mathbb{R}^2$ has the property that the projection of $A_{(x_0,y_0)}$ to $\mathbb{R}^2 \times \{0\}$ contains the open disc $U_{(x_0,y_0)} := U + (x_0, y_0)$, and also has the property that for any nongreat circle C, the projection of $A_{(x_0,y_0)}$ onto $v^\perp$ has $\mathcal{H}^2$-measure zero for $\mathcal{H}^1$-a.e. $v \in C$. Fix $(a_0, b_0) \in U$. For any $v \in S^2 \setminus \mathbb{R} \times \{0\} \times \mathbb{R}$, choose $(a_v, b_v) \in \mathbb{R}^2$ such that $(a_v + a_0, 1, (b_v + b_0)/2)$ is parallel to v. Then $K(A_{(a_v,b_v)})$ contains a horizontal unit line segment with direction $1/(a^2+1+(b/2)^2)(a, 1, b/2)$ for all $(a, b) \in U + (a_v, b_v)$. This shows that for any $v \in S^2 \setminus \mathbb{R} \times \{0\} \times \mathbb{R}$, there is an open neighbourhood $U_v \subseteq S^2$ of v and a measure zero compact set $K(A_{(a_v,b_v)})$ containing a horizontal unit line segment with every direction in U_v. Since the collection of horizontal lines is invariant under rotations of the (x, y)-plane (which fix the third coordinate), this implies that for any $v \in S^2 \setminus \{(0, 0, \pm 1\}$, there is an open neighbourhood U_v of v and a measure zero compact set containing a horizontal unit line segment with every direction in U_v. This implies that for any $\epsilon > 0$, there exists a measure zero compact set containing a horizontal unit line segment in every direction in $S^2 \setminus \mathcal{N}_\epsilon(\{(0, 0, \pm 1)\})$. This verifies the first claim.

For the second claim, take a countable sequence of translates $A_1, A_2, \ldots$ of A by vectors in $\mathbb{R}^2 \times \{0\}$ such that $d(A_j, 0) \to \infty$, and such that the projections of the A_j's onto $\mathbb{R}^2 \times \{0\}$ cover $\mathbb{R}^2$. Let

$$K_1 = (\mathbb{R} \times \{0\} \times \mathbb{R}) \cup \bigcup_{(a,b,c) \in \bigcup_j A_j} K(A_j)$$
$$= (\mathbb{R} \times \{0\} \times \mathbb{R}) \cup \bigcup_{(a,b,c) \in \bigcup_j A_j} \left\{\left(as + b, s, c + \tfrac{1}{2}bs\right) : |s| \le \frac{1}{\sqrt{4 + 4a^2 + b^2}}\right\},$$

and let K_2 be a rotation of K_1 by $\pi/2$ in the (x, y)-plane. It has already been shown that $K = K_1 \cup K_2$ has measure zero and contains a horizontal unit line segment in every direction in $S^2 \setminus \{(0, 0, \pm 1)\}$. By the condition $d(0, A_j) \to \infty$, together with the compactness of A and the definition of K, it is straightforward to check that K is closed. This verifies the second claim.

The third claim follows from a similar argument with the same A_j's; by taking

$$K_1 = (\mathbb{R} \times \{0\} \times \mathbb{R}) \cup \bigcup_{(a,b,c) \in \bigcup_j A_j} \big\{\big(as + b, s, c + \tfrac{1}{2}bs\big) : s \in \mathbb{R}\big\}.$$

Then $K = K_1 \cup K_2$ has measure zero by a similar argument to the above, and is a closed set of measure zero containing a horizontal line in every direction in $S^2 \setminus \{(0, 0, \pm 1)\}$.

Finally, the proof for SL_2 lines is similar. □

Corollary 2.3. *If $p \in (1, \infty]$ and*

$$\Big\| \sum_{T \in \mathbb{T}} \chi_T \Big\|_{L^p(B(0,1))} \leq C_\delta \tag{2-1}$$

for any set $\mathbb{T}$ of $\delta \times 1$ SL_2 tubes with δ-separated directions in S^2, then $\lim_{\delta \to 0} C_\delta = \infty$.

Let U be any Borel subset of $S^2 \setminus \{(0, 0, \pm 1)\}$ with $\mathcal{H}^2(U) > 0$. If, for some fixed $p \in [1, \infty)$,

$$\| f^*_{\delta, \mathrm{SL}_2} \|_{L^p(U)} \leq C_{\delta, U} \| f \|_p$$

for all $\delta > 0$ and for any nonnegative Borel function f on $\mathbb{R}^3$, then $\lim_{\delta \to 0} C_{\delta, U} = \infty$.

Proof. For the first part, using Theorem 2.2 let B be a compact set of measure zero containing a unit line segment in every direction in $U = S^2 \setminus \mathcal{N}_{0.01}\{(0, 0, \pm 1)\}$. Let $\widetilde{U} = \{v_1, \ldots, v_N\}$ be a maximal δ-separated subset of U, where $N \sim \delta^{-2}$. For each $v_k \in \widetilde{U}$, let T_k be a $\delta \times \delta \times 1$ SL_2 tube with direction v_k such that $T_k \subseteq \mathcal{N}_\delta(B)$. Then (for a sufficiently large absolute constant C)

$$1 \lesssim \int_{B(0,C)} \sum_{k=1}^N \chi_{T_k} \leq \mathcal{H}^3(\mathcal{N}_\delta(B))^{1-1/p} \Big\| \sum_{k=1}^N \chi_{T_k} \Big\|_{L^p(B(0,C))}.$$

Hence, by (2-1),

$$1 \lesssim \mathcal{H}^3(\mathcal{N}_\delta(B))^{1-1/p} C_\delta.$$

Letting $\delta \to 0$ proves the first part.

For the second part, let $V \subseteq U$ be a compact subset of U such that $\mathcal{H}^2(V) > 0$. Let B be a compact set of measure zero containing an SL_2 unit line segment in every direction in V, which exists by Theorem 2.2. Then

$$1 \lesssim \| (\chi_{\mathcal{N}_\delta}(B))^*_{\delta, \mathrm{SL}_2} \|_{L^p(V)} \leq C_{\delta, U} \mathcal{H}^3(\mathcal{N}_\delta(B))^{1/p}.$$

Letting $\delta \to 0$ proves the second part. □

3. Train tracks of planks

Given a Borel measure μ on $\mathbb{R}^3$ and $\alpha \geq 0$, recall that

$$c_\alpha(\mu) = \sup_{x \in \mathbb{R}^3,\, r > 0} \frac{\mu(B(x, r))}{r^\alpha}.$$

Let

$$I_\alpha(\mu) = \iint \frac{1}{|x - y|^\alpha} \, d\mu(x) \, d\mu(y).$$

Recall that $\gamma : [0, 2\pi) \to S^2$ is defined by $\gamma(\theta) = \frac{1}{\sqrt{2}}(\cos\theta, \sin\theta, 1)$, and recall that for each $\theta \in [0, 2\pi)$, π_θ is the orthogonal projection onto the plane $\gamma(\theta)^\perp$. Oberlin and Oberlin [2015] used Erdoğan's L^2 bound for the decay of conical averages of Fourier transforms of fractal measures from [Erdoğan 2004] to prove a bound on the average L^2 norms of pushforwards of fractal measures under π_θ. When $\alpha > \frac{5}{2}$, the uncertainty principle suggests that there is no loss in the approach from [Oberlin and Oberlin 2015] (up to the endpoint, provided only L^2 norms are considered), and this suggests that the Knapp examples from [Erdoğan 2004] used to prove sharpness of the L^2 conical decay rates can also be used to prove sharpness of the bound from [Oberlin and Oberlin 2015]. The following proposition verifies this intuition.

Proposition 3.1. *Let $\alpha \in [2, 3]$. If*

$$\int_0^{2\pi} \|\pi_{\theta\sharp}\mu\|^2_{L^2(\gamma(\theta)^\perp)}\, d\theta \le C_\alpha \mu(\mathbb{R}^3) c_\alpha(\mu) \tag{3-1}$$

for some nonzero finite C_α depending only on α, for all Borel measures μ on $B_3(0, 1)$, then $\alpha \ge \frac{5}{2}$. The same is true if $\mu(\mathbb{R}^3)c_\alpha(\mu)$ in (3-1) *is replaced by $I_\alpha(\mu)$.*

Remark 3.2. Oberlin and Oberlin [2015] proved that (3-1) does hold if $\alpha > \frac{5}{2}$, with either $\mu(\mathbb{R}^3)c_\alpha(\mu)$ or $I_\alpha(\mu)$ on the right-hand side.

Proof of Proposition 3.1. Let $\alpha \in [2, 3]$ and suppose that (3-1) holds. Let $\delta > 0$ be small. Fix any $\theta_0 \in [0, 2\pi)$. Let ψ be a smooth nonnegative bump function supported in $B(0, 1)$, with $\psi \sim 1$ on $B\big(0, \frac{1}{2}\big)$, such that $\int \psi = 1$. Let μ be the measure with Radon–Nikodym derivative equal to

$$\mu(x) = \delta^{-3/2}\psi\big(\langle x, \gamma(\theta_0)\rangle, \langle \delta^{-1/2}x, \sqrt{2}\gamma'(\theta_0)\rangle, \langle \delta^{-1}x, \sqrt{2}(\gamma \times \gamma')(\theta_0)\rangle\big),$$

i.e., μ is a Schwartz function of L^1 norm 1 supported on a plank of dimensions $\sim 1 \times \delta^{1/2} \times \delta$ centred at the origin, such that the longest direction is parallel to $\gamma(\theta_0)$, the medium direction is parallel to $\gamma'(\theta_0)$ and the shortest direction is parallel to $(\gamma \times \gamma')(\theta_0)$. If $|y_1| \le \frac{1}{100}\delta^{1/2}$, $|y_2| \le \frac{1}{100}\delta$, and $|y_3| \le \frac{1}{100}$, then

$$\mu\big(y_1\gamma'(\theta_0) + y_2(\gamma \times \gamma')(\theta_0) + y_3\gamma(\theta)\big) \gtrsim \delta^{-3/2}.$$

By second-order Taylor approximation, it follows that if $|\theta - \theta_0| < c\delta^{1/2}$ for $c > 0$ a sufficiently small absolute constant, then for all $|x_1| \le 10^{-3}\delta^{1/2}$ and $|x_2| \le 10^{-3}\delta$,

$$(\pi_{\theta\sharp}\mu)\big(x_1\gamma'(\theta) + x_2(\gamma \times \gamma')(\theta)\big) \ge \int_{-10^{-3}}^{10^{-3}} \mu\big(x_1\gamma'(\theta) + x_2(\gamma \times \gamma')(\theta) + t\gamma(\theta)\big)\, dt \gtrsim \delta^{-3/2}.$$

Therefore

$$\begin{aligned}\int_0^{2\pi} &\|\pi_{\theta\sharp}\mu\|^2_{L^2(\gamma(\theta)^\perp)}\, d\theta \\ &\ge \int_{|\theta-\theta_0|\le c\delta^{1/2}} \int_{[-10^{-3}\delta^{1/2}, 10^{-3}\delta^{1/2}]} \int_{[-10^{-3}\delta, 10^{-3}\delta]} \big|(\pi_{\theta\sharp}\mu)\big(x_1\gamma'(\theta) + x_2(\gamma \times \gamma')(\theta)\big)\big|^2\, dx_1\, dx_2\, d\theta \\ &\gtrsim \delta^{-1}.\end{aligned}$$

But $\mu(\mathbb{R}^3) = 1$, and it is straightforward to check that

$$c_\alpha(\mu) \sim \delta^{3/2-\alpha}.$$

Similarly $I_\alpha(\mu) \sim \delta^{3/2-\alpha}$ (e.g., via the Plancherel formula for the energy). Hence, by the assumed inequality (3-1),

$$\delta^{-1} \lesssim \delta^{3/2-\alpha}.$$

Letting $\delta \to 0$ gives $-1 \geq \frac{3}{2} - \alpha$ or $\alpha \geq \frac{5}{2}$. □

In the previous example, $c_\alpha(\mu)$ is much larger than $\mu(\mathbb{R}^3)$. But if ν is the sum of N copies of μ translated in the $\gamma'(\theta_0)$ and $(\gamma \times \gamma')(\theta_0)$ directions, such that the projections of these translated copies under π_θ are pairwise disjoint for $|\theta - \theta_0| \lesssim \delta^{1/2}$, then the lower bound for the left-hand side will be multiplied by N. If the translations are chosen sufficiently sparse, to ensure that $c_\alpha(\nu)$ is not larger than $c_\alpha(\mu)$, then the right-hand side will also be scaled by N, since $\mu(\mathbb{R}^3)$ will be scaled by N. If N could be chosen large enough to make $\nu(\mathbb{R}^3) \sim \delta^{3/2-\alpha} \sim c_\alpha(\nu) \sim c_\alpha(\mu)$, then this would show that no inequality is possible even with the larger right-hand side of $c_\alpha(\nu)^2$ instead of $\nu(\mathbb{R}^3)c_\alpha(\nu)$ in (3-1). The proposition below shows that this is possible using "parallel train tracks of planks". The spacing is based on the example of "parallel train tracks" from [Guth et al. 2020, Proposition 6.1], where the long spaces are a $\delta^{-1/2}$-multiple of the short spaces (the R in [Guth et al. 2020] corresponds to δ^{-1} here).

Proposition 3.3. *Let $\alpha \in [2, 3]$. If*

$$\int_0^{2\pi} \|\pi_{\theta\sharp}\nu\|^2_{L^2(\gamma(\theta)^\perp)}\, d\theta \leq C_\alpha c_\alpha(\nu)^2 \tag{3-2}$$

for some nonzero finite C_α depending only on α, for all Borel measures ν on $B_3(0, 1)$, then $\alpha \geq \frac{5}{2}$.

Proof. Let $\alpha \in [2, 3]$ and suppose that (3-2) holds. Let $\delta > 0$ be small, and fix any $\theta_0 \in [0, 2\pi)$. Let $\mu = \mu_{\delta,\theta_0}$ be the measure from the proof of Proposition 3.1, given by

$$\mu(x) = \delta^{-3/2}\psi\big(\langle x, \gamma(\theta_0)\rangle, \langle \delta^{-1/2}x, \sqrt{2}\gamma'(\theta_0)\rangle, \langle \delta^{-1}x, \sqrt{2}(\gamma \times \gamma')(\theta_0)\rangle\big),$$

where ψ is a smooth nonnegative bump function supported in $B(0, 1)$, with $\psi \sim 1$ on $B\big(0, \frac{1}{2}\big)$, such that $\int \psi = 1$. Let ν be the sum of translated copies of μ; spacing $\delta^{\alpha/2-1/2}$ in the short direction $(\gamma \times \gamma')(\theta_0)$, and spacing $\delta^{\alpha/2-1}$ in the medium direction $\gamma'(\theta_0)$ given by

$$\nu(x) = \sum_{|m|\leq 10^{-3}\delta^{1/2-\alpha/2}} \sum_{|n|\leq 10^{-3}\delta^{1-\alpha/2}} \mu\big(x - \delta^{\alpha/2-1/2}m(\gamma \times \gamma')(\theta_0) - n\delta^{\alpha/2-1}\gamma'(\theta_0)\big).$$

There are $\sim \delta^{3/2-\alpha}$ such copies, and the supports of the projections of these translated copies under π_θ are pairwise disjoint for $|\theta - \theta_0| \leq c\delta^{1/2}$, for a sufficiently small absolute constant c (using $\alpha \leq 3$ and second-order Taylor approximation), so the lower bound of δ^{-1} from the proof of Proposition 3.1 is multiplied by $\delta^{3/2-\alpha}$ to get

$$\int_0^{2\pi} \|\pi_{\theta\sharp}\nu\|^2_{L^2(\gamma(\theta)^\perp)}\, d\theta \gtrsim \delta^{1/2-\alpha}.$$

Moreover, $\nu(\mathbb{R}^3) \sim \delta^{3/2-\alpha}$, and

$$c_\alpha(\nu) \sim \delta^{3/2-\alpha}. \tag{3-3}$$

The lower bound $c_\alpha(\nu) \gtrsim \delta^{3/2-\alpha}$ in (3-3) follows from $c_\alpha(\nu) \gtrsim \nu(\mathbb{R}^3)$. The upper bound can be shown by considering different ranges of r separately, as follows. The maximum of $\nu(B(x,r))/r^\alpha$ over the range $0 < r \le \delta$ occurs near $r = \delta$ and is $\lesssim \delta^{3/2-\alpha}$. The maximum for the range $\delta \le r \le \delta^{\alpha/2-1/2}$ also occurs near $r = \delta$ (since the ball cannot intersect multiple "slats"), and is $\lesssim \delta^{3/2-\alpha}$. The maximum in the range $\delta^{\alpha/2-1/2} \le r \le \delta^{1/2}$ occurs near $r = \delta^{1/2}$, and is $\lesssim \delta^{3/2-\alpha}$. The maximum for the range $\delta^{1/2} \le r \le \delta^{\alpha/2-1}$ also occurs near $r = \delta^{1/2}$ (since the ball cannot intersect multiple "tracks"), and is $\lesssim \delta^{3/2-\alpha}$. The maximum for the range $\delta^{\alpha/2-1} \le r \le 1$ occurs near $r = 1$, and is $\lesssim \delta^{3/2-\alpha}$. Applying the assumed inequality (3-2) gives

$$\delta^{1/2-\alpha} \lesssim \delta^{3-2\alpha}.$$

Letting $\delta \to 0$ gives $\frac{1}{2} - \alpha \ge 3 - 2\alpha$ or $\alpha \ge \frac{5}{2}$. □

We will see that it is possible to weaken the requirement $\alpha \ge \frac{5}{2}$ by replacing the L^2 norm on the left in the previous examples by an L^p norm with $1 < p < 2$. Proposition 3.4 below shows that at the critical exponent $\alpha = 2$, the average L^p norms of δ-discretised measures cannot be bounded by $C_\epsilon \delta^{-\epsilon} c_2(\mu)^p$ unless $p \le \frac{3}{2}$. The example is similar to the "train tracks of planks" used in Proposition 3.3, except that as α approaches 2, the spaces between the parallel "train tracks" tends to a distance ~ 1, and the example reduces to a single "train track of planks", which is a plank version of the original "train track" example from [Wolff 1999, p. 563; Katz and Tao 2001, p. 151]. When $\alpha = 2$, a uniform bound (without the $C_\epsilon \delta^{-\epsilon}$ factor) is known to not be possible, by considering a purely 2-unrectifiable set in $\mathbb{R}^3$ and applying the Besicovitch–Federer projection theorem. Finally, the counterexample for projections onto lines below is based on the Knapp example from [Erdoğan 2004] used to prove sharpness of the L^2 conical decay rates of 1-dimensional fractal measures.

Proposition 3.4. *Let $p \in [1, \infty)$ and let $\epsilon > 0$.*

(1) *If, for any $\delta > 0$,*

$$\int_0^{2\pi} \|\pi_{\theta\sharp}\mu\|_{L^p(\gamma(\theta)^\perp)}^p \, d\theta \le C_\epsilon \delta^{-\epsilon} c_2(\mu)^p \tag{3-4}$$

for all Borel measures μ on $B_3(0,1)$ of the form

$$\mu = \frac{1}{\delta^3 \mathcal{H}^3(B(0,1))} \sum_{B \in \mathcal{B}} a_B \chi_B,$$

where $a_B > 0$ for all $B \in \mathcal{B}$ and where $\mathcal{B}$ is a disjoint family of δ-balls, then $p \le \frac{3}{2} + \epsilon$.

(2) *If, for any $\delta > 0$,*

$$\int_0^{2\pi} \|\rho_{\theta\sharp}\mu\|_{L^p(\mathrm{span}(\gamma(\theta)))}^p \, d\theta \le C_\epsilon \delta^{-\epsilon} c_1(\mu)^p \tag{3-5}$$

for all Borel measures μ on $B_3(0,1)$ of the form

$$\mu = \frac{1}{\delta^3 \mathcal{H}^3(B(0,1))} \sum_{B \in \mathcal{B}} a_B \chi_B,$$

where $a_B > 0$ for all $B \in \mathcal{B}$ and where $\mathcal{B}$ is a disjoint family of δ-balls, then $p \le \frac{3}{2} + \epsilon$.

Proof. Let $\delta > 0$ be small, fix $\theta_0 \in [0, 2\pi)$ and assume that (3-4) holds. Let $\mu = \mu_{\delta,\theta_0}$ be the measure from the proof of Proposition 3.1 given by

$$\mu(x) = \delta^{-3/2}\psi\big(\langle x, \gamma(\theta_0)\rangle, \langle \delta^{-1/2}x, \sqrt{2}\gamma'(\theta_0)\rangle, \langle \delta^{-1}x, \sqrt{2}(\gamma \times \gamma')(\theta_0)\rangle\big),$$

where ψ is a smooth nonnegative bump function supported in $B(0, 1)$, with $\psi \sim 1$ on $B\big(0, \frac{1}{2}\big)$, such that $\int \psi = 1$. Similarly to the proof of Proposition 3.1,

$$\int_0^{2\pi} \|\pi_{\theta\sharp}\mu\|^p_{L^p(\gamma(\theta)^\perp)}\, d\theta \geq \int_{|\theta-\theta_0|\leq c\delta^{1/2}} \int_{-10^{-3}\delta^{1/2}}^{10^{-3}\delta^{1/2}} \int_{-10^{-3}\delta}^{10^{-3}\delta} \delta^{-3p/2}\, dx_1\, dx_2\, d\theta \gtrsim \delta^{2-3p/2}. \tag{3-6}$$

But $\mu(\mathbb{R}^3) = 1$, and

$$c_2(\mu) \sim \delta^{-1/2}.$$

Moreover, μ is essentially a sum of indicator functions over a boundedly overlapping family of δ-balls. Hence, if (3-4) holds with the smaller $\mu(\mathbb{R}^3)c_2(\mu)^{p-1}$ on the right-hand side instead of $c_2(\mu)^p$, then

$$\delta^{2-3p/2} \lesssim \delta^{1/2-p/2-\epsilon}.$$

Letting $\delta \to 0$ gives $2 - \frac{3p}{2} \geq \frac{1}{2} - \frac{p}{2} - \epsilon$ or $p \leq \frac{3}{2} + \epsilon$.

In order to get the same restriction on p if the larger right-hand side $c_2(\mu)^p$ is assumed, let ν be the sum of translated copies of μ spacing $\delta^{1/2}$ in the short direction $(\gamma \times \gamma')(\theta_0)$, given by

$$\nu(x) = \sum_{|m|\leq 10^{-3}\delta^{-1/2}} \mu\big(x - \delta^{1/2}m(\gamma \times \gamma')(\theta_0)\big).$$

There are $\sim \delta^{-1/2}$ such copies, and the supports of the projections of these translated copies under π_θ are pairwise disjoint for $|\theta - \theta_0| \leq c\delta^{1/2}$, for a sufficiently small absolute constant c, so the lower bound of $\delta^{2-3p/2}$ from (3-6) is multiplied by $\delta^{-1/2}$ to get

$$\int_0^{2\pi} \|\pi_{\theta\sharp}\nu\|^p_{L^p(\gamma(\theta)^\perp)}\, d\theta \gtrsim \delta^{(3-3p)/2}. \tag{3-7}$$

But $\nu(\mathbb{R}^3) \sim \delta^{-1/2} \sim c_2(\nu)$ (as shown in the proof of Proposition 3.3), and ν is essentially a sum of indicator functions over a boundedly overlapping family of δ-balls. Hence, by the assumed (3-4),

$$\delta^{(3-3p)/2} \lesssim \delta^{-p/2-\epsilon}.$$

Letting $\delta \to 0$ gives $\frac{1}{2}(3 - 3p) \geq -\frac{1}{2}p - \epsilon$ or $p \leq \frac{3}{2} + \epsilon$. This proves the restriction on p for projections onto planes.

It remains to consider the restriction on p for projections onto lines. Let $\mu = \mu_{\delta,\theta_0}$ be the measure similar to the one from the proof of Proposition 3.1, given by

$$\mu(x) = \delta^{-3/2}\psi\big(\langle \delta^{-1}x, \gamma(\theta_0)\rangle, \langle \delta^{-1/2}x, \sqrt{2}\gamma'(\theta_0)\rangle, \langle x, \sqrt{2}(\gamma \times \gamma')(\theta_0)\rangle\big),$$

where ψ is a smooth nonnegative bump function supported in $B(0, 1)$, with $\psi \sim 1$ on $B\big(0, \frac{1}{2}\big)$, such that $\int \psi = 1$. Then $\mu(\mathbb{R}^3) \sim c_1(\mu) \sim 1$. For any $|\theta - \theta_0| \lesssim \delta^{1/2}$, μ is roughly $\delta^{-3/2}$ times the indicator

function of a $\sim \delta \times \delta^{1/2} \times 1$ plank in directions $\gamma(\theta)$, $\gamma'(\theta)$, and $(\gamma \times \gamma')(\theta)$, respectively, and so $\rho_{\theta\sharp}\mu$ is roughly a δ^{-1} multiple of the indicator function of an interval of length $\sim \delta$. Hence

$$\int_{[\theta_0-\delta^{1/2},\theta_0+\delta^{1/2}]} \int |\rho_{\theta\sharp}\mu|^p \, d\theta \gtrsim \delta^{3/2-p}.$$

Hence, by the assumed (3-5),

$$\delta^{3/2-p} \lesssim \delta^{-\epsilon}.$$

Letting $\delta \to 0$ gives $p \le \frac{3}{2} + \epsilon$. This proves the restriction on p for projections onto lines. □

For $(x, y, t) \in \mathbb{R}^3$, define $\ell^*(x, y, t)$ to be the line

$$\ell^*(x, y, t) = \big(0, x, t - \tfrac{1}{2}xy\big) + L_y,$$

where L_y is the "light ray" in the light cone

$$\widetilde{\Gamma} := \{\eta \in \mathbb{R}^3 : \eta_2^2 = 2\eta_1\eta_3\}$$

given by

$$L_y = \big\{\lambda\big(1, -y, \tfrac{1}{2}y^2\big) : \lambda \in \mathbb{R}\big\}.$$

The cone $\widetilde{\Gamma}$ is the image of the light cone

$$\Gamma := \{\xi \in \mathbb{R}^3 : \xi_3^2 = \xi_1^2 + \xi_2^2\}$$

in $\mathbb{R}^3$, under the orthogonal transformation $U(\xi_1, \xi_2, \xi_3) = (\eta_1, \eta_2, \eta_3)$ given by

$$\eta_1 = \frac{\xi_1 + \xi_3}{\sqrt{2}}, \quad \eta_3 = \frac{-\xi_1 + \xi_3}{\sqrt{2}}, \quad \eta_2 = \xi_2.$$

The restriction of ℓ^* to $|y| \le \sqrt{2}$, followed by U^* (an anticlockwise rotation by $\pi/4$ in the (ξ_1, ξ_3)-plane), parametrises the family of light rays parallel to some $\gamma(\theta)$ with $|\theta| \le \pi/2$.

Given $(a, b, c) \in \mathbb{R}^3$, define $\ell(a, b, c)$ to be the horizontal line

$$\ell(a, b, c) = \{(b, 0, c) + s(a, 1, b/2) : s \in \mathbb{R}\}. \tag{3-8}$$

The following is the point-line duality principle.

Lemma 3.5 [Fässler and Orponen 2023b, Lemma 4.11]. *Let $p \in \mathbb{R}^3$ and $p^* \in \mathbb{H}$. Then*

$$p \in \ell^*(p^*) \quad \textit{if and only if} \quad p^* \in \ell(p).$$

The symmetry of the above lemma means that many of the ideas from [Fässler and Orponen 2023b] for vertical projections in the Heisenberg group can be reversed as in the lemma below.

Lemma 3.6. *The measure $\mathfrak{m}$ given by*

$$\mathfrak{m}(F) = \int_{-\pi/2}^{\pi/2} \mathcal{H}^2\big\{w \in \gamma(\theta)^\perp : \pi_\theta^{-1}(w) \in U^*F\big\} \, d\theta,$$

for a Borel set F of light rays parallel to lines in $\widetilde{\Gamma}$, is comparable to the pushforward of the Lebesgue measure on $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : |x_2| \le \sqrt{2}\}$ under ℓ^*, meaning that

$$(\ell^*_\sharp \mathcal{H}^3 \chi_{|x_2|\le\sqrt{2}})(F) \sim \mathfrak{m}(F) \tag{3-9}$$

for any Borel set F of light rays parallel to lines in $\widetilde{\Gamma}$. As a consequence, for any nonnegative Borel function f on the set of light rays parallel to lines in $\widetilde{\Gamma}$,

$$\int_{-\pi/2}^{\pi/2} \int_{\gamma(\theta)^\perp} f(U\pi_\theta^{-1}(w))\, d\mathcal{H}^2(w)\, d\theta \sim \int f\, d(\ell^*_\sharp \mathcal{H}^3 \chi_{|x_2|\le\sqrt{2}}). \tag{3-10}$$

Proof. Let F be given. Let m be the Lebesgue measure on $\mathbb{R}^3 \cap \{|x_2| \le \sqrt{2}\}$. Then

$$\begin{aligned}
(\ell^*_\sharp m)(F) &= m\big\{(x, y, t) \in \mathbb{R}^3 : |y| \le \sqrt{2},\ \big(0, x, t - \tfrac{1}{2}xy\big) + L_y \in F\big\} \\
&= \int_{-\sqrt{2}}^{\sqrt{2}} \mathcal{H}^2\big\{(x, t) \in \mathbb{R}^2 : \big(0, x, t - \tfrac{1}{2}xy\big) + \operatorname{span}(1, -y, y^2/2) \in F\big\}\, dy \\
&\sim \int_{-\sqrt{2}}^{\sqrt{2}} \mathcal{H}^2\big\{w \in (1, -y, y^2/2)^\perp : \pi^{-1}_{(1,-y,y^2/2)^\perp}(w) \in F\big\}\, dy \\
&= \int_{-\sqrt{2}}^{\sqrt{2}} \mathcal{H}^2\big\{w \in (U^*(1, -y, y^2/2))^\perp : \pi^{-1}_{(U^*(1,-y,y^2/2))^\perp}(w) \in U^*F\big\}\, dy \\
&\sim \int_{-\pi/2}^{\pi/2} \mathcal{H}^2\big\{w \in (\cos\theta, \sin\theta, 1)^\perp : \pi^{-1}_{(\cos\theta,\sin\theta,1)^\perp}(w) \in U^*F\big\}\, d\theta \\
&= \mathfrak{m}(F).
\end{aligned}$$

This proves (3-9) and yields (3-10) whenever $f = \chi_F$ for a Borel set F of light rays. The equivalence (3-10) for general Borel functions follows by approximating f with simple functions and then applying the monotone convergence theorem. □

Proposition 3.7. *If, for any $\epsilon > 0$ and $\delta > 0$,*

$$\left\| \sum_{T\in\mathbb{T}} \chi_T \right\|^p_{L^p(B(0,1))} \le C_\epsilon \delta^{-\epsilon} \tag{3-11}$$

for any family $\mathbb{T}$ of SL_2 δ-tubes of length 1 with δ-separated directions, then $p \le \frac{3}{2}$.

Proof. It is first shown that if, for any $\epsilon > 0$, (3-11) holds for any family $\mathbb{T}$ of SL_2 δ-tubes of length 1 satisfying the *weaker* property that

$$|\{T \in \mathbb{T} : T \subseteq S\}| \le (r/\delta)^2,$$

for any tube S of length 2 and radius r, and any $r \in [\delta, 1]$, then $p \le \frac{3}{2}$.

Let $\delta > 0$ be small and choose $\theta_0 = \pi/2$. Let $\mu = \mu_{\delta,\theta_0}$ be a $\delta^{1/2}$-multiple of the (slightly modified) measure from the proof of Proposition 3.1, given by

$$\mu(x) = \delta^{-1}\psi\big(\langle x, \gamma(\theta_0)\rangle, \langle \delta^{-1/2}x, \sqrt{2}\gamma'(\theta_0)\rangle, \langle \delta^{-1}x, \sqrt{2}(\gamma \times \gamma')(\theta_0)\rangle\big),$$

where ψ is a smooth nonnegative bump function supported in $B(0,c)$, with $\psi \sim 1$ on $B(0,c/2)$, such that $\int \psi = 1$ (here c is a small absolute constant to be chosen; this is the only difference from the proof of Proposition 3.1 where $c = 1$). As in the proof of Proposition 3.4, let ν be the sum of translated copies of μ spacing $\delta^{1/2}$ in the short direction $(\gamma \times \gamma')(\theta_0)$, given by

$$\nu(x) = \sum_{|m| \le c\delta^{-1/2}} \mu\big(x - \delta^{1/2} m(\gamma \times \gamma')(\theta_0)\big).$$

Then ν is supported in a ball around the origin of radius $\sim c$, and ν is very similar to a $\delta^{1/2}$-multiple of the measure from the proof of Proposition 3.4, so by a similar argument to (3-7),

$$\delta^{3/2-p} \lesssim \int_{-\pi/2}^{\pi/2} \|\pi_{\theta\sharp}\nu\|^p_{L^p(\gamma(\theta)^\perp)}\, d\theta.$$

Let $\mathcal{B}$ be a family of δ-balls, such that the centres of the balls in $\mathcal{B}$ form a maximal δ-separated subset of the support of ν. Then

$$\nu \le \sum_{B \in \mathcal{B}} \delta^{-1} \chi_B.$$

Hence, by Lemma 3.6,

$$\begin{aligned}
\int_{-\pi/2}^{\pi/2} \|\pi_{\theta\sharp}\nu\|^p_{L^p(\gamma(\theta)^\perp)}\, d\theta &= \int_{-\pi/2}^{\pi/2} \int_{\gamma(\theta)^\perp} \left| \int_{\pi_\theta^{-1}(y)} \nu \, d\mathcal{H}^1 \right|^p d\mathcal{H}^2(y)\, d\theta \\
&= \int_{-\pi/2}^{\pi/2} \int_{\gamma(\theta)^\perp} \left| \int_{U\pi_\theta^{-1}(y)} (U_\sharp \nu)\, d\mathcal{H}^1 \right|^p d\mathcal{H}^2(y)\, d\theta \\
&\sim \int \left| \int_L U_\sharp \nu \, d\mathcal{H}^1 \right|^p d(\ell^*_\sharp \mathcal{H}^3 \chi_{|x_2| \le \sqrt{2}})(L) \\
&\lesssim \int |\{B \in \mathcal{B} : UB \cap L \ne \varnothing\}|^p \, d(\ell^*_\sharp \mathcal{H}^3 \chi_{|x_2| \le \sqrt{2}})(L) \\
&= \int_{|p_2^*| \le \sqrt{2}} |\{B \in \mathcal{B} : UB \cap \ell^*(p^*) \ne \varnothing\}|^p \, d\mathcal{H}^3(p^*). \qquad (3\text{-}12)
\end{aligned}$$

If $|p_2^*| \le \sqrt{2}$ and $\ell^*(p^*)$ intersects $B(0,c)$, then $|p^*| \le 2$ (provided c is now chosen sufficiently small; this follows easily from the definition of ℓ^*). Hence

$$(3\text{-}12) \le \int_{|p^*| \le 2} |\{B \in \mathcal{B} : UB \cap \ell^*(p^*) \ne \varnothing\}|^p \, d\mathcal{H}^3(p^*). \qquad (3\text{-}13)$$

By Lemma 3.5 (point-line duality),

$$(3\text{-}13) \le \int_{|p^*| \le 2} \left| \sum_{B \in \mathcal{B}} \chi_{\ell(UB)}(p^*) \right|^p d\mathcal{H}^3(p^*).$$

The measure ν satisfies $c_2(\nu) \lesssim 1$, and similarly so does the measure $\sum_{B \in \mathcal{B}} \delta^{-1} \chi_B$ (which is roughly the same as ν), which means that for any $\delta \le r \le 1$, the number of δ-balls from $\mathcal{B}$ intersecting any r-ball is

$\lesssim (r/\delta)^2$. By the formula for ℓ, this implies that for any $r \in [\delta, 1]$ and any tube r-tube T of length 2,

$$|\{\ell(UB) : B \in \mathcal{B}, \ell(UB) \subseteq T\}| \lesssim (r/\delta)^2.$$

Therefore, if (3-11) holds for any family of SL_2 tubes satisfying the 2-dimensional ball condition above, then

$$\delta^{3/2-p} \lesssim \int_{|p^*|\le 2} \left| \sum_{B\in\mathcal{B}} \chi_{\ell(UB)}(p^*) \right|^p d\mathcal{H}^3(p^*) \le C_\epsilon \delta^{-\epsilon}.$$

This yields $p \le \frac{3}{2} + \epsilon$.

To get the same restriction on p when (3-11) holds only for those sets $\mathbb{T}$ of SL_2 δ-tubes of length 1 with δ-separated directions, it suffices to show that the collection of tubes $\ell(UB)$ have $\sim \delta$-separated directions, for any subfamily of balls B in $\mathcal{B}$ separated by 100δ. Let $B_1, B_2 \in \mathcal{B}$ be 100δ-separated balls in $\mathcal{B}$. Let (a_j, b_j, c_j) be the centre of UB_j, where $j \in \{1, 2\}$. The direction of $\ell(UB_j)$ is parallel to $(a_j, 1, b_j/2)$, so it is required to show that

$$|a_1 - a_2| + |b_1 - b_2| \gtrsim \delta.$$

Write

$$B_2 = B_1 + \lambda_1\gamma(\pi/2) + \lambda_2\gamma'(\pi/2) + \lambda_3(\gamma \times \gamma')(\pi/2),$$

where $|\lambda_3| \le 1$, $|\lambda_2| \le c\delta^{1/2}$ and $|\lambda_1| \le 1$, and $|(\lambda_1, \lambda_2, \lambda_3)| \ge 100\delta$. This can be written as

$$B_2 = B_1 + \frac{\lambda_1}{\sqrt{2}}(0, 1, 1) + \frac{\lambda_2}{\sqrt{2}}(-1, 0, 0) + \frac{\lambda_3}{2}(0, -1, 1).$$

Hence

$$UB_2 = UB_1 + \frac{\lambda_1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, 1, \frac{1}{\sqrt{2}}\right) + \frac{\lambda_2}{\sqrt{2}}\left(\frac{-1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) + \frac{\lambda_3}{2}\left(\frac{1}{\sqrt{2}}, -1, \frac{1}{\sqrt{2}}\right).$$

It follows that

$$(a_2, b_2) = (a_1, b_1) + \left(\frac{\lambda_1}{2} - \frac{\lambda_2}{2} + \frac{\lambda_3}{2\sqrt{2}}, \frac{\lambda_1}{\sqrt{2}} - \frac{\lambda_3}{2}\right).$$

By the triangle inequality,

$$\left|\frac{\lambda_1}{2} - \frac{\lambda_2}{2} + \frac{\lambda_3}{2\sqrt{2}}\right| \ge \left|\frac{\lambda_3}{\sqrt{2}} - \frac{\lambda_2}{2}\right| - \frac{1}{\sqrt{2}}\left|\frac{\lambda_1}{\sqrt{2}} - \frac{\lambda_3}{2}\right|.$$

Hence

$$|a_1 - a_2| + \frac{|b_1 - b_2|}{\sqrt{2}} \ge \left|\frac{\lambda_3}{\sqrt{2}} - \frac{\lambda_2}{2}\right|.$$

If B_1 and B_2 lie in different planks, then $|\lambda_3| \ge \delta^{1/2}$, and since $|\lambda_2| \le c\delta^{1/2}$, this yields

$$|a_1 - a_2| + |b_1 - b_2| \gtrsim \delta^{1/2}.$$

It remains to consider the case where B_1 and B_2 lie in the same plank. In this case, $|\lambda_3| \le \delta$. If $|\lambda_1| \ge 10\delta$, then

$$|b_1 - b_2| \sim \left|\frac{\lambda_1}{\sqrt{2}} - \frac{\lambda_3}{2}\right| \gtrsim \delta.$$

Otherwise $|\lambda_1| < 10\delta$ and $|\lambda_3| \le \delta$ imply that $|\lambda_2| \ge 50\delta$, and this gives

$$|a_1 - a_2| = \left|\frac{\lambda_1}{2} - \frac{\lambda_2}{2} + \frac{\lambda_3}{2\sqrt{2}}\right| \gtrsim \delta.$$

This shows that the directions of 100δ-separated balls in $\mathcal{B}$ are $\sim \delta$-separated. It follows that if (3-11) holds for any family $\mathbb{T}$ of SL_2 δ-tubes of length 1 with δ-separated directions, then it can be applied to the family $\{\ell(UB) : B \in \mathcal{B}\}$, and therefore

$$\delta^{3/2-p} \lesssim \int_{|p^*|\le 2} \left|\sum_{B\in\mathcal{B}} \chi_{\ell(UB)}(p^*)\right|^p d\mathcal{H}^3(p^*) \le C_\epsilon \delta^{-\epsilon}.$$

This yields $p \le \frac{3}{2} + \epsilon$. Letting $\epsilon \to 0$ gives $p \le \frac{3}{2}$. □

4. An $L^{3/2}$ inequality for projections onto planes

To begin this section, we will outline a decomposition of the outer part of the light cone

$$\{(\xi_1, \xi_2, \xi_3) : |\xi_3| = |(\xi_1, \xi_2)|\},$$

which will be important in the proof of Theorem 1.2. The "outer part" means the set

$$\{(\xi_1, \xi_2, \xi_3) : |(\xi_1, \xi_2)| > |\xi_3|\}.$$

For each $\theta \in [0, 2\pi)$ and $0 \le k < j$, let

$$\tau(\theta, j, k, +) = \big\{\lambda_1(\gamma \times \gamma')(\theta) + \lambda_2\gamma'(\theta) + \lambda_3\gamma(\theta) : |\lambda_1| \sim 2^j,\ |\lambda_2| \lesssim 2^{j-k/2},\ |\lambda_3| \sim 2^{j-k},\ \lambda_1 > 0,\ \lambda_3 < 0\big\},$$

and for $k = j$ let

$$\tau(\theta, j, j, +) = \big\{\lambda_1(\gamma \times \gamma')(\theta) + \lambda_2\gamma'(\theta) + \lambda_3\gamma(\theta) : |\lambda_1| \sim 2^j,\ |\lambda_2| \lesssim 2^{j/2},\ |\lambda_3| \lesssim 1,\ \lambda_1 > 0\big\}.$$

To be precise, $|\lambda_1| \sim 2^j$ can be taken to mean $2^j/100 \le |\lambda_1| \le 100 \cdot 2^j$, and similarly for the other implied constants here; the exact constants we choose are not important. Similarly, we define

$$\tau(\theta, j, k, -) = \big\{\lambda_1(\gamma \times \gamma')(\theta) + \lambda_2\gamma'(\theta) + \lambda_3\gamma(\theta) : |\lambda_1| \sim 2^j,\ |\lambda_2| \lesssim 2^{j-k/2},\ |\lambda_3| \sim 2^{j-k},\ \lambda_1 < 0,\ \lambda_3 > 0\big\},$$

and for $k = j$,

$$\tau(\theta, j, j, -) = \big\{\lambda_1(\gamma \times \gamma')(\theta) + \lambda_2\gamma'(\theta) + \lambda_3\gamma(\theta) : |\lambda_1| \sim 2^j,\ |\lambda_2| \lesssim 2^{j/2},\ |\lambda_3| \lesssim 1,\ \lambda_1 < 0\big\}.$$

For each such τ, we write $\theta = \theta_\tau$. For each k, choose a maximal $c2^{-k/2}$-separated subset Θ_k of $[0, 2\pi)$, let $\Lambda_{j,k} = \{\tau(\theta, j, k, \pm) : \theta \in \Theta_k\}$, where c is a small constant such that the set $\Lambda = \bigcup_{j\ge1} \bigcup_{0\le k\le j} \Lambda_{j,k}$ forms a boundedly overlapping cover of the set $\big\{\xi \in \mathbb{R}^3 : \frac{1}{3} \le |\xi_3| \le |(\xi_1, \xi_2)| \le 3|\xi_3|\big\}$.

The covering property can be seen as follows. If ξ is such that $\frac{1}{3} \le |\xi_3| \le |(\xi_1, \xi_2)| \le 3|\xi_3|$, we can choose $j \ge 1$ such that $2^j/2 \le |\xi| \le 2 \cdot 2^j$, and let $\omega = 2^{-j}\xi$. Then ω lies outside the light cone and

has a distance ~ 1 to the origin, and distance $\lesssim 1$ to the light cone. Therefore, we can find $0 \leq k \leq j$ such that the distance from ω to the light cone is either $\lesssim 2^{-j}$ or $\sim 2^{-k}$. We can then express ω as $\lambda_1(\gamma \times \gamma')(\theta_0) + \lambda_3\gamma(\theta_0)$ for some $\theta_0 \in [0, 2\pi)$, $|\lambda_1| \sim 1$ and either $|\lambda_3| \lesssim 2^{-j}$ or $\lambda_3 \sim 2^{-k}$ (with appropriate signs depending on whether ω is closer to the forward or backward light cone). By writing $\theta_0 = \theta + O(2^{-k/2})$ for some $\theta \in \Theta_k$ and applying a Taylor expansion, and then scaling up $\xi = 2^j\omega$, we get that ξ is in one of the sets $\tau(\theta, j, k, \pm)$, which verifies the covering property.

Let $\{\psi_\tau\}_{\tau\in\Lambda}$ be a smooth partition of unity such that each ψ_τ is supported in τ and such that the ψ_τ's sum to 1 on the set

$$\left\{\xi \in \mathbb{R}^3 : \tfrac{1}{2} \leq |\xi_3| \leq |(\xi_1, \xi_2)| \leq 2|\xi_3|\right\}.$$

The partition of unity can be constructed so that the derivatives of each ψ_τ satisfy natural bounds; meaning that for any $m \geq 1$, the size of the m-th derivative of ψ_τ in the direction $(\gamma \times \gamma')(\theta_\tau)$ is $\lesssim_m 2^{-jm}$, the size of the m-th derivative of ψ_τ in the direction $\gamma'(\theta_\tau)$ is $\lesssim_m 2^{-(j-k/2)m}$, and the size of the m-th derivative of ψ_τ in the direction $\gamma(\theta_\tau)$ is $\lesssim_m 2^{-(j-k)m}$ (see [Harris 2023b, Lemma 3.3] for a more precise statement).

Given $\delta > 0$ (which may be thought of as vanishingly small) and $0 \leq k \leq j$, for each $\tau \in \Lambda_{j,k}$ cover $\mathbb{R}^3$ by a boundedly overlapping collection of planks $T \in \mathbb{T}_\tau$ of dimensions

$$2^{k\delta-(j-k)} \times 2^{k(-1/2+\delta)-(j-k)} \times 2^{k(-1+\delta)-(j-k)}$$

dual to τ (meaning that the long direction of T is the short direction of τ, and vice-versa). Given a measure μ on $B_3(0, 1)$ and a plank T, define

$$M_T\mu = \eta_T(\mu * \check{\psi}_\tau),$$

where $\sum_{T\in\mathbb{T}_\tau} \eta_T = 1$, and such that for each T, η_T is a bump function supported on T, with Fourier transform essentially supported in a dual box to T.

Lemma 4.2 below states that the projections of measures of dimension greater than 2 have small L^p norms on small unions of rectangles, on average, for exponents $1 \leq p \leq \frac{3}{2}$. The main use of Lemma 4.2 will be in proving that pushforwards of measures of dimension greater than 2 are almost surely in $L^{3/2}$; by splitting μ into "good" and "bad" parts, and bounding the $L^{3/2}$ norm of the projections of the "bad" part of a measure by the $L^{3/2}$ norm of the projections of the original measure on small unions of rectangles. Before we state the lemma, we give a definition which makes "small unions of rectangles" more precise. Let ϕ be a smooth bump function equal to 1 on $B_3(0, 1)$ and vanishing outside $B_3(0, 2)$. For each $R > 0$ let $\phi_R(x) = R^3\phi(Rx)$.

Definition 4.1. If $\alpha \in [0, 3]$, $\beta > 0$, $R \geq 1$, and μ is a Borel measure on $B_3(0, 1)$ with $c_\alpha(\mu) < \infty$, then a triple $(\mu, R, \{\mathbb{D}_\theta\})$ is called β-admissible if $\{\mathbb{D}_\theta\}$ is a family of sets, where each $\mathbb{D}_\theta$ is a set of rectangles of dimensions $R^{-1/2} \times R^{-1}$ in $\gamma(\theta)^\perp$ with long direction parallel to $\gamma'(\theta)$ and short direction parallel to $(\gamma \times \gamma')(\theta)$, with $(\pi_{\theta\sharp}(\mu * \phi_R))(D) \geq c_\alpha(\mu)R^{\beta-3/2}$ for each $D \in \mathbb{D}_\theta$, such that the set

$$\left\{(\theta, x) \in [0, 2\pi) \times \mathbb{R}^3 : x \in \bigcup_{D\in\mathbb{D}_\theta} D\right\}$$

is Borel measurable.

Lemma 4.2. *Let $\alpha > 2$ and $1 \le p \le \frac{3}{2}$. Then for any $\beta > 0$, there exists a $\kappa > 0$ (which may depend on α, β, and p) such that*

$$\int_0^{2\pi} \int_{\bigcup_{D\in\mathbb{D}_\theta}} |\pi_{\theta\sharp}(\mu * \phi_R)|^p \, d\mathcal{H}^2 \, d\theta \le C_{\alpha,\beta,p} R^{-\kappa} \mu(\mathbb{R}^3) c_\alpha(\mu)^{p-1} \tag{4-1}$$

for all β-admissible datum $(\mu, R, \{\mathbb{D}_\theta\})$.

Remark 4.3. Having $\mu * \phi_R$ instead of μ in (4-4) is similar to assuming that μ is a linear combination of characteristic functions over a boundedly overlapping collection of R^{-1}-balls.

Proof. Apparently, the result follows from the endpoint cases $p = 1$ and $p = \frac{3}{2}$ and interpolation. Instead of proving the two endpoint cases directly, we will in fact show that the claimed result holds true for a sequence of increasing exponents $\{p_n\}$. Briefly speaking, the proof is based on a self-improving argument: if (4-1) holds for the exponent p_n, then it holds for p_{n+1}.

Here are the details. Define the sequence $\{p_n\}_{n\in\mathbb{N}\cup\{0,\infty\}}$ by

$$\frac{1}{p_n} = \frac{2}{3} + \frac{1}{3\cdot 4^n}, \quad n \in \mathbb{N} \cup \{0, \infty\}. \tag{4-2}$$

We claim that the following holds true: if either

(1) $n \in \{0, 1\}$, or if

(2) $2 \le n \le \infty$, and for all $m \in \mathbb{N} \setminus \{0\}$ with $m < n$, for any $\beta > 0$, there exists a $\kappa > 0$ (depending on α, β, m) such that

$$\int_0^{2\pi} \int_{\bigcup_{D\in\mathbb{D}_\theta}} |\pi_{\theta\sharp}(\mu * \phi_R)|^{p_m} \, d\mathcal{H}^2 \, d\theta \le C_{\alpha,\beta,m} R^{-\kappa} \mu(\mathbb{R}^3) c_\alpha(\mu)^{p_m - 1} \tag{4-3}$$

for all β-admissible datum $(\mu, R, \{\mathbb{D}_\theta\})$,

then for any $\beta > 0$, there exists a $\kappa > 0$ (which may depend on α, β, and n) such that

$$\int_0^{2\pi} \int_{\bigcup_{D\in\mathbb{D}_\theta}} |\pi_{\theta\sharp}(\mu * \phi_R)|^{p_n} \, d\mathcal{H}^2 \, d\theta \le C_{\alpha,\beta,n} R^{-\kappa} \mu(\mathbb{R}^3) c_\alpha(\mu)^{p_n - 1} \tag{4-4}$$

for all β-admissible datum $(\mu, R, \{\mathbb{D}_\theta\})$.

It is easy to see that the claim would imply the desired result. Indeed, the endpoint case $p = 1$ follows directly from setting $n = 0$ (for which (4-4) holds unconditionally). Induction then shows that (4-4) holds for all finite n. Lastly, applying the claim again to $n = \infty$ proves the endpoint case $p = \frac{3}{2}$.

We now prove the claim. For a given n, the proof is a bootstrapping argument which gradually lowers the value of β, which for simplicity will be written as a proof by contradiction. Define $p_m = p_0 = 1$ for $m < 0$. Let $n \in \mathbb{N} \cup \{0, \infty\}$ be given and let β_{inf} be the infimum over all positive β for which the conclusion of the lemma is true for $p = p_n$. Suppose for a contradiction that $\beta_{\text{inf}} > 0$. Clearly $\beta_{\text{inf}} \le \frac{3}{2}$ (in particular it is finite); this follows from the inequality $\mu(\mathbb{R}^3) \le c_\alpha(\mu)$. Choose a constant $c \in (0, 1)$ sufficiently close to 1 to ensure that $1 - c < \frac{1}{1000}(\alpha - 2)$, then let $\beta_0 \in (\beta_{\text{inf}}, 2\beta_{\text{inf}})$ be sufficiently close to β_{inf} so that $c\beta_0 < \beta_{\text{inf}}$. To obtain a contradiction, it will suffice to show that the lemma holds for all

$\beta > c\beta_0$. Fix such a $\beta > c\beta_0$. Let $\delta > 0$ be such that $\delta \ll \kappa(\beta_0)$ and $\delta \ll \beta$. The measure $\mu * \phi_R$ will be relabelled as μ to simplify the notation (so it may be assumed that $\widehat{\mu}$ is negligible outside $B(0, R^{1+\delta})$).

We first derive an estimate of the measure of the integration domain. Let

$$F = \left\{(\theta, x) \in [0, 2\pi) \times \mathbb{R}^3 : x \in \bigcup_{D \in \mathbb{D}_\theta} D\right\}.$$

By the assumption that $(\pi_{\theta\sharp}\mu)(D) \geq R^{\beta-3/2} c_\alpha(\mu)$ and the Vitali covering lemma (to handle the possibly nondisjoint case),

$$(\mathcal{H}^1 \times \mathcal{H}^2)(F) \lesssim c_\alpha(\mu)^{-1} R^{-\beta} \int_0^{2\pi} \int_{\bigcup_{D \in \mathbb{D}_\theta} D} |\pi_{\theta\sharp}\mu| \, d\mathcal{H}^2 \, d\theta.$$

If $n < \infty$, then by Hölder's inequality applied to χ_F and $|\pi_{\theta\sharp}\mu|$,

$$(\mathcal{H}^1 \times \mathcal{H}^2)(F)^{1/p_{n-1}} \lesssim c_\alpha(\mu)^{-1} R^{-\beta} \left(\int_0^{2\pi} \int_{\bigcup_{D \in \mathbb{D}_\theta} D} |\pi_{\theta\sharp}\mu|^{p_{n-1}} \, d\mathcal{H}^2 \, d\theta \right)^{1/p_{n-1}}.$$

Therefore

$$(\mathcal{H}^1 \times \mathcal{H}^2)(F) \lesssim c_\alpha(\mu)^{-p_{n-1}} R^{-\beta p_{n-1}} \int_0^{2\pi} \int_{\bigcup_{D \in \mathbb{D}_\theta} D} |\pi_{\theta\sharp}\mu|^{p_{n-1}} \, d\mathcal{H}^2 \, d\theta.$$

By the assumed (4-3) (which also holds trivially without the $R^{-\kappa}$ decay if $m \in \{-1, 0\}$) this gives

$$(\mathcal{H}^1 \times \mathcal{H}^2)(F) \lesssim \mu(\mathbb{R}^3) c_\alpha(\mu)^{-1} R^{-\beta p_{n-1}}.$$

To include the case $n = \infty$, one can replace the above by a slightly weaker estimate

$$(\mathcal{H}^1 \times \mathcal{H}^2)(F) \lesssim \mu(\mathbb{R}^3) c_\alpha(\mu)^{-1} R^{\delta - \beta p_{n-1}}, \tag{4-5}$$

which, when $n = \infty$, follows by rerunning the above argument with p_{n-1} replaced by p_m for some sufficiently large finite m $\big($more precisely, taking $m < \infty$ such that $p_m > \frac{3}{2} - \frac{\delta}{\beta}\big)$. Here, $n - 1 := \infty$ if $n = \infty$. Since the measure $\mathcal{H}^2$ on $\mathbb{R}^3$ is not σ-finite, to avoid issues of applying Fubini's theorem above and below we identify the measure $\mathcal{H}^1 \times \mathcal{H}^2$ on $[0, 2\pi) \times \mathbb{R}^3$ with that obtained by integrating first in $x \in \gamma(\theta)^\perp$ then in $\theta \in [0, 2\pi)$ (rather than in the other order; we could alternatively rewrite the measure as a measure on $[0, 2\pi) \times \mathbb{R}^2$ by using coordinates).

Note that the above is the only place in the proof that the assumption (4-3) is used. From now on, estimate (4-5) of the integration domain is all we need.

Let J be such that $2^J \sim R^\epsilon$, where $\epsilon > 0$ is chosen so that $\delta \ll \epsilon \ll \min\{\beta, \alpha - 2\}$. Given (j, k) and $\tau \in \Lambda_{j,k}$, let

$$\mathbb{T}_{\tau,b} = \big\{T \in \mathbb{T}_\tau : \mu(4T) \geq 2^{-k(3/2-\beta_0)} 2^{-(j-k)\alpha + 1000 j\epsilon} c_\alpha(\mu)\big\}, \quad \mathbb{T}_{\tau,g} = \mathbb{T}_\tau \setminus \mathbb{T}_{\tau,b}, \tag{4-6}$$

and define the bad and good part of the measure as

$$\mu_b = \sum_{j \geq J} \sum_{k \in [j\epsilon, j]} \sum_{\tau \in \Lambda_{j,k}} \sum_{T \in \mathbb{T}_{\tau,b}} M_T \mu, \quad \mu_g = \mu - \mu_b. \tag{4-7}$$

Note that by the trivial upper bound on each $\mu(4T)$ implied by the Frostman condition of μ, $\mathbb{T}_{\tau,b}$ is empty whenever $\tau \in \Lambda_{j,k}$ and $0 \le k \le j\epsilon$, so the range of k in the definition of μ_b in (4-7) could be replaced by $k \in [0, j]$. This will be used later.

By the triangle inequality, followed by Hölder's inequality (using (4-5)),

$$\int_0^{2\pi} \int_{\bigcup_{D\in\mathbb{D}_\theta}} |\pi_{\theta\sharp}|^{p_n}\, d\mathcal{H}^2\, d\theta \lesssim \int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_b|^{p_n}\, d\mathcal{H}^2\, d\theta + \left(\frac{\mu(\mathbb{R}^3)}{c_\alpha(\mu)}\right)^{1-p_n/2} R^{\delta-\beta p_{n-1}(1-p_n/2)} \left(\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_g|^2\, d\mathcal{H}^2\, d\theta\right)^{p_n/2}. \quad (4\text{-}8)$$

In what follows, negligible tail terms of the form $R^{-N}\mu(\mathbb{R}^3)$ (for arbitrarily large N) will be left out, since whenever such terms dominate, the desired inequalities follow trivially. In particular, this means that it suffices to consider scales $J \le j \lesssim \log R$.

We first estimate the bad part, which doesn't rely on (4-5). By the triangle inequality, followed by Hölder's inequality again (noting that $j \lesssim \log R$),

$$\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_b|^{p_n}\, d\mathcal{H}^2\, d\theta \lesssim (\log R)^{O(1)} \sum_{j\ge J} \sum_{k\in[j\epsilon, j]} \int_0^{2\pi}\int \left| \sum_{\tau\in\Lambda_{j,k}} \sum_{T\in\mathbb{T}_{\tau,b}} \pi_{\theta\sharp} M_T\mu \right|^{p_n} d\mathcal{H}^2\, d\theta. \quad (4\text{-}9)$$

Fix a pair (j, k) occurring in the above sum. We will prove the following:

Claim.

$$\int_0^{2\pi}\int \left| \sum_{\tau\in\Lambda_{j,k}} \sum_{T\in\mathbb{T}_{\tau,b}} \pi_{\theta\sharp} M_T\mu \right|^{p_n} d\mathcal{H}^2\, d\theta \lesssim 2^{kO(\delta)} \int_0^{2\pi} \sum_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|\le 2^{k(\delta-1/2)}}} \sum_{T\in\mathbb{T}_{\tau,b}} \mu(2T)^{p_n} (m(T)2^{j-k})^{-(p_n-1)}\, d\theta, \quad (4\text{-}10)$$

where θ_τ is defined by $\tau = \tau(\theta_\tau, j, k)$, and $m(T)$ is the Lebesgue measure of T.

Proof of the Claim. In order to prove (4-10), we will sometimes need to make use of the intermediate inequality, for any $\theta \in [0, 2\pi]$ with $|\theta_\tau - \theta| \le 2^{k(\delta-1/2)}$,

$$\|\pi_{\theta\sharp} M_T\mu\|_{L^{p_n}(\mathcal{H}^2)}^{p_n} \lesssim 2^{kO(\delta)} \mu(2T)^{p_n} (m(T)2^{j-k})^{-(p_n-1)}. \quad (4\text{-}11)$$

The inequality (4-11) holds up to a rapidly decaying tail, and is a consequence of the uncertainty principle (see Bernstein's inequality in [Wolff 2003, Proposition 5.3]). The idea is that, since each $M_T\mu$ is a wave packet that is concentrated on T and locally constant, up to a rapidly decaying tail, one has $\|\pi_{\theta\sharp} M_T\mu\|_{L^{p_n}}^{L^{p_n}} \sim \|\pi_{\theta\sharp} M_T\mu\|_{L^1}^{p_n} m(\pi_\theta(T))^{1-p_n}$. Since $\|\pi_{\theta\sharp} M_T\mu\|_{L^1} \le \|M_T\mu\|_{L^1}$, one can then use Lemma 4 of [Gan et al. 2022] to get $\|\pi_{\theta\sharp} M_T\mu\|_{L^1} \lesssim 2^{kO(\delta)}\mu(2T)$. The condition $|\theta_\tau - \theta| \le 2^{k(\delta-1/2)}$ implies that $\pi_\theta(T)$ is contained in a rectangle of area $\lesssim 2^{kO(\delta)} 2^{j-k} m(T)$, and putting this all together yields (4-11).

An elementary way to obtain (4-11) which avoids using Lemma 4 of [Gan et al. 2022] or Bernstein's inequality, is to use the definition $M_T\mu = \eta_T(\mu * \check{\psi}_\tau)$ and the definition of π_θ to get

$$\|\pi_{\theta\sharp} M_T\mu\|_{L^\infty} \lesssim 2^{k-j+kO(\delta)} \|M_T\mu\|_{L^\infty} \lesssim 2^{k-j+kO(\delta)} \|\mu * \check{\psi}_\tau\|_{L^\infty(T)}.$$

If $x \in T$, then apart from a negligible error term, since $\check{\psi}_\tau$ is negligible outside the plank parallel to T centred at the origin,

$$|\mu * \check{\psi}_\tau(x)| \le \int |\check{\psi}_\tau(x-y)|\, d\mu(y) \lesssim m(\tau)\mu(2T) \lesssim 2^{kO(\delta)} m(T)^{-1}\mu(2T).$$

Hence

$$\|\pi_{\theta\sharp} M_T\mu\|_{L^\infty} \lesssim 2^{k-j+kO(\delta)} m(T)^{-1}\mu(2T). \tag{4-12}$$

Again the condition $|\theta_\tau - \theta| \le 2^{k(\delta-1/2)}$ implies that $\pi_\theta(T)$ is contained in a rectangle with

$$\text{area} \lesssim 2^{kO(\delta)} 2^{j-k} m(T),$$

so substituting (4-12) into the left-hand side of (4-11) yields

$$\|\pi_{\theta\sharp} M_T\mu\|^{p_n}_{L^{p_n}(\mathcal{H}^2)}\, d\theta \lesssim 2^{kO(\delta)} 2^{j-k} m(T) (2^{k-j} m(T)^{-1}\mu(2T))^{p_n} \lesssim 2^{kO(\delta)} \mu(2T)^{p_n} (2^{k-j} m(T)^{-1})^{p_n - 1},$$

which matches (4-11).

We now prove (4-10). First, note that the left-hand side of (4-10) is

$$\lesssim \int_0^{2\pi} \int \Bigg| \sum_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|\le 2^{k(\delta-1/2)}}} \sum_{T\in\mathbb{T}_{\tau,b}} \pi_{\theta\sharp} M_T\mu \Bigg|^{p_n} d\mathcal{H}^2\, d\theta + \int_0^{2\pi} \int \Bigg| \sum_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|> 2^{k(\delta-1/2)}}} \sum_{T\in\mathbb{T}_{\tau,b}} \pi_{\theta\sharp} M_T\mu \Bigg|^{p_n} d\mathcal{H}^2\, d\theta. \tag{4-13}$$

The second term is negligible by the trivial L^∞ bound on each $\pi_{\theta\sharp} M_T\mu$, followed by the same stationary phase bound used in the L^1 case (see Lemma 5 of [Gan et al. 2022]). So it suffices to only consider τ such that θ_τ is close to θ.

Assume first that $k = j$. By the $\lesssim 2^{kO(\delta)}$-overlapping property of the $\pi_\theta(T)$'s (this property uses $j = k$), one gets the bound

$$\int_0^{2\pi} \int \Bigg| \sum_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|\le 2^{k(\delta-1/2)}}} \sum_{T\in\mathbb{T}_{\tau,b}} \pi_{\theta\sharp} M_T\mu \Bigg|^{p_n} d\mathcal{H}^2\, d\theta \lesssim 2^{kO(\delta)} \int_0^{2\pi} \sum_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|\le 2^{k(\delta-1/2)}}} \sum_{T\in\mathbb{T}_{\tau,b}} \|\pi_{\theta\sharp} M_T\mu\|^{p_n}_{L^{p_n}(\mathcal{H}^2)}\, d\theta,$$

and then (4-10) follows by applying (4-11) to the above.

Now consider the case $k < j$. If p_n were replaced by $p = 2$ then (4-10) would follow by orthogonality. More precisely, by applying Plancherel in two dimensions, and then changing variables to an integral over $\mathbb{R}^3$ (fixing (j, k) ensures that the Jacobian of the change of variables is the constant $2^{-j+k/2}$; we

use (4-18) below and the straightforward observation that if $\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta)$ intersects $\tau\in\Lambda_{j,k}$, then $|\eta_1|\sim 2^{j-k/2}$, whose justification can be found in [Gan et al. 2022, equation 106]),

$$\int_0^{2\pi}\int\left|\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,b}}\pi_{\theta\sharp}M_T\mu\right|^2 d\mathcal{H}^2\,d\theta\lesssim 2^{-j+k/2}\int_{\mathbb{R}^3}\left|\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,b}}M_T\mu(x)\right|^2 dx. \tag{4-14}$$

We then apply Plancherel's theorem in $\mathbb{R}^3$ to decouple the L^2 norms (as in [Gan et al. 2022, equation 113]), to bound the above by

$$2^{-j+k/2}\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,b}}\int_{\mathbb{R}^3}|M_T\mu(x)|^2\,dx.$$

Each $M_T\mu$ is just a wave packet, so it is easy to bound the above by

$$2^{-j+k/2+kO(\delta)}\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,b}}\mu(2T)^2m(T)^{-1},$$

and by Fubini, the above is comparable to the right-hand side of (4-10) if p_n is replaced by $p=2$.

If p_n were replaced by $p=1$, then (4-10) would follow by applying the triangle inequality and then Lemma 4 of [Gan et al. 2022] to the first term in (4-13), and again the stationary phase bound from Lemma 5 of [Gan et al. 2022] to the second term in (4-13). This proves (4-10) if $p=1$.

The arguments used above to prove (4-10) when p_n is replaced by $p\in\{1,2\}$ apply similarly if the collections of T's in the left-hand side is replaced by any subcollection, so (4-10) holds more generally for any subcollection of T's in the left-hand side, if p_n is replaced by $p\in\{1,2\}$.

The inequality (4-10) for the exponent $p=p_n$ will be shown via the following application of the interpolation argument from [Demeter 2020, Exercise 9.21]. By pigeonholing, there exists a subcollection $\mathbb{T}$ of T's in the left-hand side of (4-10), such that $\mu(2T)$ is constant over $T\in\mathbb{T}$ up to a factor of 2, and such that

$$\int_0^{2\pi}\int\left|\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,b}}\pi_{\theta\sharp}M_T\mu\right|^{p_n} d\mathcal{H}^2\,d\theta\lesssim 2^{kO(\delta)}\int_0^{2\pi}\int\left|\sum_{T\in\mathbb{T}}\pi_{\theta\sharp}M_T\mu\right|^{p_n} d\mathcal{H}^2\,d\theta.$$

If we write $p_n=(1-\phi)\cdot 1+\phi\cdot 2$ and apply Hölder's inequality with $q=1/(1-\phi)$ and $q'=1/\phi$, this bounds the above by

$$2^{kO(\delta)}\left(\int_0^{2\pi}\int\left|\sum_{T\in\mathbb{T}}\pi_{\theta\sharp}M_T\mu\right| d\mathcal{H}^2\,d\theta\right)^{1-\phi}\left(\int_0^{2\pi}\int\left|\sum_{T\in\mathbb{T}}\pi_{\theta\sharp}M_T\mu\right|^2 d\mathcal{H}^2\,d\theta\right)^{\phi}.$$

If we apply (4-10) in the cases $p=1$ and $p=2$ (with the sum in left-hand side replaced by a sum over $\mathbb{T}$), the above is bounded by

$$2^{kO(\delta)}\left(\int_0^{2\pi}\sum_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|\le 2^{k(\delta-1/2)}}}\sum_{T\in\mathbb{T}_{\tau,b}\cap\mathbb{T}}\mu(2T)\,d\theta\right)^{1-\phi}\times\left(\int_0^{2\pi}\sum_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|\le 2^{k(\delta-1/2)}}}\sum_{T\in\mathbb{T}_{\tau,b}\cap\mathbb{T}}\mu(2T)^2(m(T)2^{j-k})^{-1}\,d\theta\right)^{\phi}.$$

By Fubini's theorem and the essentially constant property of $\mu(2T)$, the above is bounded by

$$2^{kO(\delta)}(|\mathbb{T}|\mu(2T)2^{-k/2})^{1-\phi}\big(|\mathbb{T}|\mu(2T)^2(m(T)2^{j-k})^{-1}2^{-k/2}\big)^{\phi}.$$

This simplifies to

$$2^{kO(\delta)}|\mathbb{T}|\mu(2T)^{p_n}2^{-k/2}(m(T)2^{j-k})^{-(p_n-1)},$$

which, by Fubini's theorem again, is bounded by

$$2^{kO(\delta)}\int_0^{2\pi}\sum_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|\le 2^{k(\delta-1/2)}}}\sum_{T\in\mathbb{T}_{\tau,b}}\mu(2T)^{p_n}(m(T)2^{j-k})^{-(p_n-1)}\,d\theta.$$

This matches the right-hand side of (4-10), so this finishes the proof of (4-10). □

We continue with the proof of the theorem, and the bound on the "bad" part. The right-hand side of (4-10) is bounded by

$$2^{kO(\delta)}\int_0^{2\pi}\sum_m\sum_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|\le 2^{k(\delta-1/2)}}}\sum_{\substack{T\in\mathbb{T}_{\tau,b}:\\ T\cap B_m\neq\varnothing}}\mu(2T)^{p_n}(m(T)2^{j-k})^{-(p_n-1)}\,d\theta, \tag{4-15}$$

where, for each pair (j,k), $\{B_m\}_m$ is a boundedly overlapping cover of $\mathbb{R}^3$ by balls of radius $2^{-(j-k)+k\delta}$. The size of the balls B_m is chosen so that, for any fixed B_m, after a rescaling later, the badness of planks T (intersecting B_m) implies the β_0-admissibility, i.e., the rescaled version of the measure μ restricted on B_m and the rescaled planks satisfy the hypotheses for (4-4).

For fixed (j,k), m, τ, and a T occurring in the sum above, recalling that the shortest sidelength of the plank T is $2^{k\delta-j}$, so for any $\theta\in[0,2\pi)$,

$$\mu(2T)^{p_n}(m(T)2^{j-k})^{-(p_n-1)}\lesssim\left(\int_{\pi_\theta(2T)}\big|\pi_{\theta\sharp}(\mu_{m,j,k}*\phi_{2^{j-k\delta}})\big|\,d\mathcal{H}^2\right)^{p_n}(m(T)2^{j-k})^{-(p_n-1)},$$

where $\mu_{m,j,k}$ is the restriction of μ to $100B_m$. By Hölder's inequality applied to the above,

$$\mu(2T)^{p_n}(m(T)2^{j-k})^{-(p_n-1)}\lesssim 2^{kO(\delta)}\int_{\pi_\theta(2T)}|\pi_{\theta\sharp}(\mu_{m,j,k}*\phi_{2^{j-k\delta}})|^{p_n}\,d\mathcal{H}^2,\quad\text{when } |\theta-\theta_\tau|\le 2^{k(\delta-1/2)}.$$

If we substitute this inequality into (4-15), and then (4-15) into (4-10), and then sum (4-10) over j and k in (4-9), we get

$$\begin{aligned}&\int_0^{2\pi}\int|\pi_{\theta\sharp}\mu_b|^{p_n}\,d\mathcal{H}^2\,d\theta\\ &\quad\lesssim(\log R)^{O(1)}\sum_{j\ge J}\sum_{k\in[j\epsilon,j]}2^{kO(\delta)}\sum_m\int_0^{2\pi}\sum_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|\le 2^{k(\delta-1/2)}}}\sum_{\substack{T\in\mathbb{T}_{\tau,b}:\\ T\cap B_m\neq\varnothing}}\int_{\pi_\theta(2T)}|\pi_{\theta\sharp}(\mu_{m,j,k}*\phi_{2^{j-k\delta}})|^{p_n}\,d\mathcal{H}^2\,d\theta.\end{aligned}$$

The rectangles $\pi_\theta(2T)$ in the sum above are $\lesssim 2^{kO(\delta)}$-overlapping over the two innermost sums, so the above implies that

$$\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_b|^{p_n}\,d\mathcal{H}^2\,d\theta \lesssim (\log R)^{O(1)}\sum_{j\geq J}\sum_{k\in[j\epsilon,j]} 2^{kO(\delta)}\sum_m \int_0^{2\pi}\int_{E_{m,j,k}(\theta)} |\pi_{\theta\sharp}(\mu_{m,j,k}*\phi_{2^{j-k\delta}})|^{p_n}\,d\mathcal{H}^2\,d\theta,$$

where

$$E_{m,j,k}(\theta) := \bigcup_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|\leq 2^{k(\delta-1/2)}}}\ \bigcup_{\substack{T\in\mathbb{T}_{\tau,b}:\\ T\cap B_m\neq\varnothing}} \pi_\theta(2T).$$

We will abbreviate $\nu_{m,j,k}=\mu_{m,j,k}*\phi_{2^{j-k\delta}}$ to simplify the notation. If

$$F_{m,j,k}(\theta) := 2^{j-k-k\delta}E_{m,j,k}(\theta) = \bigcup_{\substack{\tau\in\Lambda_{j,k}:\\ |\theta_\tau-\theta|\leq 2^{k(\delta-1/2)}}}\ \bigcup_{\substack{T\in\mathbb{T}_{\tau,b}:\\ T\cap B_m\neq\varnothing}} \pi_\theta(2^{j-k-k\delta}2T),$$

then by rescaling each ball B_m to a ball of radius 1, using that orthogonal projections commute with rescalings,

$$\begin{aligned}\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_b|^{p_n}\,d\mathcal{H}^2\,d\theta&\\ \lesssim (\log R)^{O(1)}\sum_{j\geq J}\sum_{k\in[j\epsilon,j]} 2^{kO(\delta)}\sum_m 2^{2(j-k)(p_n-1)}&\int_0^{2\pi}\int_{F_{m,j,k}(\theta)} |\pi_{\theta\sharp}2^{j-k-k\delta}_\sharp \nu_{m,j,k}|^{p_n}\,d\mathcal{H}^2\,d\theta.\end{aligned}$$

Each rescaled measure $2^{j-k-k\delta}_\sharp\nu_{m,j,k}$ and collection of rescaled planks $2^{j-k-k\delta}2T$ satisfy the hypotheses for (4-4); by the definition of β_0 and of the "bad" planks. Moreover,

$$c_\alpha(2^{j-k-k\delta}_\sharp\nu_{m,j,k}) \leq 2^{-\alpha(j-k)+kO(\delta)}c_\alpha(\mu).$$

Hence, by the conclusion of (4-4) for β_0 and the assumption that $\alpha>2$,

$$\begin{aligned}\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_b|^{p_n}\,d\mathcal{H}^2\,d\theta &\lesssim (\log R)^{O(1)}\sum_{j\geq J}\sum_{k\in[j\epsilon,j]}\sum_m \mu(100B_m)2^{kO(\delta)}c_\alpha(\mu)^{p_n-1}2^{-k\kappa}\\ &\lesssim 2^{-J\kappa/2}\mu(\mathbb{R}^3)c_\alpha(\mu)^{p_n-1}\lesssim R^{-\epsilon\kappa/100}\mu(\mathbb{R}^3)c_\alpha(\mu)^{p_n-1}.\end{aligned}$$

It remains to bound the "good" part. The L^2 method we use is similar to the second part of [Gan et al. 2022], but we include all details for completeness. By (4-8), it suffices to show that

$$\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_g|^2\,d\mathcal{H}^2\,d\theta \lesssim \mu(\mathbb{R}^3)c_\alpha(\mu)R^{-100\delta+p_{n-1}\beta(2/p_n-1)}. \tag{4-16}$$

By Plancherel's theorem in two dimensions and the projection-slice theorem,

$$\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_g|^2\,d\mathcal{H}^2\,d\theta = 2\int_0^{2\pi}\int \left|\widehat{\mu}_g(\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta))\right|^2 d\eta_1\,d\eta_2\,d\theta.$$

Therefore, to prove (4-16), it suffices to show that

$$\int_0^{2\pi}\int_{\{|\eta|\geq 2^{2J}\}} \left|\widehat{\mu}_g(\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta))\right|^2 d\eta_1\, d\eta_2\, d\theta \lesssim \mu(\mathbb{R}^3)c_\alpha(\mu)R^{-100\delta+p_{n-1}\beta(2/p_n-1)};$$

since the part near the origin can be handled by a trivial bound on $\|\widehat{\mu}_g\|_{L^\infty(B(0,2^{2J})}$ by

$$\|\widehat{\mu}\|_{L^\infty}+\|\widehat{\mu}_b\|_{L^\infty(B(0,2^{2J}))}\lesssim 2^{O(1)J}\lesssim R^{O(1)\epsilon}$$

(this is where we use that $\epsilon\ll\beta$). We break the left-hand side above into two integrals:

$$\int_0^{2\pi}\int_{\{|\eta_1|\geq|\eta|^{1-\epsilon}\}\cap\{|\eta|\geq 2^{2J}\}} \left|\widehat{\mu}_g(\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta))\right|^2 d\eta_1\, d\eta_2\, d\theta + \int_0^{2\pi}\int_{\{|\eta_1|\leq|\eta|^{1-\epsilon}\}\cap\{|\eta|\geq 2^{2J}\}} \left|\widehat{\mu}_g(\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta))\right|^2 d\eta_1\, d\eta_2\, d\theta.$$

For the first integral, $\widehat{\mu}_b$ is negligible on the domain of integration, since $\mathbb{T}_{\tau,b}$ is empty whenever $\tau\in\Lambda_{j,k}$ and $k\leq 10j\epsilon$ (as mentioned previously, this follows from the definition of $\mathbb{T}_{\tau,b}$ and the fractal property of μ). Therefore, to bound the first integral, it suffices to prove that

$$\int_0^{2\pi}\int_{\{|\eta_1|\geq|\eta|^{1-\epsilon}\}\cap\{|\eta|\geq 2^{2J}\}} \left|\widehat{\mu}(\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta))\right|^2 d\eta_1\, d\eta_2\, d\theta \lesssim \mu(\mathbb{R}^3)c_\alpha(\mu). \tag{4-17}$$

The change of variables

$$(\xi_1,\xi_2,\xi_3)=\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta)$$

has Jacobian

$$\left|\frac{\partial(\eta_1,\eta_2,\theta)}{\partial(\xi_1,\xi_2,\xi_3)}\right|\sim|\eta_1|^{-1}, \tag{4-18}$$

so the integral in (4-17) is bounded by

$$\int|\xi|^{\epsilon-1}|\widehat{\mu}(\xi)|^2\,d\xi.$$

By the Fourier formula for the energy (see, e.g., [Mattila 2015, Theorem 3.10]), the above is comparable to $I_{2+\epsilon}(\mu)$, which is bounded by $\mu(\mathbb{R}^3)c_\alpha(\mu)$ since $\alpha>2+\epsilon$ (see for instance the inequality at the end of page 19 in [Mattila 2015]).

Therefore, it remains to prove that

$$\int_0^{2\pi}\int_{\{|\eta_1|\leq|\eta|^{1-\epsilon}\}\cap\{|\eta|\geq 2^{2J}\}} \left|\widehat{\mu}_g(\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta))\right|^2 d\eta_1\, d\eta_2\, d\theta \lesssim \mu(\mathbb{R}^3)c_\alpha(\mu)R^{-100\delta+p_{n-1}\beta(2/p_n-1)}. \tag{4-19}$$

As remarked above, that $\mathbb{T}_{\tau,b}$ is empty whenever $k\leq 10j\epsilon$ implies that the definition of μ_b in (4-7) can be written as

$$\mu_b=\sum_{j\geq J}\sum_{k\in[0,j]}\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,b}} M_T\mu,\quad \mu_g=\mu-\mu_b.$$

It follows that on the domain of integration in (4-19), μ_g can be replaced by the sum over the good planks (ignoring rapidly decaying error terms), and it suffices to prove that

$$\int_0^{2\pi}\int_{\{|\eta_1|\le|\eta|^{1-\epsilon}\}\cap\{|\eta|\ge 2^{2J}\}}\left|\sum_{j\ge J}\sum_{k\in[j\epsilon/1000,\,j]}\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,g}}\widehat{M_T\mu}(\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta))\right|^2 d\eta_1\,d\eta_2\,d\theta \lesssim \mu(\mathbb{R}^3)c_\alpha(\mu)R^{-100\delta+p_{n-1}\beta(2/p_n-1)}.$$

By disjointness, it is enough to show that

$$\sum_{j\ge J}\sum_{k\in[j\epsilon/1000,\,j]}\sum_{\tau\in\Lambda_{j,k}}\int_0^{2\pi}\int_{\{|\eta_1|\le|\eta|^{1-\epsilon}\}\cap\{|\eta|\ge 2^{2J}\}}\left|\sum_{T\in\mathbb{T}_{\tau,g}}\widehat{M_T\mu}(\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta))\right|^2 d\eta_1\,d\eta_2\,d\theta \lesssim \mu(\mathbb{R}^3)c_\alpha(\mu)R^{-100\delta+p_{n-1}\beta(2/p_n-1)}.$$

If $k<j$ and $\tau\in\Lambda_{j,k}$, then by a change of variables and the formula (4-18),

$$\int_0^{2\pi}\int_{\{|\eta_1|\le|\eta|^{1-\epsilon}\}\cap\{|\eta|\ge 2^{2J}\}}\left|\sum_{T\in\mathbb{T}_{\tau,g}}\widehat{M_T\mu}(\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta))\right|^2 d\eta_1\,d\eta_2\,d\theta \lesssim 2^{-j+k/2}\int_{\mathbb{R}^3}\left|\sum_{T\in\mathbb{T}_{\tau,g}}\widehat{M_T\mu}(\xi)\right|^2 d\xi,$$

where we used that if $\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta)$ intersects $\tau\in\Lambda_{j,k}$, then $|\eta_1|\sim 2^{j-k/2}$. We then use essential orthogonality of wave packets to get

$$\int_0^{2\pi}\int_{\{|\eta_1|\le|\eta|^{1-\epsilon}\}\cap\{|\eta|\ge 2^{2J}\}}\left|\sum_{T\in\mathbb{T}_{\tau,g}}\widehat{M_T\mu}(\eta_1\gamma'(\theta)+\eta_2(\gamma\times\gamma')(\theta))\right|^2 d\eta_1\,d\eta_2\,d\theta \lesssim 2^{-j+k/2+kO(\delta)}\sum_{T\in\mathbb{T}_{\tau,g}}\int_{\mathbb{R}^3}|\widehat{M_T\mu}(\xi)|^2\,d\xi. \quad (4\text{-}20)$$

The above inequality also holds if $k=j$ by the following argument. Use Plancherel's theorem in two dimensions to change the integral in the left-hand side of (4-20) back to an integral of $\sum_{T\in\mathbb{T}_{\tau,g}}\pi_{\theta\sharp}M_T\mu$. The only significant θ in the integration over θ satisfies $|\theta-\theta_\tau|\lesssim 2^{k(-1/2+\delta)}$, so the sets $\pi_\theta(T)$ are $\lesssim 2^{kO(\delta)}$-overlapping. This means that the L^2 norms can be decoupled, and then (4-20) follows by a calculation using that each $M_T\mu$ is a wave packet.

Therefore, it suffices to prove that

$$\sum_{j\ge J}\sum_{k\in[\epsilon j/1000,\,j]}2^{-j+k/2+kO(\delta)}\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,g}}\|M_T\mu\|_2^2\lesssim\mu(\mathbb{R}^3)c_\alpha(\mu)R^{-100\delta+p_{n-1}\beta(2/p_n-1)}. \quad (4\text{-}21)$$

This will be shown via

$$\sum_{j\ge J}\sum_{k\in[\epsilon j/1000,\,j]}2^{-j+k/2+kO(\delta)}\sum_{m}\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,m,g}}\|M_T\mu\|_2^2\lesssim\mu(\mathbb{R}^3)c_\alpha(\mu)R^{-100\delta+p_{n-1}\beta(2/p_n-1)},$$

where, for each (j,k), $\{B_m\}$ is a boundedly overlapping covering of $\mathbb{R}^3$ by balls of radius $2^{-(j-k)+k\delta}$, and where $\mathbb{T}_{\tau,g,m}=\{T\in\mathbb{T}_{\tau,g}:T\cap B_m\ne\varnothing\}$. Fix j, k, and m. By writing out the L^2 norm as an inner product

and expanding out the definition of one of the $M_T\mu$'s, the innermost double sum can be written as

$$\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,g,m}}\|M_T\mu\|_2^2=\int\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,g,m}}(\eta_T M_T\mu)*\check{\psi}_\tau\,d\mu. \tag{4-22}$$

By Cauchy–Schwarz,

$$(4\text{-}22)\le\mu(100B_m)^{1/2}\bigg(\int\bigg|\sum_{\tau\in\Lambda_{j,k}}\sum_{T\in\mathbb{T}_{\tau,g,m}}(\eta_T M_T\mu)*\check{\psi}_\tau\bigg|^2 d(\mu*\phi_{2^j})\bigg)^{1/2}, \tag{4-23}$$

where we used the Fourier support of the integrand in (4-22) to replace μ by its smoothed out version $\mu*\phi_{2^j}$ at the scale 2^{-j}. Let $p_{\mathrm{dec}}\in(2,6]$ be an exponent to be chosen. We abbreviate $f_T=(\eta_T M_T\mu)*\check{\psi}_\tau$. By dyadic pigeonholing, we can find a subset $\mathbb{W}\subseteq\bigcup_{\tau\in\Lambda_{j,k}}\mathbb{T}_{\tau,g,m}$ such that $\|f_T\|_{p_{\mathrm{dec}}}$ is constant over $T\in\mathbb{W}$ up to a factor of 2, such that

$$(4\text{-}23)\le O(j)\mu(100B_m)^{1/2}\bigg(\int\bigg|\sum_{T\in\mathbb{W}}f_T\bigg|^2 d(\mu*\phi_{2^j})\bigg)^{1/2}. \tag{4-24}$$

By partitioning the domain of integration into boundedly overlapping balls of radius 2^{-j}, we can find a dyadic number M and a union Y of balls of radius 2^{-j} such that each ball $B\subseteq Y$ intersects $\sim M$ planks $2T$ with $T\in\mathbb{W}$, and such that

$$(4\text{-}24)\lesssim O(j)\mu(100B_m)^{1/2}\bigg(\int_Y\bigg|\sum_{T\in\mathbb{W}}f_T\bigg|^2 d(\mu*\phi_{2^j})\bigg)^{1/2}. \tag{4-25}$$

By Hölder's inequality with respect to Lebesgue measure,

$$\bigg(\int_Y\bigg|\sum_{T\in\mathbb{W}}f_T\bigg|^2 d(\mu*\phi_{2^j})\bigg)^{1/2}\le\bigg(\int_Y\bigg|\sum_{T\in\mathbb{W}}f_T\bigg|^{p_{\mathrm{dec}}}\bigg)^{1/p_{\mathrm{dec}}}\bigg(\int_Y(\mu*\phi_{2^j})^{p_{\mathrm{dec}}/p_{\mathrm{dec}}-2}\bigg)^{1/2-1/p_{\mathrm{dec}}}, \tag{4-26}$$

By a rescaling, followed by an application of the refined decoupling inequality (Theorem A.1), the first factor in the right-hand side of (4-26) satisfies

$$\bigg(\int_Y\bigg|\sum_{T\in\mathbb{W}}f_T\bigg|^{p_{\mathrm{dec}}}\bigg)^{1/p_{\mathrm{dec}}}\lesssim 2^{k\epsilon}\bigg(\frac{M}{|\mathbb{W}|}\bigg)^{1/2-1/p_{\mathrm{dec}}}\bigg(\sum_{T\in\mathbb{W}}\|f_T\|_{p_{\mathrm{dec}}}^2\bigg)^{1/2}. \tag{4-27}$$

By the definition of M, and the definition of "bad" planks in (4-6), the second factor in the right-hand side of (4-26) satisfies

$$\begin{aligned}\bigg(\int_Y(\mu*\phi_{2^j})^{p_{\mathrm{dec}}/p_{\mathrm{dec}}-2}\bigg)^{1/2-1/p_{\mathrm{dec}}}&\lesssim\bigg(\int_Y d(\mu*\phi_{2^j})\bigg)^{1/2-1/p_{\mathrm{dec}}}2^{j(3-\alpha)/p_{\mathrm{dec}}}c_\alpha(\mu)^{1/p_{\mathrm{dec}}}\\&\lesssim\bigg(\frac{1}{M}\sum_{T\in\mathbb{W}}\int_{Y\cap T}d(\mu*\phi_{2^j})\bigg)^{1/2-1/p_{\mathrm{dec}}}2^{j(3-\alpha)/p_{\mathrm{dec}}}c_\alpha(\mu)^{1/p_{\mathrm{dec}}}\end{aligned}\qquad\text{(cont.)}$$

$$\lesssim \left(\frac{|\mathbb{W}|}{M} 2^{-\alpha(j-k)-k(3/2-\beta_0)+1000j\epsilon} c_\alpha(\mu)\right)^{1/2-1/p_{\mathrm{dec}}} 2^{j(3-\alpha)/p_{\mathrm{dec}}} c_\alpha(\mu)^{1/p_{\mathrm{dec}}}$$

$$= \left(\frac{|\mathbb{W}|}{M}\right)^{1/2-1/p_{\mathrm{dec}}} 2^{(1/2-1/p_{\mathrm{dec}})\left(-\alpha(j-k)-k(3/2-\beta_0)+1000j\epsilon\right)+j(3-\alpha)/p_{\mathrm{dec}}} c_\alpha(\mu)^{1/2}. \quad (4\text{-}28)$$

Substituting (4-27) and (4-28) into (4-26) yields

$$\left(\int_Y \left|\sum_{T\in\mathbb{W}} f_T\right|^2 d\mu\right)^{1/2} \lesssim 2^{k\epsilon+(1/2-1/p_{\mathrm{dec}})\left(-\alpha(j-k)-k(3/2-\beta_0)+1000j\epsilon\right)+j(3-\alpha)/p_{\mathrm{dec}}} c_\alpha(\mu)^{1/2} \left(\sum_{T\in\mathbb{W}} \|f_T\|_{p_{\mathrm{dec}}}^2\right)^{1/2}. \quad (4\text{-}29)$$

By recalling that $f_T = (\eta_T M_T \mu) * \check{\psi}_\tau$ and applying the Hausdorff–Young inequality, followed by Hölder's inequality,

$$\|f_T\|_{p_{\mathrm{dec}}} \lesssim \|M_T\mu\|_2 2^{(3j-3k/2)(1/2-1/p_{\mathrm{dec}})}.$$

Therefore, (4-29) becomes

$$\left(\int_Y \left|\sum_{T\in\mathbb{W}} f_T\right|^2 d\mu\right)^{1/2} \lesssim 2^{k\epsilon+(1/2-1/p_{\mathrm{dec}})\left(-\alpha(j-k)-k(3/2-\beta_0)+1000j\epsilon\right)+j(3-\alpha)/p_{\mathrm{dec}}+(3j-3k/2)(1/2-1/p_{\mathrm{dec}})} \times c_\alpha(\mu)^{1/2} \left(\sum_{\tau\in\Lambda_{j,k}} \sum_{T\in\mathbb{T}_{\tau,g,m}} \|M_T\mu\|_2^2\right)^{1/2}, \quad (4\text{-}30)$$

where we now sum over all "good" planks in $\mathbb{T}_{\tau,g,m}$ rather than just $\mathbb{W}$, since the property of $\mathbb{W}$ will no longer be needed. By substituting (4-30) into (4-25), then into (4-24), then into (4-23), and then finally into (4-22), we get

$$\sum_{\tau\in\Lambda_{j,k}} \sum_{T\in\mathbb{T}_{\tau,g,m}} \|M_T\mu\|_2^2 \lesssim 2^{jO(\epsilon)+(1/2-1/p_{\mathrm{dec}})\left(-\alpha(j-k)-k(3/2-\beta_0)+1000j\epsilon\right)+j(3-\alpha)/p_{\mathrm{dec}}+(3j-3k/2)(1/2-1/p_{\mathrm{dec}})} \times \mu(100B_m)^{1/2} c_\alpha(\mu)^{1/2} \left(\sum_{\tau\in\Lambda_{j,k}} \sum_{T\in\mathbb{T}_{\tau,g,m}} \|M_T\mu\|_2^2\right)^{1/2}.$$

By cancelling the common factor in both sides, this simplifies to

$$\sum_{\tau\in\Lambda_{j,k}} \sum_{T\in\mathbb{T}_{\tau,g,m}} \|M_T\mu\|_2^2 \lesssim 2^{jO(\epsilon)+j(3-\alpha)-2k(1/2-1/p_{\mathrm{dec}})(3-\alpha-\beta_0)} \mu(100B_m) c_\alpha(\mu).$$

Summing over m gives

$$\sum_{\tau\in\Lambda_{j,k}} \sum_{T\in\mathbb{T}_{\tau,g}} \|M_T\mu\|_2^2 \lesssim 2^{jO(\epsilon)+j(3-\alpha)-2k(1/2-1/p_{\mathrm{dec}})(3-\alpha-\beta_0)} \mu(\mathbb{R}^3) c_\alpha(\mu).$$

If we substitute this bound into the left-hand side of (4-21), we get

$$\sum_{j\geq J} \sum_{k\in[\epsilon j/1000,\, j]} \sum_{\tau\in\Lambda_{j,k}} \sum_{T\in\mathbb{T}_{\tau,g}} 2^{-j+k/2+kO(\delta)} \|M_T\mu\|_2^2 \lesssim \mu(\mathbb{R}^3) c_\alpha(\mu) \sum_{J\leq j\leq(1+\delta)\log_2(R)} \sum_{k\in[\epsilon j/1000,\, j]} 2^{-j+k/2+jO(\epsilon)} 2^{j(3-\alpha)-2k(1/2-1/p_{\mathrm{dec}})(3-\alpha-\beta_0)}. \quad (4\text{-}31)$$

By taking $p_{\mathrm{dec}} = 4$ and summing the two geometric series, (4-31) simplifies to

$$(4\text{-}31) \lesssim \mu(\mathbb{R}^3) c_\alpha(\mu) R^{O(\epsilon)} \max\{1, R^{\beta_0/2-(\alpha-2)/2}\}. \tag{4-32}$$

Using that $\delta \ll \epsilon \ll \min\{\beta, \alpha-2\}$, to verify that (4-32) implies (4-21), it suffices to check that

$$\tfrac{1}{2}\beta_0 \le p_{n-1}\beta\Big(\frac{2}{p_n} - 1\Big) + (\alpha-2)/100.$$

Since $\beta > c\beta_0$, where $1 - c \ll \alpha - 2$, it suffices to check that

$$\tfrac{1}{2} \le p_{n-1}\Big(\frac{2}{p_n} - 1\Big). \tag{4-33}$$

But by (4-2), equality[3] holds in (4-33) for any $n \in \mathbb{N} \cup \{\infty\}$ (using that $p_{-1} = 1$ when $n = 0$). We have verified (4-16), so this finishes the proof. □

We are now ready to prove the following main result of the section, which includes Theorem 1.2 as a special case.

Theorem 4.4. *Let $\alpha > 2$ and $1 \le p \le \frac{3}{2}$. Then*

$$\int_0^{2\pi} \int |\pi_{\theta\sharp}\mu|^p \, d\mathcal{H}^2 \, d\theta \le C_{\alpha,p}\mu(\mathbb{R}^3) c_\alpha(\mu)^{p-1} \tag{4-34}$$

for any Borel measure μ on $B_3(0,1)$ with $c_\alpha(\mu) < \infty$.

Proof. Since (4-34) is trivial if the projections are a.e. absolutely continuous, it suffices to prove (4-34) for $p > 1$. By interpolation (or just the log convexity of L^p norms, or Hölder's inequality), it suffices to prove the endpoint case $p = \frac{3}{2}$. Since $p = \frac{3}{2} > 1$, by a standard limiting argument (which uses that the continuous compactly supported functions are dense in the dual of L^p), it may be assumed that μ is a smooth function.

Take $\beta > 0$ such that $\beta \ll \alpha - 2$, and let $\delta, \epsilon > 0$ be such that $\delta \ll \epsilon \ll \beta$ and $\delta \ll \kappa = \kappa(\beta)$, the parameter in Lemma 4.2. Write $\mu = \mu_b + \mu_g$, where the "bad" part is defined as in (4-7), except that we take $J = 1$ and the parameter β is used instead of β_0 in (4-6). By the triangle inequality, and then by following the same argument as in the proof of Lemma 4.2,

$$\begin{aligned}
&\left(\int_0^{2\pi} \int |\pi_{\theta\sharp}\mu_b|^p \, d\mathcal{H}^2 \, d\theta\right)^{1/p} \\
&\quad \lesssim \sum_{j\ge 1} \sum_{k\in[j\epsilon, j]} \left(O(j) 2^{2(j-k)(p-1)+kO(\delta)} \sum_m \int_0^{2\pi} \int_{F_{m,j,k}(\theta)} \left|\pi_{\theta\sharp} 2^{j-k-k\delta}_\sharp (\mu_{m,j,k} * \phi_{2^{j-k\delta}})\right|^p d\mathcal{H}^2 \, d\theta\right)^{1/p},
\end{aligned}$$

where

$$F_{m,j,k}(\theta) := \bigcup_{\substack{\tau\in\Lambda_{j,k}: \\ |\theta_\tau - \theta| \le 2^{k(\delta-1/2)}}} \bigcup_{\substack{T\in\mathbb{T}_{\tau,b}: \\ T\cap B_m \ne \varnothing}} \pi_\theta(2^{j-k-k\delta} 2T).$$

[3]This step could only be improved by using $p_{\mathrm{dec}} > 4$ if we could assume that β is small, but for β close to 1 there is no gain from taking $p_{\mathrm{dec}} > 4$ in (4-31). From the train-track example in Proposition 3.4, we expect the $\beta = 1$ case to be critical. Equality at $p_{n-1} = p_n = \frac{3}{2}$ explains why the exponent $\frac{3}{2}$ is important and cannot be increased.

By the $p = \frac{3}{2}$ case of Lemma 4.2,

$$\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_b|^p \, d\mathcal{H}^2 \, d\theta \lesssim c_\alpha(\mu)^{p-1}\mu(\mathbb{R}^3)\bigg(\sum_{j\geq 1} 2^{-j\epsilon\kappa/1000}\bigg)^p \lesssim c_\alpha(\mu)^{p-1}\mu(\mathbb{R}^3). \tag{4-35}$$

This bounds the bad part, so it remains to bound the good part.

If we follow the same argument as in Lemma 4.2 to bound the good part, the proof works without substantial changes, and yields

$$\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_g|^2 \, d\mathcal{H}^2 \, d\theta \lesssim I_{2+100\epsilon}(\mu) + \mu(\mathbb{R}^3)c_\alpha(\mu)\sum_{j\geq 1}\sum_{k\in[\epsilon j/1000,\, j]} 2^{-j+k/2+jO(\epsilon)}2^{j(3-\alpha)-2k(1/2-1/p_{\mathrm{dec}})(3-\alpha-\beta)}.$$

By taking $p_{\mathrm{dec}} = 4$ and using $\alpha > 2$ and $\epsilon \ll \beta \ll \alpha - 2$, the second term is a decaying geometric series, which yields

$$\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_g|^2 \, d\mathcal{H}^2 \, d\theta \lesssim \mu(\mathbb{R}^3)c_\alpha(\mu).$$

Note that the argument above for the bad part in (4-35) also works if one sets $p = 1$ and applies the corresponding conclusion in Lemma 4.2. Hence, the above shows that $\pi_{\theta\sharp}\mu_b \in L^1$ for a.e. $\theta \in [0, 2\pi)$, and that

$$\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_b| \, d\mathcal{H}^2 \, d\theta \lesssim \mu(\mathbb{R}^3).$$

By the trivial L^1 estimate of $\pi_{\theta\sharp}\mu$ (using that μ is a smooth function) and the triangle inequality, this gives

$$\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_g| \, d\mathcal{H}^2 \, d\theta \lesssim \mu(\mathbb{R}^3).$$

By writing $\frac{3}{2} = p = (1-\phi)1 + \phi \cdot 2$, where $\phi = p - 1$, and then applying Hölder's inequality to the function

$$|\pi_{\theta\sharp}\mu_g|^p = |\pi_{\theta\sharp}\mu_g|^{(1-\phi)\cdot 1} \cdot |\pi_{\theta\sharp}\mu_g|^{\phi\cdot 2},$$

with $q = \frac{1}{1-\phi}$ and $q' = \frac{1}{\phi}$, this yields

$$\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu_g|^p \, d\mathcal{H}^2 \, d\theta \lesssim c_\alpha(\mu)^{p-1}\mu(\mathbb{R}^3). \tag{4-36}$$

Combining (4-35) and (4-36) gives, by the triangle inequality,

$$\int_0^{2\pi}\int |\pi_{\theta\sharp}\mu|^p \, d\mathcal{H}^2 \, d\theta \lesssim c_\alpha(\mu)^{p-1}\mu(\mathbb{R}^3). \qquad \square$$

Proof of Theorem 1.1. The class of horizontal (or SL_2) lines is invariant under rotations of the plane $\mathbb{R}^2 \times \{0\}$, so by the triangle inequality it may be assumed that all tubes $T \in \mathbb{T}$ have $\ell(T) = p * \mathbb{V}_\theta$ with

$|\theta| \geq \pi/4$, where we use the convention $-\pi/2 < \theta \leq \pi/2$. Recall that $U(\xi_1, \xi_2, \xi_3) = (\eta_1, \eta_2, \eta_3)$ is the orthogonal transformation given by

$$\eta_1 = \frac{\xi_1 + \xi_3}{\sqrt{2}}, \quad \eta_3 = \frac{-\xi_1 + \xi_3}{\sqrt{2}}, \quad \eta_2 = \xi_2.$$

For each $T \in \mathbb{T}$, let w_T be such that $\ell(Uw_T)$ is the centre line of T. By the assumption that $|\theta| \geq \pi/4$ for the angle of the tubes T, and by the definition of ℓ in (3-8), the property that T intersects $B_3(0, 1)$ implies that $|w_T| \lesssim 1$. This ensures that the measure $\mu := \delta^{-1} \sum_{T \in \mathbb{T}} \chi_{B(w_T, C\delta)}$, where C is a large constant to be chosen, is contained in a ball around the origin of radius $\lesssim 1$. If we run the point-line duality argument from (3-12) in reverse, we get

$$\begin{aligned}
\int_{-\pi/2}^{\pi/2} \|\pi_{\theta\sharp}\mu\|^{3/2}_{L^{3/2}(\gamma(\theta)^\perp)}\, d\theta &= \int_{-\pi/2}^{\pi/2} \int_{\gamma(\theta)^\perp} \left| \int_{\pi_\theta^{-1}(y)} \mu \, d\mathcal{H}^1 \right|^{3/2} d\mathcal{H}^2(y)\, d\theta \\
&= \int_{-\pi/2}^{\pi/2} \int_{\gamma(\theta)^\perp} \left| \int_{U\pi_\theta^{-1}(y)} (U_\sharp \mu)\, d\mathcal{H}^1 \right|^{3/2} d\mathcal{H}^2(y)\, d\theta \\
&\sim \int \left| \int_L U_\sharp \mu \, d\mathcal{H}^1 \right|^{3/2} d(\ell^*_\sharp \mathcal{H}^3 \chi_{|x_2| \leq \sqrt{2}})(L) \qquad \text{(by Lemma 3.6)} \\
&\gtrsim \int \left| \{T \in \mathbb{T} : UB(w_T, 100\delta) \cap L \neq \varnothing\} \right|^{3/2} d(\ell^*_\sharp \mathcal{H}^3 \chi_{|x_2| \leq \sqrt{2}})(L) \qquad (4\text{-}37) \\
&= \int_{|p_2^*| \leq \sqrt{2}} \left| \{T \in \mathbb{T} : UB(w_T, 100\delta) \cap \ell^*(p^*) \neq \varnothing\} \right|^{3/2} d\mathcal{H}^3(p^*) \\
&\geq \int_{|p^*| \leq 1} \left| \{T \in \mathbb{T} : UB(w_T, 100\delta) \cap \ell^*(p^*) \neq \varnothing\} \right|^{3/2} d\mathcal{H}^3(p^*) \\
&= \int_{|p^*| \leq 1} \left| \{T \in \mathbb{T} : p^* \in \ell(UB(w_T, 100\delta))\} \right|^{3/2} d\mathcal{H}^3(p^*) \qquad (4\text{-}38) \\
&= \int_{|p^*| \leq 1} \left| \sum_{T \in \mathbb{T}} \chi_{\ell(UB(w_T, 100\delta))}(p^*) \right|^{3/2} d\mathcal{H}^3(p^*) \\
&\geq \int_{B_3(0,1)} \left| \sum_{T \in \mathbb{T}} \chi_T \right|^{3/2}, \qquad (4\text{-}39)
\end{aligned}$$

where we now choose C large enough to ensure that (4-37) holds. In more detail, if C is large enough then for any SL_2 line L,

$$\int_L U_\sharp \mu \, d\mathcal{H}^1 \gtrsim \left| \{T \in \mathbb{T} : UB(w_T, 100\delta) \cap L \neq \varnothing\} \right|.$$

By the formula defining μ, to verify the above it suffices to check that for any $T \in \mathbb{T}$ such that $UB(w_T, 100\delta) \cap L \neq \varnothing$,

$$\mathcal{H}^1(L \cap UB(w_T, C\delta)) \gtrsim \delta,$$

which is clearly true if $C > 1000$. The inequality (4-38) used Lemma 3.5. The last line of (4-39) used

$$\ell(UB(w_T, 100\delta)) \cap B_3(0, 1) \supseteq \mathcal{N}_\delta(\ell(Uw_T)) \cap B_3(0, 1) \supseteq T \cap B_3(0, 1),$$

where the first part follows from the definition of ℓ in (3-8) (it is locally Lipschitz), and the second part follows from the assumption that $\ell(Uw_T)$ is the centre line of T.

If we apply Theorem 1.2 to the left-hand side of (4-39) with $\alpha = 2+\epsilon$, we get

$$\left\| \sum_{T\in\mathbb{T}} \chi_T \right\|^{3/2}_{L^{3/2}(B(0,1))} \lesssim \int \|\pi_{\theta\sharp}\mu\|^{3/2}_{3/2}\, d\theta \le C_\epsilon \delta^{-\epsilon/2} \mu(\mathbb{R}^3) c_2(\mu)^{1/2}. \tag{4-40}$$

The above inequality used that $c_{2+\epsilon}(\mu) \lesssim \delta^{-\epsilon} c_2(\mu)$ whenever μ is a linear combination of indicator functions over a boundedly overlapping collection of δ-balls. But $\mu(\mathbb{R}^3) \lesssim \delta^2|\mathbb{T}|$, and $c_2(\mu) \lesssim 1$ (this is where we use the 2-dimensional ball condition (1-1) on the family of tubes $\mathbb{T}$), so (4-40) implies (1-2) and proves Theorem 1.1. □

5. An L^p inequality for projections onto lines

Lemma 5.2 below is the analogue of Lemma 4.2 for lines instead of planes. To state it, as in the case of projections to planes, we first need a definition for a "small collection of bad intervals".

Definition 5.1. Given $R > 1$ and $\beta > 0$, and a Borel measure μ on $B_3(0,1)$ with $c_1(\mu) < \infty$, we call the triple $(\mu, R, \{\mathbb{D}_\theta\})$ "β-admissible" if $\{\mathbb{D}_\theta\}$ is a family of sets, where each $\mathbb{D}_\theta$ is a set of intervals of radius R^{-1} in $\mathrm{span}(\gamma(\theta))$, with $(\rho_{\theta\sharp}(\mu * \phi_R))(D) \ge c_1(\mu) R^{\beta-1}$ for each $D \in \mathbb{D}_\theta$, and such that the set $\left\{(\theta, x) \in [0, 2\pi) \times \mathbb{R}^3 : x \in \bigcup_{D\in\mathbb{D}_\theta} D\right\}$ is Borel measurable.

Lemma 5.2. *Let p be such that $1 \le p < \frac{3}{2}$. For any $\beta > 0$, there exists an $\kappa > 0$ (which may depend on β and p), such that*

$$\int_0^{2\pi} \int_{\bigcup_{D\in\mathbb{D}_\theta}} |\rho_{\theta\sharp}(\mu * \phi_R)|^p \, d\mathcal{H}^1 \, d\theta \le C_{\beta,p} R^{-\kappa} \mu(\mathbb{R}^3) c_1(\mu)^{p-1} \tag{5-1}$$

for any β-admissible triple $(\mu, R, \{\mathbb{D}_\theta\})$.

Proof. To prove (5-1), by pigeonholing we can replace μ by a sum of indicator functions over a disjoint family of R^{-1}-balls, and replace $\mu * \phi_R$ with μ in (5-1), i.e.,

$$\mu = \frac{R^3 \mu(\mathbb{R}^3)}{|\mathcal{B}|(4\pi/3)} \sum_{B\in\mathcal{B}} \chi_B. \tag{5-2}$$

By scaling we can additionally assume that $c_1(\mu) = 1$. Then (5-1) is equivalent to

$$\left(\frac{R^3\mu(\mathbb{R}^3)}{|\mathcal{B}|}\right)^p \int_0^{2\pi} \int_{\bigcup_{D\in\mathbb{D}_\theta}} \left| \rho_{\theta\sharp}\left(\sum_{B\in\mathcal{B}} \chi_B\right)\right|^p d\mathcal{H}^1 \, d\theta \le C_{\beta,p} R^{-\kappa} \mu(\mathbb{R}^3).$$

By integrating out the projections ρ_θ, it is enough to show that

$$\left(\frac{R\mu(\mathbb{R}^3)}{|\mathcal{B}|}\right)^p \int_0^{2\pi} \int_{\bigcup_{D\in\mathbb{D}_\theta}} \left| \sum_{B\in\mathcal{B}} \chi_{\rho_\theta(B)}\right|^p d\mathcal{H}^1 \, d\theta \le C_{\beta,p} R^{-\kappa} \mu(\mathbb{R}^3). \tag{5-3}$$

For each $z \in \mathbb{R}^3$, let $\Gamma_z := \{(\theta, \langle z, \gamma(\theta)\rangle) : \theta \in [0, 2\pi)\}$, which is a planar curve. Let $\delta = R^{-1}$, and let Γ_z^δ be the δ-neighbourhood of the planar curve Γ_z. Let Z'_δ be the set of centres of balls in $\mathcal{B}$. It is

straightforward to verify that

$$\sum_{B\in\mathcal{B}} \chi_{\rho_\theta(B)}(y) \le \sum_{z\in Z'_\delta} \chi_{\Gamma^\delta_z}(\theta, \langle y, \gamma(\theta)\rangle),$$

which was observed in [Pramanik et al. 2022, equation 2.12], and also in [Käenmäki et al. 2025]. To show (5-3) it is therefore enough to check that

$$\left(\frac{R\mu(\mathbb{R}^3)}{|\mathcal{B}|}\right)^p \int_0^{2\pi}\int_{\bigcup_{D\in\mathbb{D}_\theta}} \left|\sum_{z\in Z'_\delta} \chi_{\Gamma^\delta_z}(\theta, \langle y, \gamma(\theta)\rangle)\right|^p d\mathcal{H}^1(y)\, d\theta \le C_{\beta,p} R^{-\kappa}\mu(\mathbb{R}^3).$$

By Hölder's inequality applied to the above, it suffices to check that, for any $\epsilon > 0$ (where ϵ is much smaller than κ),

$$\int_0^{2\pi}\int \left|\sum_{z\in Z'_\delta} \chi_{\Gamma^\delta_z}(\theta, \langle y, \gamma(\theta)\rangle)\right|^{3/2} d\mathcal{H}^1(y)\, d\theta \le C_\epsilon R^\epsilon \mu(\mathbb{R}^3)^{-1/2} R^{-3/2}|\mathcal{B}|^{3/2}, \tag{5-4}$$

where we used the Vitali covering lemma to estimate the measure of

$$\left\{(\theta, x) \in [0, 2\pi)\times\mathbb{R}^3 : x \in \bigcup_{D\in\mathbb{D}_\theta} D\right\}.$$

By the Frostman condition on the measure μ, (5-4) follows from the general $L^{3/2}$ maximal function estimate in [Pramanik et al. 2022] (more precisely see Remark 5.10 on page 32 of [Pramanik et al. 2022]) which states (as a special case of [Pramanik et al. 2022, equation 5.63]) with $\alpha = \zeta = 1$)

$$\int_0^{2\pi}\int \left|\sum_{z\in Z'_\delta} \chi_{\Gamma^\delta_z}(\theta, \langle y, \gamma(\theta)\rangle)\right|^{3/2} d\mathcal{H}^1(y)\, d\theta \le C_\epsilon \delta^{-\epsilon} A^{1/2}\delta|Z'_\delta|, \tag{5-5}$$

whenever Z'_δ is a collection of points in $B_3(0, 1)$ such that

$$|\{z \in Z'_\delta \cap B(z_0, r)\}| \le A(r/\delta), \quad r \in [\delta, 1]. \tag{5-6}$$

The formula (5-2) and the assumption that $c_1(\mu) \le 1$ yields that (5-6) holds with $A \sim \delta|Z'_\delta|\mu(\mathbb{R}^3)^{-1}$, so (5-5) implies (5-4), and this finishes the proof. □

Proof of Theorem 1.3. The theorem for $p = 1$ follows from the theorem for any $p > 1$, so it may be assumed that $p > 1$. Therefore, by a limiting argument, it may be assumed that μ is a smooth function. Take $\beta > 0$ such that $\beta \ll \alpha - 1$, and let $\delta, \epsilon > 0$ be such that $\delta \ll \epsilon \ll \beta$, and $\delta \ll \kappa(\beta)$, the parameter in Lemma 5.2. Write $\mu = \mu_b + \mu_g$, where

$$\mu_b := \sum_{j\ge1}\sum_{\tau\in\Lambda_j}\sum_{T\in\mathbb{T}_{\tau,b}} M_T\mu, \quad \mu_g := \mu - \mu_b, \quad M_T\mu = \eta_T(\mu * \check{\psi}_\tau),$$

and where $\{\psi_\tau\}_{\tau\in\Lambda}$ is a smooth partition of unity on the ~ 1 neighbourhood of the light cone minus the set $B_3(0, 1)$, $\Lambda = \bigcup_{j\ge1}\Lambda_j$, and each Λ_j is a boundedly overlapping cover of the ~ 1 neighbourhood of the light cone intersected with $B(0, 2^j)\setminus B(0, 2^{j-1})$, by boxes of dimensions $\sim 1\times 2^{j/2}\times 2^j$. For each τ,

the set $\mathbb{T}_\tau$ is a boundedly overlapping cover of $\mathbb{R}^3$ by planks T of dimensions $2^{j\delta} \times 2^{-j(1/2-\delta)} \times 2^{-j(1-\delta)}$ dual to τ, and

$$\mathbb{T}_{\tau,b} = \big\{T \in \mathbb{T}_\tau : \mu(4T) \geq 2^{-j(1-\beta)} c_\alpha(\mu)\big\}, \quad \mathbb{T}_{\tau,g} = \mathbb{T}_\tau \setminus \mathbb{T}_{\tau,b}. \tag{5-7}$$

The functions η_T form a smooth partition of unity subordinate to the cover $\mathbb{T}_\tau$ of $\mathbb{R}^3$, and each $\widehat{\eta}_T$ is rapidly decaying outside the box T_0 around the origin dual to T.

Fix $j \geq 1$. It will be shown that

$$\int_0^{2\pi} \int \Big| \sum_{\tau \in \Lambda_j} \sum_{T \in \mathbb{T}_{\tau,b}} \rho_{\theta\sharp} M_T \mu \Big|^p \, d\mathcal{H}^1 \, d\theta \lesssim 2^{jO(\delta)} \int_0^{2\pi} \sum_{\substack{\tau \in \Lambda_j : \\ |\theta_\tau - \theta| \leq 2^{j(\delta-1/2)}}} \sum_{T \in \mathbb{T}_{\tau,b}} \mu(2T)^p 2^{j(p-1)} \, d\theta, \tag{5-8}$$

where θ_τ is such that $\gamma(\theta_\tau)$ is parallel to the long axis of τ (the medium axis of τ is parallel to $\gamma'(\theta_\tau)$ and the short axis of τ is parallel to $(\gamma \times \gamma')(\theta_\tau)$). If p were equal to 2 then (5-8) would be a consequence of the following argument. For each θ, apply Plancherel's theorem in one dimension to change the left-hand side of (5-8) to a surface integral over the light cone at distance $\sim 2^j$ from the origin. The uncertainty principle, using that μ is supported in the unit ball, changes this integral to (2^{-j} times) an integral over $\mathbb{R}^3$ (essentially supported in the ~ 1 neighbourhood of the light cone in $\mathbb{R}^3$ at distance $\sim 2^j$ from the origin). Plancherel's theorem in $\mathbb{R}^3$ can then be used to decouple the L^2 norms (this is similar to the argument from equations 2.16 to 2.18 in [Harris 2024]), and then (5-8) follows from the uncertainty principle or direct calculation (and (5-8) continues to hold with the set of planks in the left-hand side of (5-8) replaced by any subcollection). If p were equal to 1 then (5-8) would follow from the triangle inequality and stationary phase (see, e.g., Lemma 2.2 of [Harris 2024]). This implies (5-8) if either $p = 1$ or $p = 2$ (and similarly for any subcollection of T's in the left-hand side). The case $p \in (1, 2)$ of (5-8) then holds by interpolation; by pigeonholing it may be assumed that the sum over the T's in the left-hand side of (5-8) is over a subset of the T's such that $\mu(2T)$ is constant up to a factor of 2, and then (5-8) follows by writing $p = (1-\phi) \cdot 1 + \phi \cdot 2$ and applying Hölder's inequality with $q = 1/(1-\phi)$ and $q' = 1/\phi$. This proves (5-8).

By Hölder's inequality, the integrand in the right-hand side of (5-8) satisfies, for any $\theta \in [0, 2\pi)$,

$$\sum_{\substack{\tau \in \Lambda_j : \\ |\theta_\tau - \theta| \leq 2^{j(\delta-1/2)}}} \sum_{T \in \mathbb{T}_{\tau,b}} \mu(2T)^p 2^{j(p-1)} \lesssim 2^{jO(\delta)} \int_{F_j(\theta)} |\rho_{\theta\sharp}(\mu * \phi_{2^{j(1-\delta)}})|^p \, d\mathcal{H}^1, \tag{5-9}$$

where we used the inequality $\|x\|_p \leq \|x\|_1$ for sequences to deal with the possibility of multiple planks in a slab,[4] and

$$F_j(\theta) := \bigcup_{\substack{\tau \in \Lambda_j : \\ |\theta_\tau - \theta| \leq 2^{j(\delta-1/2)}}} \bigcup_{T \in \mathbb{T}_{\tau,b}} \rho_\theta(100T).$$

[4]This step was not needed in the case of projections onto planes, and was overlooked in [Harris 2023a].

By the triangle inequality, followed by (5-8), and then (5-9),

$$\left(\int_0^{2\pi}\int |\rho_{\theta\sharp}\mu_b|^p\,d\mathcal{H}^1\,d\theta\right)^{1/p} \lesssim \sum_{j\geq 1}\left(2^{jO(\delta)}\int_0^{2\pi}\int_{F_j(\theta)} |\rho_{\theta\sharp}(\mu * \phi_{2^{j(1-\delta)}})|^p\,d\mathcal{H}^1\,d\theta\right)^{1/p}.$$

By the definition of the sets $\mathbb{T}_{\tau,b}$, the collection of intervals formed by $F_j(\theta)$ generate an admissible triple satisfying the assumptions of Lemma 5.2, so applying this yields

$$\int_0^{2\pi}\int |\rho_{\theta\sharp}\mu_b|^p\,d\mathcal{H}^1\,d\theta \lesssim c_\alpha(\mu)^{p-1}\mu(\mathbb{R}^3)\left(\sum_{j\geq 1} 2^{-j\kappa/10^{10}}\right)^p \lesssim c_\alpha(\mu)^{p-1}\mu(\mathbb{R}^3). \tag{5-10}$$

This bounds the "bad" part.

It remains to bound the "good" part. From this point on, the L^2 method we use will be similar to [Harris 2024], but we include the details for completeness. It will be shown that

$$\int_0^{2\pi}\int |\rho_{\theta\sharp}\mu_g|^2\,d\mathcal{H}^1\,d\theta \lesssim \mu(\mathbb{R}^3)c_\alpha(\mu)\sum_{j\geq 1} 2^{j\left((5-2\alpha)/p_{\mathrm{dec}}-1/2+\beta(p_{\mathrm{dec}}-2/p_{\mathrm{dec}})\right)+jO(\delta)}, \tag{5-11}$$

where $2\leq p_{\mathrm{dec}}\leq 6$. Before proving (5-11), we will show that it implies the theorem. By taking $p_{\mathrm{dec}}=6$ and using $\delta\ll\beta\ll\alpha-1$, (5-11) is a decaying geometric series, which yields

$$\int_0^{2\pi}\int |\rho_{\theta\sharp}\mu_g|^2\,d\mathcal{H}^1\,d\theta \lesssim \mu(\mathbb{R}^3)c_\alpha(\mu).$$

By taking $p=1$, the previous argument for the bad part to obtain (5-10) shows that $\rho_{\theta\sharp}\mu_b\in L^1$ for a.e. $\theta\in[0,2\pi)$, and that

$$\int_0^{2\pi}\int |\rho_{\theta\sharp}\mu_b|\,d\mathcal{H}^1\,d\theta \lesssim \mu(\mathbb{R}^3).$$

By the triangle inequality, this implies that

$$\int_0^{2\pi}\int |\rho_{\theta\sharp}\mu_g|\,d\mathcal{H}^1\,d\theta \lesssim \mu(\mathbb{R}^3).$$

For general p with $1\leq p<\frac{3}{2}$, by writing $p=(1-\phi)1+\phi\cdot 2$, where $\phi=p-1$, and by applying Hölder's inequality to the function

$$|\rho_{\theta\sharp}\mu_g|^p = |\rho_{\theta\sharp}\mu_g|^{(1-\phi)\cdot 1}\cdot|\rho_{\theta\sharp}\mu_g|^{\phi\cdot 2},$$

with $q=\frac{1}{1-\phi}$ and $q'=\frac{1}{\phi}$, this yields

$$\int_0^{2\pi}\int |\rho_{\theta\sharp}\mu_g|^p\,d\mathcal{H}^1\,d\theta \lesssim c_\alpha(\mu)^{p-1}\mu(\mathbb{R}^3). \tag{5-12}$$

Combining (5-10) and (5-12) gives

$$\int_0^{2\pi}\int |\rho_{\theta\sharp}\mu|^p\,d\mathcal{H}^1\,d\theta \lesssim c_\alpha(\mu)^{p-1}\mu(\mathbb{R}^3).$$

This proves the theorem, assuming (5-11).

It remains to prove (5-11). By the triangle inequality and by symmetry, to prove (5-11) it suffices to show that for any $j \geq 1$,

$$\int_0^{2\pi} \int_{2^{j-1}}^{2^j} |\widehat{\mu}_g(t\gamma(\theta))|^2 \, dt \, d\theta \lesssim \mu(\mathbb{R}^3) c_\alpha(\mu) 2^{j\left((5-2\alpha)/p_{\mathrm{dec}} - 1/2 + \beta(p_{\mathrm{dec}} - 2/p_{\mathrm{dec}})\right) + jO(\epsilon)}, \tag{5-13}$$

for any $p_{\mathrm{dec}} \in (2, 6]$. By the definition of μ_g, it suffices to show that for any $j \geq 1$,

$$\int_0^{2\pi} \int_{2^{j-10}}^{2^{j+10}} \left| \sum_{\tau \in \Lambda_j} \sum_{T \in \mathbb{T}_{\tau,g}} \widehat{M_T \mu}(t\gamma(\theta)) \right|^2 dt \, d\theta \lesssim \mu(\mathbb{R}^3) c_\alpha(\mu) 2^{j\left((5-2\alpha)/p_{\mathrm{dec}} - 1/2 + \beta(p_{\mathrm{dec}} - 2/p_{\mathrm{dec}})\right) + jO(\epsilon)},$$

where, above and below, we can assume that rapidly decaying tail terms do not dominate (since in this case (5-13) would be trivial). Since μ is supported in the unit ball, the uncertainty principle can be used to change the above surface integral to an integral over $\mathbb{R}^3$. In more detail, we partition the part of the light cone over which we are integrating into balls of radius $2^{-10j\delta}$, and then use the locally constant property of the integrand over each such ball to replace the surface integral $dt \, d\theta$ over the ball intersected with the light cone with the 3-dimensional integral dx over the ball. Because the $dt \, d\theta$ measure above of such a ball intersected with the light cone is $\lesssim 2^{-j}$ (as $t \sim 2^j$ and is varying on an interval of length ~ 1, the angle θ such that $t\gamma(\theta)$ intersects the ball must lie in an interval of size $\sim 2^{-j}$), we gain a factor of 2^{-j} when changing $dt \, d\theta$ to dx. We then replace the integral over the $2^{-10j\delta}$-neighbourhood of the light cone with an integral over all of $\mathbb{R}^3$, and then use Plancherel's theorem in $\mathbb{R}^3$ to decouple the L^2 norms. Due to the 2^{-j} gain, it is enough to show that

$$\sum_{\tau \in \Lambda_j} \sum_{T \in \mathbb{T}_{\tau,g}} \int_{\mathbb{R}^3} |M_T \mu|^2 \, dx \lesssim \mu(\mathbb{R}^3) c_\alpha(\mu) 2^{j\left((5-2\alpha)/p_{\mathrm{dec}} + 1/2 + \beta(p_{\mathrm{dec}} - 2/p_{\mathrm{dec}})\right) + jO(\epsilon)}. \tag{5-14}$$

By writing out the L^2 norm above as an inner product, applying the definition $M_T \mu = \eta_T (\mu * \check{\psi}_\tau)$ to one of the $M_T \mu$'s, and using Fubini's theorem to interchange the μ integral with all the others, the left-hand side of (5-14) is equal to

$$\int \sum_{\tau \in \Lambda_j} \sum_{T \in \mathbb{T}_{\tau,g}} (\eta_T M_T \mu) * \check{\psi}_\tau \, d\mu. \tag{5-15}$$

By the Cauchy–Schwarz inequality (with respect to the measure μ),

$$(5\text{-}15) \lesssim \mu(\mathbb{R}^3)^{1/2} \left(\int \left| \sum_{\tau \in \Lambda_j} \sum_{T \in \mathbb{T}_{\tau,g}} (\eta_T M_T \mu) * \check{\psi}_\tau \right|^2 d\mu \right)^{1/2}. \tag{5-16}$$

For each T above, abbreviate $f_T = (\eta_T M_T \mu) * \check{\psi}_\tau$. By pigeonholing, the sum over T in the left-hand side above can be restricted to a subset $\mathbb{W}$ of $\bigcup_{\tau \in \Lambda_j} \mathbb{T}_{\tau,g}$ on which $\|f_T\|_{p_{\mathrm{dec}}}$ is constant over $T \in \mathbb{W}$ up to a factor of 2. Because the integrand on the left-hand side of (5-16) has frequencies of size $\lesssim 2^j$, the measure μ being integrated against can be replaced by its smoothed out version $\mu * \phi_{2^j}$ at scale 2^{-j}. We then partition the domain into balls B of radius 2^{-j}, and pigeonhole these balls B according to the

number M of planks $T \in \mathbb{W}$ such that $2T$ has nonempty intersection with B. Let Y be the union over these balls B. With this setup, (5-16) will be bounded by

$$\log(2^j)^{O(1)}\mu(\mathbb{R}^3)^{1/2}\bigg(\int_Y \bigg|\sum_{T\in\mathbb{W}} f_T\bigg|^2 d(\mu * \phi_j)\bigg)^{1/2}. \tag{5-17}$$

By Hölder's inequality (with respect to Lebesgue measure),

$$\bigg(\int_Y \bigg|\sum_{T\in\mathbb{W}} f_T\bigg|^2 d(\mu * \phi_j)\bigg)^{1/2} \le \bigg(\int_Y \bigg|\sum_{T\in\mathbb{W}} f_T\bigg|^{p_{\mathrm{dec}}}\bigg)^{1/p_{\mathrm{dec}}} \bigg(\int_Y (\mu * \phi_j)^{p_{\mathrm{dec}}/p_{\mathrm{dec}}-2}\bigg)^{1/2-1/p_{\mathrm{dec}}}. \tag{5-18}$$

By the refined decoupling theorem (see Theorem A.1), the first factor in (5-18) satisfies

$$\bigg(\int_Y \bigg|\sum_{T\in\mathbb{W}} f_T\bigg|^{p_{\mathrm{dec}}}\bigg)^{1/p_{\mathrm{dec}}} \le C_\epsilon 2^{j\epsilon}\bigg(\frac{M}{|\mathbb{W}|}\bigg)^{1/2-1/p_{\mathrm{dec}}} \bigg(\sum_{T\in\mathbb{W}} \|f_T\|_{p_{\mathrm{dec}}}^2\bigg)^{1/2}. \tag{5-19}$$

By using (5-7), the second factor in (5-18) satisfies

$$\begin{aligned}\bigg(\int_Y (\mu * \phi_j)^{p_{\mathrm{dec}}/(p_{\mathrm{dec}}-2)}\bigg)^{1/2-1/p_{\mathrm{dec}}} &\le \bigg(\frac{1}{M}\sum_{T\in\mathbb{W}} \int_{Y\cap 2T} (\mu * \phi_j)^{p_{\mathrm{dec}}/(p_{\mathrm{dec}}-2)}\bigg)^{1/2-1/p_{\mathrm{dec}}} \\ &\lesssim \bigg(\frac{1}{M}\sum_{T\in\mathbb{W}} \mu(4T)\bigg)^{1/2-1/p_{\mathrm{dec}}} c_\alpha(\mu)^{1/p_{\mathrm{dec}}} 2^{j(3-\alpha)/p_{\mathrm{dec}}} \\ &\lesssim c_\alpha(\mu)^{1/2}\bigg(\frac{|\mathbb{W}|}{M}\bigg)^{1/2-1/p_{\mathrm{dec}}} 2^{-j(1-\beta)(1/2-1/p_{\mathrm{dec}})+j(3-\alpha)/p_{\mathrm{dec}}}.\end{aligned} \tag{5-20}$$

Substituting (5-19) and (5-20) into (5-17) and then (5-18) yields

$$\begin{aligned}&\bigg(\int \bigg|\sum_{\tau\in\Lambda_j}\sum_{T\in\mathbb{T}_{\tau,g}} (\eta_T M_T\mu) * \check{\psi}_\tau\bigg|^2 d\mu\bigg)^{1/2} \\ &\qquad\lesssim c_\alpha(\mu)^{1/2} 2^{-j(1-\beta)(1/2-1/p_{\mathrm{dec}})+j(3-\alpha)/p_{\mathrm{dec}}+jO(\epsilon)}\bigg(\sum_{\tau\in\Lambda_j}\sum_{T\in\mathbb{T}_{\tau,b}} \|f_T\|_{p_{\mathrm{dec}}}^2\bigg)^{1/2}.\end{aligned} \tag{5-21}$$

By comparing the left-hand side of the above with (5-15) and then (5-14), and then by using

$$\|f_T\|_{p_{\mathrm{dec}}} \lesssim 2^{jO(\delta)}\|M_T\mu\|_2 2^{3j/2(1/2-1/p_{\mathrm{dec}})}$$

(which follows from the definition of f_T and the Hausdorff–Young inequality, followed by Hölder's inequality), (5-21) implies that

$$\begin{aligned}&\sum_{\tau\in\Lambda_j}\sum_{T\in\mathbb{T}_{\tau,g}} \|M_T\mu\|_2^2 \\ &\qquad\lesssim \bigg(\sum_{\tau\in\Lambda_j}\sum_{T\in\mathbb{T}_{\tau,g}} \|M_T\mu\|_2^2\bigg)^{1/2} \mu(\mathbb{R}^3)^{1/2} c_\alpha(\mu)^{1/2} 2^{(3j/2-j(1-\beta))(1/2-1/p_{\mathrm{dec}})+j(3-\alpha)/p_{\mathrm{dec}}+jO(\epsilon)}.\end{aligned}$$

This simplifies to

$$\sum_{\tau\in\Lambda_j}\sum_{T\in\mathbb{T}_{\tau,g}}\|M_T\mu\|_2^2\lesssim\mu(\mathbb{R}^3)c_\alpha(\mu)2^{j(1+2\beta)(1/2-1/p_{\mathrm{dec}})+2j(3-\alpha)/p_{\mathrm{dec}}+jO(\epsilon)}.$$

This is the same as (5-14), so this finishes the proof. □

6. Some open problems

(1) An open problem that lies in-between the stated intersection and projection theorems, is whether there is a "restricted" Besicovitch–Federer projection theorem in $\mathbb{R}^3$. For example, if $A\subseteq\mathbb{R}^3$ is a Borel set with $0<\mathcal{H}^2(A)<\infty$, such that the projection $\pi_{v^\perp}(A)$ of A onto $v^\perp$ has $\mathcal{H}^2(\pi_{v^\perp}(A))=0$ for a.e. $v\in S^2$, where $\pi_{v^\perp}$ is orthogonal projection onto $v^\perp$, then is $\mathcal{H}^2(\pi_\theta(A))=0$ for a.e. $\theta\in[0,2\pi)$?

(2) We do not know if $p<\frac{3}{2}$ in Theorem 1.3 can be replaced by $p=\frac{3}{2}$. The exponent $p=\frac{3}{2}$ is only sharp in Theorems 1.2 and 1.3 if p is not allowed to depend on α. In Theorem 1.2, if $p=p(\alpha)$ is allowed to depend on α, then up to the endpoint the sharp exponent is likely the one which naturally interpolates between $p=\frac{3}{2}$ at $\alpha=2$ and $p=2$ at $\alpha=\frac{5}{2}$, and similarly for Theorem 1.3.

(3) Another open problem is whether $B_3(0,1)$ in Theorem 1.1 can be replaced by $\mathbb{R}^3$; unlike for the standard Kakeya maximal function, the local and global versions do not seem to be obviously equivalent, since horizontal or SL_2 lines are not preserved by Euclidean translations. For the global problem on $\mathbb{R}^3$, the ball condition (1-1) is no longer equivalent to the Wolff axioms in the SL_2 case. Moreover, the example consisting of $\sim\delta^{-2}$ many $1\times\delta$ tube segments inside a vertical $1\times\delta\times1$ slab parallel to the xz-plane, centred at the point $(\delta^{-1},0,0)$, shows that the global version of (1-2) is *false* if only the ball condition (1-1) is assumed. However, the global version of (1-2) with $B_3(0,1)$ replaced by $\mathbb{R}^3$ may still be true if the full Wolff axioms are assumed instead of just the ball condition (1-1).

Appendix: Refined decoupling

Given $R\geq1$ and a small constant $c>0$, let Θ_R be a $\sim cR^{-1/2}$-separated subset of $[0,2\pi)$. For each $\theta\in\Theta_R$, let $\tau(\theta)$ be a plank of dimensions $1\times R^{-1/2}\times R^{-1}$ of distance ~1 from the origin, tangent to the cone at $\gamma(\theta)$. Given $\delta>0$, for each τ, let $\mathbb{T}_\tau$ be a collection of planks of dimensions $R^\delta\times R^{1/2+\delta}\times R^{1+\delta}$ dual to τ, forming a boundedly overlapping cover of $\mathbb{R}^3$.

Theorem A.1 [Gan et al. 2022, Theorem 9]. *Let $p_{\mathrm{dec}}\in[2,6]$. If $c>0$ is sufficiently small, then for any $\epsilon>0$, there exists $\delta_0>0$ such that the following holds for all $0<\delta<\delta_0$. Let $R>1$, and suppose that $f=\sum_{T\in\mathbb{W}}f_T$, where $\mathbb{W}\subseteq\bigcup_{\theta\in\Theta_R}\mathbb{T}_{\tau(\theta)}$, and each f_T has $\widehat{f_T}$ supported in τ, with f_T essentially supported on T in the sense that $\|f_T\|_{L^\infty(\mathbb{R}^3\setminus T)}\leq AR^{-10000}$, for a large constant A. Suppose that $\|f_T\|_p$ is constant over $T\in\mathbb{W}$ up to a factor of 2, and that Y is a disjoint union of radius 1 balls in $B(0,R)$, such that each radius 1 ball $B\subseteq Y$ intersects at most M planks $2T$ with $T\in\mathbb{W}$. Then*

$$\Big\|\sum_{T\in\mathbb{W}}f_T\Big\|_{L^{p_{\mathrm{dec}}}(Y)}\leq C_{\epsilon,\delta,A,c}R^\epsilon\Big(\frac{M}{|\mathbb{W}|}\Big)^{1/2-1/p_{\mathrm{dec}}}\Big(\sum_{T\in\mathbb{W}}\|f_T\|_{p_{\mathrm{dec}}}^2\Big)^{1/2}.$$

Acknowledgements

The "microlocal interpolation" argument from [Demeter 2020, Exercise 9.21] used in Section 4 was explained to Harris in a different context by Alex Barron. A possible connection between the L^p SL_2 Kakeya maximal inequality and L^p bounds for restricted families of projections was suggested by Josh Zahl.

We thank the reviewer for carefully reading the manuscript and providing many helpful comments and suggestions.

Harris was partially supported by funding from the American Institute of Mathematics (AIM). Ou was partially supported by NSF DMS-2142221 and NSF DMS-2055008.

References

[Davies 1952] R. O. Davies, "Subsets of finite measure in analytic sets", *Indag. Math.* **14** (1952), 488–489. MR

[Demeter 2020] C. Demeter, *Fourier restriction, decoupling, and applications*, Cambridge Studies in Advanced Mathematics **184**, Cambridge Univ. Press, 2020. MR

[Dąbrowski et al. 2022] D. Dąbrowski, T. Orponen, and M. Villa, "Integrability of orthogonal projections, and applications to Furstenberg sets", *Adv. Math.* **407** (2022), art. id. 108567. MR

[Erdoğan 2004] M. B. Erdoğan, "A note on the Fourier transform of fractal measures", *Math. Res. Lett.* **11**:2-3 (2004), 299–313. MR

[Fässler and Orponen 2014] K. Fässler and T. Orponen, "On restricted families of projections in $\mathbb{R}^3$", *Proc. Lond. Math. Soc.* (3) **109**:2 (2014), 353–381. MR

[Fässler and Orponen 2023a] K. Fässler and T. Orponen, "A note on Kakeya sets of horizontal and SL(2) lines", *Bull. Lond. Math. Soc.* **55**:5 (2023), 2195–2204. MR

[Fässler and Orponen 2023b] K. Fässler and T. Orponen, "Vertical projections in the Heisenberg group via cinematic functions and point-plate incidences", *Adv. Math.* **431** (2023), art. id. 109248. MR

[Fässler et al. 2025] K. Fässler, A. Pinamonti, and P. Wald, "A Kakeya maximal inequality in the Heisenberg group", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **26**:3 (2025), 1451–1474. MR

[Gan et al. 2022] S. Gan, S. Guo, L. Guth, T. L. J. Harris, D. Maldague, and H. Wang, "On restricted projections to planes in $\mathbb{R}^3$", preprint, 2022. arXiv 2207.13844v1

[Gan et al. 2024] S. Gan, L. Guth, and D. Maldague, "An exceptional set estimate for restricted projections to lines in $\mathbb{R}^3$", *J. Geom. Anal.* **34**:1 (2024), art. id. 15. MR

[Guth et al. 2020] L. Guth, A. Iosevich, Y. Ou, and H. Wang, "On Falconer's distance set problem in the plane", *Invent. Math.* **219**:3 (2020), 779–830. MR

[Harris 2022] T. L. J. Harris, "Improved bounds for restricted projection families via weighted Fourier restriction", *Anal. PDE* **15**:7 (2022), 1655–1701. MR

[Harris 2023a] T. L. J. Harris, "An $L^{4/3}$ SL_2 Kakeya maximal inequality", preprint, 2023. arXiv 2311.14667v2

[Harris 2023b] T. L. J. Harris, "Restricted families of projections onto planes: the general case of nonvanishing geodesic curvature", *Rev. Mat. Iberoam.* **39**:5 (2023), 1863–1894. MR

[Harris 2024] T. L. J. Harris, "Length of sets under restricted families of projections onto lines", pp. 1–17 in *Recent developments in harmonic analysis and its applications*, edited by S. Guo et al., Contemp. Math. **792**, Amer. Math. Soc., Providence, RI, 2024. MR

[Järvenpää et al. 2008] E. Järvenpää, M. Järvenpää, F. Ledrappier, and M. Leikas, "One-dimensional families of projections", *Nonlinearity* **21**:3 (2008), 453–463. MR

[Käenmäki et al. 2025] A. Käenmäki, T. Orponen, and L. Venieri, "A Marstrand-type restricted projection theorem in $\mathbb{R}^3$", *Amer. J. Math.* **147**:1 (2025), 81–123. MR

[Kahane 1969] J.-P. Kahane, "Trois notes sur les ensembles parfaits linéaires", *Enseign. Math.* (2) **15** (1969), 185–192. MR

[Katz and Tao 2001] N. H. Katz and T. Tao, "Some connections between Falconer's distance set conjecture and sets of Furstenburg type", *New York J. Math.* **7** (2001), 149–187. MR

[Katz and Zahl 2019] N. H. Katz and J. Zahl, "An improved bound on the Hausdorff dimension of Besicovitch sets in $\mathbb{R}^3$", *J. Amer. Math. Soc.* **32**:1 (2019), 195–259. MR

[Katz et al. 2023] N. H. Katz, S. Wu, and J. Zahl, "Kakeya sets from lines in SL_2", *Ars Inven. Anal.* (2023), art. id. 6. MR

[Li and Liu 2026] L. Li and B. Liu, "Orthogonal projections, a dual Furstenberg set problem, and discretized sum-product estimates", *J. Fourier Anal. Appl.* **32**:4 (2026), art. id. 68. MR

[Liu 2022] J. Liu, "On the dimension of Kakeya sets in the first Heisenberg group", *Proc. Amer. Math. Soc.* **150**:8 (2022), 3445–3455. MR

[Mattila 1995] P. Mattila, *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability*, Cambridge Studies in Advanced Mathematics **44**, Cambridge Univ. Press, 1995. MR

[Mattila 2015] P. Mattila, *Fourier analysis and Hausdorff dimension*, Cambridge Studies in Advanced Mathematics **150**, Cambridge Univ. Press, 2015. MR

[Mattila 2019] P. Mattila, "Hausdorff dimension, projections, intersections, and Besicovitch sets", pp. 129–157 in *New trends in applied harmonic analysis*, 2: *Harmonic analysis*, *geometric measure theory*, *and applications*, edited by A. Aldroubi et al., Springer, 2019. MR

[Mattila 2023a] P. Mattila, "Hausdorff dimension of plane sections and general intersections", preprint, 2023. arXiv 2310.07538v1

[Mattila 2023b] P. Mattila, "A survey on the Hausdorff dimension of intersections", preprint, 2023. arXiv 2301.13478v3

[Oberlin and Oberlin 2015] D. Oberlin and R. Oberlin, "Application of a Fourier restriction theorem to certain families of projections in $\mathbb{R}^3$", *J. Geom. Anal.* **25**:3 (2015), 1476–1491. MR

[Orponen 2017] T. Orponen, "Quasisymmetric maps on Kakeya sets", *Int. Math. Res. Not.* **2017**:11 (2017), 3413–3425. MR

[Orponen 2019] T. Orponen, "On the dimension and smoothness of radial projections", *Anal. PDE* **12**:5 (2019), 1273–1294. MR

[Orponen and Venieri 2020] T. Orponen and L. Venieri, "Improved bounds for restricted families of projections to planes in $\mathbb{R}^3$", *Int. Math. Res. Not.* **2020**:19 (2020), 5797–5813. MR

[Pramanik et al. 2022] M. Pramanik, T. Yang, and J. Zahl, "A Furstenberg-type problem for circles, and a Kaufman-type restricted projection theorem in $\mathbb{R}^3$", preprint, 2022. arXiv 2207.02259v2

[Stein and Shakarchi 2005] E. M. Stein and R. Shakarchi, *Real analysis*: *measure theory*, *integration*, *and Hilbert spaces*, Princeton Lectures in Analysis **3**, Princeton Univ. Press, 2005. MR

[Wang and Zahl 2026] H. Wang and J. Zahl, "Sticky Kakeya sets and the sticky Kakeya conjecture", *J. Amer. Math. Soc.* **39**:2 (2026), 515–585. MR

[Wolff 1999] T. Wolff, "Decay of circular means of Fourier transforms of measures", *Int. Math. Res. Not.* **1999**:10 (1999), 547–567. MR

[Wolff 2003] T. H. Wolff, *Lectures on harmonic analysis*, edited by I. Łaba and C. Shubin, University Lecture Series **29**, Amer. Math. Soc., Providence, RI, 2003. MR

Received 21 Jun 2024. Revised 6 May 2025. Accepted 19 Sep 2025.

JOHN GREEN: john.d.green@warwick.ac.uk
Mathematics Institute, University of Warwick, Coventry, United Kingdom

TERENCE L. J. HARRIS: terry.harris@uq.edu.au
Department of Mathematics, University of Wisconsin-Madison, Madison, WI, United States

YUMENG OU: yumengou@sas.upenn.edu
Department of Mathematics, University of Pennsylvania, Philadelphia, PA, United States

ANALYSIS AND PDE
Vol. 19 (2026), No. 7, pp. 1339–1360
DOI: 10.2140/apde.2026.19.1339

THE BRUNN–MINKOWSKI INEQUALITY FOR THE FIRST EIGENVALUE OF THE ORNSTEIN–UHLENBECK OPERATOR AND LOG-CONCAVITY OF THE RELEVANT EIGENFUNCTION

ANDREA COLESANTI, ELISA FRANCINI, GALYNA LIVSHYTS AND PAOLO SALANI

We prove that the first (nontrivial) Dirichlet eigenvalue of the Ornstein–Uhlenbeck operator

$$L(u) = \Delta u - \langle \nabla u, x \rangle,$$

as a function of the domain, is convex with respect to Minkowski addition, and we characterize the equality cases in some classes of convex sets. We also prove that the corresponding (positive) eigenfunction is log-concave if the domain is convex.

1. Introduction

We are concerned with the first Dirichlet eigenvalue λ_γ of the Ornstein–Uhlenbeck operator and the corresponding eigenfunction. Such an eigenvalue represents the counterpart, in Gauss space, of the classical principal frequency, which can also be characterized as the optimal constant in the Poincaré inequality, or as the first Dirichlet eigenvalue of the Laplace operator.

We investigate in particular two concepts: a Brunn–Minkowski-type inequality for λ_γ with respect to the Minkowski addition, and the log-concavity of the corresponding (positive) eigenfunction. As we will see, the situation in the Gauss space closely parallels that of the classical setting, with some differences due to the lack of homogeneity and translation invariance of the Gauss measure.

1.1. ***Short history in the classical Lebesgue case.*** For a bounded set Ω (with sufficiently regular boundary) in $\mathbb{R}^n$, the principal frequency of Ω is defined as

$$\lambda(\Omega) := \inf\left\{ \frac{\int_\Omega |\nabla v|^2 \, dx}{\int_\Omega v^2 \, dx} : v \in W_0^{1,2}(\Omega),\ v \not\equiv 0 \right\}$$

(here and below we adopt the standard notation for Sobolev spaces). It is well known, and follows via a variational argument, that $\lambda(\Omega)$ is the smallest positive number λ such that there exists a solution of the boundary value problem

$$\begin{cases} \Delta u = -\lambda u,\ u > 0, & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{1}$$

See, e.g., [Gilbarg and Trudinger 1983; Pólya and Szegö 1951] for a classical exposition of this topic.

MSC2020: primary 35E10; secondary 35P15, 52A40.
Keywords: Ornstein–Uhlenbeck operator, principal frequency, Gauss space, Brunn–Minkowski inequality, convex body.

© 2026 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.

The principal frequency $\lambda(\Omega)$ is monotone decreasing with respect to inclusion; that is, $\lambda(\Omega_1) \geq \lambda(\Omega_2)$ whenever $\Omega_1 \subset \Omega_2$. Moreover, λ is a positively homogeneous functional of order -2; that is,

$$\lambda(t\Omega) = t^{-2}\lambda(\Omega)$$

for any $t > 0$.

The principal frequency satisfies various inequalities of isoperimetric type. The most significant is the Faber–Krahn inequality (see, e.g., [Pólya and Szegö 1951]) which states that

$$\lambda(\Omega) \geq \lambda(\Omega'),$$

where Ω' is a ball having the same Lebesgue measure as Ω (an interesting strengthening of the Faber–Krahn inequality was proven by Kohler-Jobin [1978; 1982], see also Brasco [2014]).

A different inequality, proven by Brascamp and Lieb [1976], states that $\lambda^{-1/2}$ is concave with respect to the Minkowski addition of sets. We recall that for $A, B \subset \mathbb{R}^n$ their Minkowski addition is defined as

$$A + B = \{x + y : x \in A,\ y \in B\}.$$

Theorem 1.1 (Brascamp–Lieb). *Let Ω_0, Ω_1 be connected, open and bounded subsets of $\mathbb{R}^n$, and let $t \in [0, 1]$. Then*

$$\lambda((1-t)\Omega_0 + t\Omega_1)^{-1/2} \geq (1-t)\lambda(\Omega_0)^{-1/2} + t\lambda(\Omega_1)^{-1/2}. \tag{2}$$

We refer the reader to [Borell 2000] for an alternative proof and to [Colesanti 2005] for the characterization of equality conditions. In fact, in [Brascamp and Lieb 1976] the following weaker inequality is proved:

$$\lambda((1-t)\Omega_0 + t\Omega_1) \leq (1-t)\lambda(\Omega_0) + t\lambda(\Omega_1). \tag{3}$$

On the other hand, (2) can be proved using (3) and a standard argument based on the homogeneity of λ. This feature will mark a difference between the classical and the Gaussian case.

Inequality (2) may be compared to the classical Brunn–Minkowski inequality: for all measurable sets A_0, A_1 and $t \in [0, 1]$ such that $(1-t)A_0 + tA_1$ is measurable,

$$|(1-t)A_0 + tA_1|^{1/n} \geq (1-t)|A_0|^{1/n} + t|A_1|^{1/n}, \tag{4}$$

where $|\cdot|$ denotes the Lebesgue measure. The analogy is strengthened by the fact that the exponent $1/n$ in (4) is the reciprocal of the homogeneity order of the involved functional, like in (2). We refer to [Gardner 2002; Schneider 2014; Artstein-Avidan et al. 2015] for extensive treatises on the Brunn–Minkowski inequality.

The concavity inequality (2) is connected to a geometric property of the solution u of (1): if Ω is convex, then u is log-concave. The log-concavity of u was first proved in [Brascamp and Lieb 1976]; different proofs can be found in [Caffarelli and Friedman 1985; Caffarelli and Spruck 1982; Korevaar 1983; Korevaar and Lewis 1987; Kennington 1985] (see also [Acker et al. 1981] for a weaker version of this result). The link between the concavity of the variational functional λ and the geometric property of the solution of the corresponding Euler–Lagrange equation is not evident, but it is revealed by the

proofs of the two facts. Indeed, more than one technique successfully exploited to prove the log-concavity of u can be adapted to demonstrate the validity of (2); see [Brascamp and Lieb 1976; Korevaar 1983; Kennington 1985]. This phenomenon can also be observed for other functionals, like the electrostatic capacity and the torsional rigidity; see for instance [Colesanti 2005]. As we will see in this paper, the eigenvalue of the Ornstein–Uhlenbeck operator provides a further example in this sense.

1.2. ***The Gaussian eigenvalue problem.*** We denote by γ the Gaussian measure in $\mathbb{R}^n$, which is defined as follows: for a measurable subset A of $\mathbb{R}^n$,

$$\gamma(A) = \frac{1}{(2\pi)^{n/2}} \int_A e^{-|x|^2/2}\, \mathrm{d}x.$$

In the sequel, $\mathrm{d}\gamma$ stands for integration with respect to γ.

Let $\Omega \subset \mathbb{R}^n$ be open and bounded, and assume that $\partial\Omega$ is Lipschitz. The Gaussian version of the principal frequency for Ω can be defined as the solution of the minimization problem

$$\lambda_\gamma(\Omega) = \inf\{R_\gamma(v) : v \in W_0^{1,2}(\Omega, \gamma),\ v \not\equiv 0\}, \tag{5}$$

where R_γ is the Gaussian Rayleigh quotient, defined as

$$R_\gamma(v) = \frac{\int_\Omega |\nabla v|^2\, \mathrm{d}\gamma}{\int_\Omega v^2\, \mathrm{d}\gamma}.$$

A standard variational argument shows that the infimum in (5) is in fact a minimum and $\lambda_\gamma(\Omega)$ coincides with the positive number λ_γ such that the following boundary value problem has a solution:

$$\begin{cases} Lu = -\lambda_\gamma u,\ u > 0, & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{6}$$

Here L is the Ornstein–Uhlenbeck operator, namely, the second-order differential operator defined on $C^2(\Omega)$ by the formula

$$Lu = \Delta u - \langle \nabla u, x \rangle.$$

The solution of problem (6) is unique, up to multiplication by positive constants (that is, λ_γ is simple). Throughout the paper, we call any solution of (6) an *eigenfunction* of Ω. We mention here that the definition of λ_γ can be extended in a natural way to unbounded sets; see, for instance, [Ehrhard 1984; Livshyts 2024b].

We will refer to λ_γ as the first Dirichlet eigenvalue of the operator L in Ω, or as the Gaussian principal frequency, and occasionally we will write it as $\lambda_\gamma(\Omega)$ to emphasize its dependency on Ω. We refer, e.g., to [Carlen and Kerce 2001; Livshyts 2024b; Herscovici and Livshyts 2024] for a general study of the Gaussian principal frequency in various contexts. Note that, similarly to the Lebesgue case, $\lambda_\gamma(\Omega_1) \geq \lambda_\gamma(\Omega_2)$ whenever $\Omega_1 \subset \Omega_2$; however, the functional λ_γ lacks homogeneity properties.

A Gaussian analogue of the Faber–Krahn inequality was established by Ehrhard [1984]; see also [Carlen and Kerce 2001] for a discussion on equality conditions. This result states, for every open set $\Omega \subset \mathbb{R}^n$,

$$\lambda_\gamma(\Omega) \geq \lambda_\gamma(H),$$

where H is a half-space such that $\gamma(K) = \gamma(H)$; for the Gaussian analogue of the Kohler-Jobin inequality, see [Herscovici and Livshyts 2024]. According to the Gaussian isoperimetric inequality (see [Sudakov and Tsirel'son 1974; Borell 1975]), half-spaces are the isoperimetric sets for the Gaussian measure; it is therefore not surprising that half-spaces also appear as minimizers for the Gaussian principal frequency (in fact, these kinds of phenomena are directly related; see for instance [Carlen and Kerce 2001]).

In this paper we establish a convexity property for the Gaussian principal frequency λ_γ. In analogy with (3), we show the following result.

Theorem 1.2. *Let Ω_0, Ω_1 be open, bounded subsets of $\mathbb{R}^n$ with Lipschitz boundary. Let $t \in [0, 1]$ and set*

$$\Omega_t = (1-t)\Omega_0 + t\Omega_1.$$

Then

$$\lambda_\gamma(\Omega_t) \le (1-t)\lambda_\gamma(\Omega_0) + t\lambda_\gamma(\Omega_1). \tag{7}$$

As already observed, since λ_γ is not homogeneous, (7) does not imply a stronger inequality such as (2). A similar fact happens for the volume: by the Prékopa–Leindler inequality (see [Leindler 1972; Prékopa 1971; 1973]), for all measurable sets A_0, A_1 and $t \in [0, 1]$ such that $(1-t)A_0 + tA_1$ is measurable, we have

$$\gamma((1-t)A_0 + tA_1) \ge \gamma(A_0)^{1-t}\gamma(A_1)^t. \tag{8}$$

Inequality (8) can be considered as the Gaussian version of the Brunn–Minkowski inequality; see also [Ehrhard 1983; Borell 2003] for the Ehrhard inequality, a stronger result. On the other hand, as it was first conjectured by Gardner and Zvavitch [2010] and then proved by Eskenazis and Moschidis [2021], for convex sets A_0 and A_1 which are symmetric with respect to the origin and for $t \in [0, 1]$, we have

$$\gamma((1-t)A_0 + tA_1)^{1/n} \ge (1-t)\gamma(A_0)^{1/n} + t\gamma(A_1)^{1/n};$$

see also [Kolesnikov and Livshyts 2021] for important contributions to the conjecture by Gardner and Zvavitch. It would be interesting to understand whether a similar improvement holds for inequality (7).

When restricting to suitably regular convex sets that are symmetric with respect to the origin, we also establish equality conditions for inequality (7).

Theorem 1.3. *Let Ω_0, Ω_1 be open, bounded and convex subsets of $\mathbb{R}^n$. Assume moreover that Ω_0 and Ω_1 are symmetric with respect to the origin, and that $\partial\Omega_0$ and $\partial\Omega_1$ are of class $C^{2,\alpha}_+$ for some $\alpha \in (0, 1)$. If, for some $t \in (0, 1)$,*

$$\lambda_\gamma((1-t)\Omega_0 + t\Omega_1) = (1-t)\lambda_\gamma(\Omega_0) + t\lambda_\gamma(\Omega_1),$$

then $\Omega_0 = \Omega_1$.

Here $\partial\Omega \in C^{2,\alpha}_+$ means that the boundary of Ω is of class $C^{2,\alpha}$ and the Gauss curvature is strictly positive at every point of $\partial\Omega$.

Remark 1.4. In fact we treat equality conditions for a larger class of sets: convex sets such that the Laplacian of an eigenfunction is strictly negative. This is the content of Theorem 5.1, of which Theorem 1.3 is a corollary. An intermediate result is Corollary 5.2, which refers to convex sets such that the maximum

of an eigenfunction is at the origin. We prefer however to emphasize Theorem 1.3, because here we have a clear and simple geometric condition on the involved sets.

Similarly to other functionals satisfying a Brunn–Minkowski-type inequality (see for instance [Bianchini and Salani 2009], and [Salani 2015] for a related general theory), inequality (2) can be used to derive a Urysohn-type inequality for λ. This is a similar inequality to the previously mentioned Faber–Krahn inequality, where the volume is replaced by the *mean width*. We recall that, given an open bounded and convex set Ω and a direction $y \in \mathbb{S}^{n-1}$, the width of Ω in the direction y is the distance between the supporting hyperplanes to Ω with outer unit normal y and $-y$; we denote it by $w(\Omega, y)$. The mean width $w(\Omega)$ of Ω is then defined as

$$w(\Omega) = \frac{1}{\mathcal{H}^{n-1}(\mathbb{S}^{n-1})} \int_{\mathbb{S}^{n-1}} w(\Omega, y)\, \mathrm{d}\mathcal{H}^{n-1}(y),$$

where $\mathcal{H}^{n-1}$ is the $(n-1)$-dimensional Hausdorff measure. Then, for every open bounded convex set Ω, we have

$$\lambda(\Omega) \geq \lambda(\Omega'),$$

where

$$\Omega' \text{ is any ball such that } w(\Omega') = w(\Omega).$$

Similarly to the Lebesgue case, we get the following result.

Corollary 1.5. *Let Ω be an open, bounded and convex set in $\mathbb{R}^n$. Then*

$$\lambda_\gamma(\Omega) \geq \lambda_\gamma(\Omega^\sharp),$$

where

$$\Omega^\sharp \textit{ is a ball centered at } 0 \textit{ such that } w(\Omega^\sharp) = w(\Omega).$$

Remark 1.6. Notice that, unlike the Lebesgue case, the ball $\Omega^\sharp$ must be centered at the origin (since the Gaussian measure is rotation invariant but not translation invariant), and, more importantly, Corollary 1.5 does not follow from the Gaussian Faber–Krahn inequality. On one hand, the log-concavity and the rotation-invariance of the Gaussian measure imply (via a similar argument) that whenever $\Omega^\sharp$ is an origin-symmetric ball such that $w(\Omega) = w(\Omega^\sharp)$, then $\gamma(\Omega) \geq \gamma(\Omega^\sharp)$. On the other hand, as we mentioned earlier, half-spaces are the minimizers in the Gaussian version of the Faber–Krahn inequality, rather than balls. So, the fact that the ball (rather than the half-space) is the extremizer in Corollary 1.5 sets it apart from the cohort of inherently Gaussian isoperimetric-type inequalities, in which the half-space is extremal. Such inequalities include, of course, the Gaussian isoperimetric inequality (see, e.g., [Artstein-Avidan et al. 2021]), the Gaussian Faber–Krahn inequality (see [Ehrhard 1983; Carlen and Kerce 2001]), the Gaussian Saint-Venant inequality (see, e.g., [Ehrhard 1983; Livshyts 2024b]), the Gaussian Kohler-Jobin inequality from [Herscovici and Livshyts 2024], and a variety of other inequalities such as the Gaussian barycenter inequality, the Gaussian dilation inequality and others; see [Livshyts 2024a]. Therefore, perhaps Theorem 1.2 is a property related to rotation invariance of the measure rather than to its special Gaussian properties.

1.3. ***Log-concavity of the solution.*** Theorem 1.2 is deeply connected to a geometric property of the first eigenfunction of the operator L, like in the Lebesgue case. This is the content of the second main result of this paper.

Theorem 1.7. *Let Ω be an open, bounded and convex subset of $\mathbb{R}^n$. Let u be an eigenfunction of Ω, that is, a solution of problem* (6) *in Ω. Then the function*

$$W = -\ln(u)$$

is convex in Ω.

1.4. ***The strategy of proofs and connections between Theorems 1.2 and 1.7.*** We start noticing that several techniques, developed to prove convexity-type properties of solutions to elliptic boundary value problems, could be adapted to prove Theorems 1.2 and 1.7; see Remark 3.2.

Moreover, Theorem 1.7 could be obtained as a consequence of [Bianchini and Salani 2009, Lemma 4]. Furthermore, the Brunn–Minkowski inequality for λ_γ, in the case of convex sets, can be deduced as an application of [Salani 2015, Lemma 5.1] (where the comparison principle is not needed). To obtain Theorem 1.2 without the assumption of convexity, we use the technique of [Salani 2015] (where indeed convexity of the domains as well as power concavity properties of the involved functions are not really needed in most of the results). Then, for the convenience of the reader, we will give also a complete proof of Theorem 1.7 as a suitable modification of the one of Theorem 1.2, to emphasize the strict relation between the two results.

Although we prove inequality (7) for open bounded sets, we are able to establish its equality conditions only under more restrictive assumptions on the involved domains: central symmetry, regularity and strong convexity of the boundary. These assumptions permit us to prove (via the constant-rank method) a stronger version of Theorem 1.7, namely that the logarithms of the involved eigenfunctions have everywhere positive-definite Hessian matrices. This property allows us to exploit effectively their Legendre–Fenchel transforms. However, we believe that it should be possible to treat the equality case of the Brunn–Minkowski inequality for the Gaussian eigenvalue in the general case and we state the following conjecture.

Conjecture 1.8. *Equality holds in* (7) *if and only if Ω_0 and Ω_1 coincide and are convex, up to removing null measure sets.*

2. Preliminaries

We work in the n-dimensional Euclidean space $\mathbb{R}^n$. We denote by $|\cdot|$ and $\langle\cdot\,,\cdot\rangle$ the Euclidean norm and the standard scalar product in $\mathbb{R}^n$, respectively. Given a subset A of $\mathbb{R}^n$, we denote by $\mathrm{cl}(A)$ the closure of A.

2.1. ***The Ornstein–Uhlenbeck operator.*** Let Ω be an open subset of $\mathbb{R}^n$ and $u \in C^2(\Omega)$. We set

$$Lu(x) = \Delta u(x) - \langle \nabla u(x), x\rangle.$$

We will refer to L as to the Ornstein–Uhlenbeck operator (in the one-dimensional case L is involved also in the Hermite differential equation).

Integration by parts shows that L is self-adjoint with respect to the Gaussian measure: if Ω is an open subset of $\mathbb{R}^n$ and $\phi, \psi \in C_c^\infty(\Omega)$, then

$$\int_\Omega \varphi L\psi \, d\gamma = -\int_\Omega \langle \nabla\varphi, \nabla\psi \rangle \, d\gamma = \int_\Omega \psi L\varphi \, d\gamma.$$

2.2. ***Existence of the eigenvalue.*** The following existence result is classical; see, e.g., [Gilbarg and Trudinger 1983; Evans 2010].

Theorem 2.1. *Let Ω be an open and bounded subset of $\mathbb{R}^n$ with Lipschitz boundary. Then there exists a unique number $\lambda_\gamma > 0$ such that the problem*

$$\begin{cases} Lu = -\lambda_\gamma u, \ u > 0, & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{9}$$

*admits a solution u (which we call an **eigenfunction** of Ω). The solution is unique up to multiplication by positive constants.*

As mentioned in the introduction, λ_γ can be characterized as

$$\lambda_\gamma = \inf \frac{\int_\Omega |\nabla v|^2 \, d\gamma}{\int_\Omega v^2 \, d\gamma},$$

where the infimum is taken over all functions $v \in W_0^{1,2}(\Omega, \gamma)$ such that

$$\int_\Omega v^2 \, d\gamma > 0.$$

In particular, any solution of (9) is a minimizer of the previous problem and vice versa.

3. The proofs of Theorems 1.2 and 1.7

We start by recalling some notions from the theory of viscosity solutions, for which we refer the reader to [Crandall et al. 1992].

Let $u, \varphi : \Omega \to \mathbb{R}$, where Ω is an open subset of $\mathbb{R}^n$. We say that φ *touches u from above at a point* $\bar{x} \in \Omega$ if $\varphi(\bar{x}) = u(\bar{x})$ and $\varphi \geq u$ in a neighborhood of $\bar{x}$. Similarly, φ *touches u from below at* $\bar{x}$ if $\varphi(\bar{x}) = u(\bar{x})$ and $\varphi \leq u$ in a neighborhood of $\bar{x}$.

Notice that if φ and ψ are twice differentiable at $\bar{x}$ and φ touches ψ from above at $\bar{x}$, then

$$\varphi(\bar{x}) = \psi(\bar{x}), \quad \nabla\varphi(\bar{x}) = \nabla\psi(\bar{x}), \quad D^2\varphi(\bar{x}) \geq D^2\psi(\bar{x}). \tag{10}$$

An upper continuous function v is a *viscosity subsolution* of the equation

$$\Delta u - \langle x, \nabla u \rangle + \lambda u = 0$$

in Ω if, for every point $\bar{x} \in \Omega$ and every test function φ touching u from above at $\bar{x}$, we have

$$\Delta\varphi(\bar{x}) - \langle \bar{x}, \nabla\varphi(\bar{x}) \rangle + \lambda\varphi(\bar{x}) \geq 0.$$

In fact, we are here interested in *distributional solutions* of (9). Luckily, viscosity solutions and distributional solutions coincide in this case by a famous result of Ishii [1995], allowing us to integrate by parts when needed, as we will see.

Remark 3.1. The definition of viscosity solutions is based on the maximum principle (and especially on the fact that if $v \in C^2(\Omega)$ attains its maximum at $x \in \Omega$, then $\nabla v(x) = 0$ and $D^2 v(x) \leq 0$) and it is particularly useful when a comparison principle holds. On the other hand, here we are treating an eigenvalue problem in which no comparison principle can work since the solution is insensitive to positive scalar multiplication; that is, if u is a solution of problem (9), then au is a solution too for any $a > 0$.

Proof of Theorem 1.2.. The proof largely follows the steps of [Salani 2015, Lemma 5.1], which in turn generalizes ideas introduced in [Bianchini and Salani 2013] and is strongly inspired by [Alvarez et al. 1997].

Let Ω_0 and Ω_1 be open and bounded, with Lipschitz boundary. Let u_i be the solution of problem (6) for $i = 0, 1$. For simplicity, set

$$\lambda_i = \lambda_\gamma(\Omega_i) \ \text{ for } i = 0, 1, \quad \lambda_t = (1-t)\lambda_0 + t\lambda_1.$$

Moreover, for $x \in \Omega_t$ let

$$u_t(x) = \max\{u_0(x_0)^{1-t} u_1(x_1)^t : x_i \in \overline{\Omega}_i \text{ for } i = 0, 1,\ x = (1-t)x_0 + t x_1\}. \tag{11}$$

Notice that $u_t \in C(\mathrm{cl}(\Omega_t))$ and that

$$u_t > 0 \ \text{ in } \Omega_t, \quad u_t = 0 \ \text{ on } \partial\Omega_t,$$

as $u_i \in C(\mathrm{cl}(\Omega_i))$, $u_i > 0$ in Ω_i and $u_i = 0$ on $\partial\Omega_i$ for $i = 0, 1$. Furthermore, for every $\bar{x} \in \mathrm{cl}(\Omega_t)$ there exist $\bar{x}_0 \in \mathrm{cl}(\Omega_0)$, $\bar{x}_1 \in \mathrm{cl}(\Omega_1)$ where the maximum in the definition of (11) is attained; that is

$$\bar{x} = (1-t)\bar{x}_0 + t\bar{x}_1, \quad u_t(\bar{x}) = u_0(\bar{x}_0)^{1-t} u_1(\bar{x}_1)^t. \tag{12}$$

Notice that if $\bar{x} \in \partial\Omega_t$, then clearly $\bar{x}_i \in \partial\Omega_i$ for $i = 0, 1$, while if $\bar{x} \in \Omega_t$, then $\bar{x}_i \in \Omega_i$ for $i = 0, 1$ since $u_t(\bar{x}) > 0$. In the latter case, as u_0 and u_1 are differentiable at x_0 and x_1, respectively, we can also easily see that

$$\frac{\nabla u_0(\bar{x}_0)}{u_0(\bar{x}_0)} = \frac{\nabla u_1(\bar{x}_1)}{u_1(\bar{x}_1)} =: \theta. \tag{13}$$

The previous equation follows by applying the Lagrange multipliers theorem to the maximization problem

$$\max_{x_0, x_1} u_0(x_0)^{1-t} u_1(x_1)^t$$

when x_0 and x_1 are subject to the constraint

$$(1-t)x_0 + t x_1 = x.$$

We want to prove that u_t is a viscosity subsolution of problem (6) when $\Omega = \Omega_t$ and $\lambda = \lambda_t$. To show this, we construct, for every point $\bar{x} \in \Omega_t$, a C^2 function ψ touching u_t from below at $\bar{x}$ and such that

$$\Delta\psi(\bar{x}) - \langle \bar{x}, \nabla\psi(\bar{x})\rangle + \lambda_t \psi(\bar{x}) = 0.$$

Indeed, this yields that every test function ϕ touching u_t from above at $\bar{x}$ must also touch ψ from above at $\bar{x}$, and (10) gives

$$\Delta\phi(\bar{x}) - \langle \bar{x}, \nabla\phi(\bar{x})\rangle + \lambda_t\phi(\bar{x}) \geq 0,$$

as desired. In fact, we need to construct ψ only for the points $\bar{x}$ such that $u_t(\bar{x}) > u(\bar{x})$; indeed, if $u_t(\bar{x}) = u(\bar{x})$, every test function touching u_t from above at $\bar{x}$ also touches u from above at $\bar{x}$, and the latter inequality is straightforward. Then, let $\bar{x} \in \Omega_t$ be such that $u_t(\bar{x}) > u(\bar{x})$, and let $\bar{x}_0 \in \Omega_0$ and $\bar{x}_1 \in \Omega_1$ be the points given by (12). In a sufficiently small neighborhood B of $\bar{x}$, we define

$$\psi(x) = u_0(x - \bar{x} + \bar{x}_0)^{1-t} u_1(x - \bar{x} + \bar{x}_1)^t.$$

Then, by the choice of $\bar{x}_0$ and $\bar{x}_1$ and the definition of u_t, it is easily seen that ψ touches u_t from below at $\bar{x}$. Furthermore, $\psi \in C^2(B)$ and we can calculate

$$\nabla\psi(x) = \psi(x)\left[(1-t)\frac{\nabla u_0(x-\bar{x}+\bar{x}_0)}{u_0(x-\bar{x}+\bar{x}_0)} + t\frac{\nabla u_1(x-\bar{x}+\bar{x}_1)}{u_1(x-\bar{x}+\bar{x}_1)}\right]$$

and

$$\begin{aligned} D^2\psi(x) = {} & \frac{\nabla\psi(x)\times\nabla\psi(x)}{\psi(x)} \\ & - \psi(x)\left[(1-t)\frac{\nabla u_0(x-\bar{x}+\bar{x}_0)\times\nabla u_0(x-\bar{x}+\bar{x}_0)}{u_0(x-\bar{x}+\bar{x}_0)^2} + t\frac{\nabla u_1(x-\bar{x}+\bar{x}_1)\times\nabla u_1(x-\bar{x}+\bar{x}_1)}{u_1(x-\bar{x}+\bar{x}_1)^2}\right] \\ & + \psi(x)\left[(1-t)\frac{D^2u_0(x-\bar{x}+\bar{x}_0)}{u_0(x-\bar{x}+\bar{x}_0)} + t\frac{D^2u_1(x-\bar{x}+\bar{x}_1)}{u_1(x-\bar{x}+\bar{x}_1)}\right], \end{aligned}$$

whence

$$\nabla\psi(\bar{x}) = \psi(\bar{x})\theta \tag{14}$$

and

$$D^2\psi(\bar{x}) = \psi(\bar{x})\left[(1-t)\frac{D^2u_0(\bar{x}_0)}{u_0(\bar{x}_0)} + t\frac{D^2u_1(\bar{x}_1)}{u_1(\bar{x}_1)}\right], \tag{15}$$

where in the former we have used (13) and in the latter we have used the former and again (13). From (14) and (15) we finally get

$$\begin{aligned} \Delta\psi(\bar{x}) &= \psi(\bar{x})\left[(1-t)\frac{\Delta u_0(\bar{x}_0)}{u_0(\bar{x}_0)} + t\frac{\Delta u_1(\bar{x}_1)}{u_1(\bar{x}_1)}\right] \\ &= \psi(\bar{x})\left[(1-t)\frac{\langle \bar{x}_0, \nabla u_0(\bar{x}_0)\rangle - \lambda_0 u_0(\bar{x}_0)}{u_0(\bar{x}_0)} + t\frac{\langle \bar{x}_1, \nabla u_1(\bar{x}_1)\rangle - \lambda_1 u_1(\bar{x}_1)}{u_1(\bar{x}_1)}\right] \\ &= \psi(\bar{x})\left[(1-t)\left(\left\langle \bar{x}_0, \frac{\nabla u_0(\bar{x}_0)}{u_0(\bar{x}_0)}\right\rangle - \lambda_0\right) + t\left(\left\langle \bar{x}_1, \frac{\nabla u_1(\bar{x}_1)}{u_1(\bar{x}_1)}\right\rangle - \lambda_1\right)\right] \\ &= \psi(\bar{x})\left[\langle (1-t)\bar{x}_0 + t\bar{x}_1, \theta\rangle - ((1-t)\lambda_0 + t\lambda_1)\right] \\ &= \langle \bar{x}, \nabla\psi(\bar{x})\rangle - \lambda_t\psi(\bar{x}). \end{aligned} \tag{16}$$

In the above series of equalities we have made use of the equations satisfied by u_0 and u_1, of (13) and of (14).

So far, we proved that u_t satisfies the inequality

$$\Delta u_t(x) - \langle \nabla u_t(x), x\rangle \geq -\lambda_t u_t \quad \text{in } \Omega_t \tag{17}$$

in the viscosity sense, and consequently also in the distributional sense; see [Ishii 1995, Theorem 1].

Now multiply (17) by $-u_t(x)$ to get

$$-u_t \Delta u_t(x) + u_t \langle \nabla u_t(x), x\rangle \leq \lambda_t u_t^2$$

and integrate in $\mathrm{d}\gamma$ on Ω_t. Then, taking into account that

$$-e^{-|x|^2/2} u_t \Delta u_t + u_t \langle \nabla u_t(x), x\rangle e^{-|x|^2/2} = -\mathrm{div}(e^{-|x|^2/2} u_t \nabla u_t) + |\nabla u_t|^2 e^{-|x|^2/2}$$

and that u_t vanishes on $\partial\Omega_t$, an integration by parts gives

$$R_\gamma(u_t) \leq \lambda_t, \tag{18}$$

whence (7) follows by the very definition of $\lambda_\gamma(\Omega_t)$. □

Proof of Theorem 1.7. The proof could be obtained by applying [Bianchini and Salani 2009, Theorem 1]. However, it is useful to give here a complete proof to emphasize the relation between the Brunn–Minkowski inequality of the eigenvalue and the log-concavity property of the eigenfunction; for this, see also [Salani 2015]. We repeat the argument of the previous proof with $\Omega_0 = \Omega_1 = \Omega$, and consequently $u_0 = u_1 = u$ and $\lambda_t = \lambda_\gamma(\Omega)$. For $t \in [0, 1]$, the function u_t is defined similarly as in (11):

$$u_t(x) = \max\{u(x_0)^{1-t} u(x_1)^t : x_0, x_1 \in \Omega,\ (1-t)x_0 + t x_1 = x\}.$$

Notice that the *log-concave envelope* u^* of u, i.e., the smallest log-concave function greater than or equal to u in Ω, is given by

$$u^*(x) = \sup\{u_t(x) : t \in [0, 1]\},$$

and that $u \leq u_t \leq u^*$ in Ω for every $t \in [0, 1]$, while

$$u \text{ is log-concave (i.e., it coincides with } u^*) \text{ if and only if } u = u_t \text{ for every } t \in [0, 1].$$

With the same steps as before, it can be proved that u_t satisfies — in the viscosity (and distributional) sense — the inequality

$$\Delta u_t(x) - \langle \nabla u_t(x), x\rangle + \lambda_\gamma(\Omega) u_t \geq 0 \tag{19}$$

in Ω. Now we multiply by $-u_t$ and integrate by parts in $\mathrm{d}\gamma$, as in the proof of Theorem 1.2, to get

$$R_\gamma(u_t) \leq \lambda_\gamma(\Omega),$$

whence indeed $R_\gamma(u_t) = \lambda_\gamma(\Omega)$ and u_t is an eigenfunction, and hence $u_t = au$ for some $a > 0$. On the other hand, by the very definition if u_t, we have

$$\max_{\overline{\Omega}} u_t = \max_{\overline{\Omega}} u,$$

whence $a = 1$ and u_t coincides with u in Ω, as desired. □

Remark 3.2. Under suitable assumptions, the log-concavity of the eigenfunction of a convex domain can be deduced in several other ways, for instance by means of the concavity function (first developed by Korevaar [1983] and Kennington [1985]) or via the constant-rank method (see [Acker et al. 1981; Caffarelli and Friedman 1985; Korevaar and Lewis 1987], and [Guan and Xu 2013] for a survey on more recent developments of this method). We would also like to thank E. Milman for showing us a further elegant proof (personal communication, 2024) based on a clever transform from the Ornstein–Uhlenbeck to the Schrödinger operator and on [Brascamp and Lieb 1976, Theorem 6.1].

Proof of Corollary 1.5. The proof closely follows the steps of [Bianchini and Salani 2009, Corollary 2.2] and [Salani 2015, Section 6]. The mean width is linear with respect to Minkowski addition and clearly invariant with respect to rigid motions. Hence, given a sequence of rotations (about the origin) $\{\rho_j\}_{j\in\mathbb{N}} \subset \mathrm{SO}(n)$, we define the *rotation mean* $\Omega_m^\sharp$ of Ω as

$$\Omega_m^\sharp = \frac{1}{m}(\rho_1\Omega + \cdots + \rho_m\Omega),$$

and we get that

$$w(\Omega_m^\sharp) = w(\Omega) \quad \text{for every } m \in \mathbb{N}.$$

Moreover, the rotation-invariance of the Gaussian measure yields

$$\lambda_\gamma(\rho_j\Omega) = \lambda_\gamma(\Omega) \quad \text{for every } j \in \mathbb{N}. \tag{20}$$

Then, inequality (7) combined with (20) shows that

$$\lambda_\gamma(\Omega_m^\sharp) \text{ is a decreasing sequence.}$$

On the other hand, by a theorem of Hadwiger (see [Schneider 1993, Theorem 3.3.2]) there exists a sequence of rotation means of Ω converging in Hausdorff metric to a ball $\Omega^\sharp$, centered at the origin, with diameter equal to the mean width $w(\Omega)$ of Ω. Finally, the conclusion follows in view of the fact that λ_γ is continuous in Hausdorff metric. □

4. Strong log-concavity of the eigenfunction

In this section we see how, under additional symmetry and regularity assumptions on Ω, we can improve the conclusion of Theorem 1.7 in the following sense: the Hessian of $-\ln(u)$ is positive definite in Ω. This will allow us to investigate equality conditions of (7) in some restricted case and then to prove Theorem 1.3 in the final section of the paper.

We say that the boundary of an open set $\Omega \subset \mathbb{R}^n$ is of class $C_+^{3,\alpha}$, for some $\alpha \in (0, 1)$, if it is of class $C^{3,\alpha}$ and the Gauss curvature is strictly positive at every point of $\partial\Omega$.

Theorem 4.1. *Let Ω be an open, bounded and convex subset of $\mathbb{R}^n$ with boundary of class $C^{3,\alpha}_+$ for some $\alpha \in (0,1)$, and let u be a solution of problem* (6) *in Ω. Assume that*

$$\Delta u < 0 \quad \text{in } \Omega. \tag{21}$$

Then the function

$$W = -\ln(u)$$

*is **strongly convex** in Ω, i.e.,*

$$D^2 W(x) > 0 \quad \text{for all } x \in \Omega. \tag{22}$$

The proof is based on the so-called constant-rank method introduced in [Acker et al. 1981] and subsequently developed in [Caffarelli and Friedman 1985; Korevaar and Lewis 1987]. We also point out that recently this approach led to new important results in the theory of geometric properties of solutions to elliptic and parabolic PDEs. We refer the reader to recent papers about this subject, like [Ma and Ou 2010; Guan and Xu 2013].

Before proving Theorem 4.1 (whose proof will be indeed a straightforward application of Theorem 1.7 and two technical results, namely Propositions 4.4 and 4.5, to be proven hereafter), let us see two simple, but interesting, corollaries.

Corollary 4.2. *Let Ω be an open, bounded and convex subset of $\mathbb{R}^n$ with boundary of class $C^{3,\alpha}_+$ for some $\alpha \in (0,1)$, and let u be a solution of problem* (6) *in Ω. Assume that $u(0) = \max_{\overline{\Omega}} u$. Then the function*

$$W = -\ln(u)$$

*is **strongly convex** in Ω.*

Proof. By Theorem 1.7, u is log-concave, and hence its level sets are all convex. Since 0 is a maximum point, it is contained in every superlevel set, which is then starshaped with respect to 0. This implies $\langle x, \nabla u\rangle \le 0$ in Ω, whence $\Delta u < 0$ and we can apply Theorem 4.1. □

Corollary 4.3. *Let Ω be an open, bounded and convex subset of $\mathbb{R}^n$ with boundary of class $C^{3,\alpha}_+$ for some $\alpha \in (0,1)$, and let u be a solution of problem* (6) *in Ω. Assume furthermore that Ω is symmetric with respect to the origin. Then the function*

$$W = -\ln(u)$$

*is **strongly convex** in Ω.*

Proof. The symmetry of Ω easily implies that u is also symmetric with respect to the origin, which is then a maximum point of u and we can apply the previous corollary. □

4.1. ***Boundary behavior.*** We will first show a result on the boundary behavior that will be used in the proof of the above theorem as a companion of the constant-rank theorem.

Given $X = (X_1, \dots, X_n)$, $Y = (Y_1, \dots, Y_n) \in \mathbb{R}^n$, we denote by $X \otimes Y$ the $n \times n$ matrix with ij-th entry equal to $X_i Y_j$ for $i, j = 1, \dots, n$. Moreover, if M is a square symmetric matrix, the notation $M > 0$ means that M is positive definite.

Proposition 4.4. *Let Ω be an open, bounded and convex subset of $\mathbb{R}^n$; assume that $\partial\Omega$ is of class $C^{3,\alpha}$, $\alpha \in (0,1)$, and all its sectional curvatures are strictly positive at each point of $\partial\Omega$. Let $w \in C^3(\overline{\Omega})$ be such that*

$$w > 0 \ \text{ in } \Omega, \quad w = 0 \ \text{ on } \partial\Omega, \quad \frac{\partial w}{\partial \nu} < 0 \ \text{ on } \partial\Omega,$$

where ν denotes the outer unit normal to $\partial\Omega$. Then there exists $\varepsilon > 0$ such that

$$\nabla w(x) \otimes \nabla w(x) - w(x) D^2 w(x) > 0 \tag{23}$$

for every $x \in \overline{\Omega}$ such that $\operatorname{dist}(x, \partial\Omega) < \varepsilon$. In particular, setting

$$W = -\ln(w),$$

we have

$$D^2 W(x) > 0 \quad \text{for all } x \in \{y \in \Omega : \operatorname{dist}(y, \partial\Omega) < \varepsilon\}.$$

Proof. As

$$\frac{\partial w}{\partial \nu} \neq 0$$

on $\partial\Omega$, by continuity there exists $\varepsilon_0 > 0$ such that

$$\frac{m}{2} \leq |\nabla w(x)| \leq 2M$$

for every x which satisfies $\operatorname{dist}(x, \partial\Omega) \leq \epsilon_0$.

Let $x \in \overline{\Omega}$ be such that $\nabla w(x) \neq 0$ and $t = w(x)$; we denote by $\kappa_i(x),\ i = 1, \ldots, (n-1)$, the principal curvatures of the level set $\{w = t\}$ at the point x. In particular, if $x \in \partial\Omega$, these are the principal curvatures of $\partial\Omega$ at x. Let $\eta > 0$ be such that $\kappa_i(x) \geq \eta$ for every $i = 1, \ldots, (n-1)$ and for every $x \in \partial\Omega$. Up to reducing the value of ε_0, we may assume that

$$\kappa_i(x) \geq \frac{\eta}{2}$$

for every x such that $\operatorname{dist}(x, \partial\Omega) \leq \epsilon_0$ and for every $i = 1, \ldots, n-1$. Let

$$\Omega_0 := \{x \in \overline{\Omega} : 0 \leq \operatorname{dist}(x, \partial\Omega) \leq \varepsilon_0\}.$$

For $k = 1, \ldots, n-1$, we denote by $\sigma_k(x)$ the k-th elementary symmetric function of the numbers $\kappa_i(x)$:

$$\sigma_k(x) = \sum_{1 \leq i_1 < i_2 \cdots < i_k \leq n-1} \kappa_{i_1} \kappa_{i_2} \cdots \kappa_{i_k}.$$

For completeness, we set $\sigma_0(x) = 1$ for every x. For $x \in \overline{\Omega}$ and for $k \in \{1, \ldots, n\}$, let

$$S_k(D^2 W(x))$$

be the k-th elementary symmetric function of the eigenvalues of $D^2 W(x)$, with $S_0(D^2 W(x)) = 1$ for every x. We are going to prove that, for a suitable $\varepsilon > 0,\ \varepsilon \leq \varepsilon_0$, we have

$$S_k(D^2 W(x)) > 0 \tag{24}$$

for every $k = 1, \dots, n$, and for every $x \in \overline{\Omega}$ such that $\operatorname{dist}(x, \partial\Omega) \leq \varepsilon$. Condition (24) is equivalent to

$$D^2 W(x) > 0,$$

hence this will conclude the proof.

Let us fix $x_0 \in \Omega_0$. We may choose a coordinate system such that the first $n-1$ coordinates are parallel to the directions of the principal curvatures of the level set $\{W = W(x_0)\}$, and the n-th coordinate is parallel to $\nabla W(x_0)$. With respect to this system we have

$$D^2 W(x_0) = \begin{pmatrix} \kappa_1(x_0)|\nabla W(x_0)| & 0 & 0 & \dots & 0 & W_{1n}(x_0) \\ 0 & \kappa_2(x_0)|\nabla W(x_0)| & 0 & \dots & 0 & W_{2n}(x_0) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \kappa_{n-1}(x_0)|\nabla W(x_0)| & W_{(n-1)n}(x_0) \\ W_{n1}(x_0) & W_{n2}(x_0) & W_{n3(x_0)} & \dots & W_{n(n-1)}(x_0) & W_{nn}(x_0) \end{pmatrix}.$$

Moreover,

$$|\nabla W(x_0)| = \frac{|\nabla w(x_0)|}{w(x_0)}, \quad W_{nj}(x_0) = W_{jn}(x_0) = -\frac{w_{nj}(x_0)}{w(x_0)}, \quad W_{nn}(x_0) = \frac{|\nabla w(x_0)|^2}{w^2(x_0)} - \frac{w_{nn}(x_0)}{w(x_0)}.$$

The quantity $S_k(D^2 W(x_0))$ can be expressed as

$$S_k(D^2 W(x_0)) = \frac{1}{k} \sum_{\substack{1 \leq i_1 < \dots < i_k \leq n \\ 1 \leq j_1 < \dots < j_k \leq n}} \delta \begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} W_{i_1 j_1} \cdots W_{i_k j_k}, \tag{25}$$

where the symbol

$$\delta \begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$$

is equal to 1 (resp. to -1) if $(i_1, \dots, i_k)$ is an even (resp. odd) permutation of $(j_1, \dots, j_k)$, and it is 0 otherwise.

We write

$$S_k(D^2 W(x_0)) = A + B,$$

where A contains all the terms of the sum in (25) such that $i_k = j_k = n$ and B contains all the other terms. Then

$$A = \frac{1}{k} |\nabla W(x_0)|^{k-1} W_{nn}(x_0) \sigma_{k-1}(x_0) = \frac{|\nabla w(x_0)|^{k-1}}{k w(x_0)^{k+1}} \big(|\nabla w(x_0)|^2 - w(x_0) w_{nn}(x_0) \big) \sigma_{k-1}(x_0).$$

On the other hand

$$B = \frac{1}{w(x_0)^k} F(x_0),$$

where $F = F(x)$, $x \in \Omega_0$, is the sum of products of first and second derivatives of w. In particular there exists a constant $N > 0$ such that

$$|F(x)| \leq N$$

for every $x \in \Omega_0$. Hence, in an arbitrary point $x \in \Omega_0$, we have

$$S_k(D^2W(x)) = \frac{1}{kw(x)^{k+1}}\big[|\nabla w(x)|^{k+1}\sigma_{k-1}(x) + kw(x)F(x)\big].$$

Recall that, in Ω_0,

$$|\nabla w(x)|^{k+1}\sigma_{k-1}(x) \geq cm^{k+1}\eta^{k-1}$$

for some dimensional constant $c > 0$, where $m > 0$ is such that $|\nabla w(x)| \geq m$ for every $x \in \partial\Omega$. As

$$\lim_{x\to\partial\Omega} w(x) = 0,$$

there exists $\varepsilon \in (0, \varepsilon_0)$ such that

$$0 \leq w(x) \leq \frac{cm^{k+1}\eta^{k-1}}{4N}$$

for every x such that $\operatorname{dist}(x, \partial\Omega) \leq \varepsilon$. We have proved that

$$S_k(D^2W(x)) > 0$$

for every x such that $\operatorname{dist}(x, \partial\Omega) \leq \varepsilon$. □

4.2. ***The constant-rank theorem.*** The following result is a slight generalization of the constant-rank theorem proved in [Korevaar and Lewis 1987, Theorem 1].

Proposition 4.5. *Let $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be a function of class $C^2(\mathbb{R} \times \mathbb{R}^n)$. Assume that*

$$F_{tt}(t, y) \leq 0 \quad \textit{for all } (t, y) \in \mathbb{R} \times \mathbb{R}^n.$$

Let Ω be an open, bounded and convex subset of $\mathbb{R}^n$. Let $v \in C^4(\Omega)$ be a solution of the equation

$$\Delta v = F(v, \nabla v) + \langle \nabla v(x), x\rangle \quad \textit{in } \Omega. \tag{26}$$

Assume moreover that

$$\Delta v > 0 \quad \textit{in } \Omega.$$

If v is convex in Ω, then D^2v has constant rank in Ω.

Proof. We closely follow the scheme of [Korevaar and Lewis 1987]. Let x_0 be a point where the rank of D^2v is at a minimum. If $r = n$, there is nothing to prove. Hence, for the rest of the proof, we assume that

$$\operatorname{rank}(D^2v(x_0)) = r < n.$$

For $x \in \Omega$, we denote by $\varphi(x)$ the elementary symmetric function of the eigenvalues of $D^2v(x)$ of order $(r + 1)$. Then

$$\varphi(x) \geq 0 \text{ in } \Omega \quad \text{and} \quad \varphi(x_0) = 0.$$

We will prove that there exist positive constants c_1 and c_2 such that

$$\Delta\varphi(x) \leq c_1|\nabla\varphi(x)| + c_2\varphi(x) \quad \text{in a neighborhood of } x_0. \tag{27}$$

By the strong maximum principle, this implies that $\varphi \equiv 0$ in a neighborhood of x_0. Hence the set

$$\Omega' = \{x \in \Omega : \varphi(x) = 0\} \subset \Omega$$

is open and, obviously, closed in Ω; consequently, as Ω is connected, Ω' coincides with Ω. This implies that

$$\operatorname{rank}(D^2 v(x)) \equiv r \quad \text{in } \Omega.$$

Adopting the notation of [Korevaar and Lewis 1987] (and previously introduced in [Caffarelli and Friedman 1985]), given two functions f and g in $C^1(\Omega)$ and $x \in \Omega$, we write

$$f(x) \lesssim g(x)$$

if there exist $c_1, c_2 > 0$ such that

$$f(x) - g(x) \le c_1 |\nabla \varphi(x)| + c_2 \varphi(x).$$

The notation $f(x) \sim g(x)$ means that $f(x) \lesssim g(x)$ and $g(x) \lesssim f(x)$. If we omit the variable x in any of the previous relations, we indicate that the relation holds for every x in a neighborhood of x_0, with constants c_1 and c_2 independent of x.

Hence (27) can be written as

$$\Delta \varphi \lesssim 0. \tag{28}$$

In what follows, "locally" means in a suitable neighborhood of x_0. The matrix $D^2 v(x_0)$ has r positive eigenvalues (and $n-r$ zero eigenvalues). Hence we may assume that in a neighborhood $\mathcal{U}$ of x_0, r eigenvalues of $D^2 v(x)$ are bounded from below by some constant $c > 0$. Let $z \in \mathcal{U}$, and choose a coordinate system such that $D^2 v(z)$ is a diagonal matrix (note that the equation satisfied by v is invariant under rotations). Then we may assume that

$$v_{jj}(z) \ge c, \quad j = 1, \dots, r.$$

In what follows, we will use some formulas obtained in the proof of Theorem 1 in [Korevaar and Lewis 1987]. A comment is in order: the equation considered in [loc. cit.] is different from (26). On the other hand, the formulas that we will use, involving the derivatives of the solution v, are independent of the equation.

We set $G = \{1, \dots, r\}$ and $B = \{r+1, \dots, n\}$. The following relation is [loc. cit., (3.6)]:

$$\Delta \varphi \sim Q \sum_{i \in B} (\Delta v)_{ii} - 2 \sum_{i \in B} \sum_{j \in G} |\nabla v_{ij}|^2 Q_j - R \sum_{i,j \in B} |\nabla v_{ij}|^2, \tag{29}$$

where

$$Q = \prod_{j \in G} v_{jj}, \quad Q_j = \frac{Q}{v_{jj}}, \quad j \in G, \quad R = \sum_{j \in G} Q_j.$$

Note that, by the convexity of v, $Q \ge 0$ and $Q_j \ge 0$ for every $j \in G$, $R \ge 0$. Consequently

$$\Delta \varphi \lesssim Q \sum_{i \in B} (\Delta v)_{ii}. \tag{30}$$

As established in [loc. cit., (3.8)],

$$\sum_{i\in B}(F(v,\nabla v))_{ii} \sim F_{tt}(v,\nabla v)\sum_{i\in B} v_i^2. \tag{31}$$

Now let $H:\Omega\to\mathbb{R}$ be defined by

$$H(x)=\langle \nabla v(x), x\rangle.$$

Then, for every $i=1,\dots,n$,

$$H_{ii}=2v_{ii}+\sum_{k=1}^{n} x_k v_{kii}.$$

Therefore

$$\sum_{i\in B} H_{ii}=2\sum_{i\in B} v_{ii}+\sum_{k=1}^{n} x_k \sum_{i\in B} v_{kii}. \tag{32}$$

On the other hand (see again [loc. cit., (3.3) and (3.4)]),

$$\sum_{i\in B} v_{ii}\sim 0 \quad\text{and}\quad \sum_{i\in B} v_{kii}\sim 0 \tag{33}$$

for every $k=1,\dots,n$.

Now (28) follows from (30), (31), (32) and (33). □

Proof of Theorem 4.1. Set $v=-\log u$ and write the equation for v:

$$\Delta v=\lambda_\gamma+|\nabla v|^2+\langle x,\nabla v\rangle.$$

Notice that

$$\Delta v=-\frac{\Delta u}{u}+\frac{|\nabla u|^2}{u^2}>0 \quad \text{in } \Omega,$$

thanks to assumption (21). Then the proof is a straightforward consequence of the combination of Theorem 1.7 and Propositions 4.5 and 4.4. □

5. The equality conditions of Theorem 1.2

Hereafter we will discuss the equality case of inequality (7), first giving a more general result, then obtaining as a consequence an intermediate result, which in turn has Theorem 1.3 as a corollary.

With this aim, we restrict to a special class of convex sets, so that we can exploit Theorem 4.1 and follow similar ideas to those used in [Colesanti 2005, Section 5] for the standard case, that is, when the Gauss measure is replaced by the Lebesgue measure.

Theorem 5.1. *Let Ω_0 and Ω_1 be open, bounded and convex subsets of $\mathbb{R}^n$ with boundary of class $C_+^{2,\alpha}$ for some $\alpha\in(0,1)$. Let u_0 and u_1 be solutions of* (6) *with $\Omega=\Omega_0$ and $\Omega=\Omega_1$, respectively, and assume that*

$$\Delta u_0<0 \ \text{ in } \Omega_0 \quad \text{and} \quad \Delta u_1<0 \ \text{ in } \Omega_1. \tag{34}$$

Let $t \in (0,1)$ and set

$$\Omega_t = (1-t)\Omega_0 + t\Omega_1.$$

Then, equality holds in (7)*, i.e.,*

$$\lambda_\gamma(\Omega_t) = (1-t)\lambda_\gamma(\Omega_0) + t\lambda_\gamma(\Omega_1), \tag{35}$$

only if Ω_0 and Ω_1 coincide up to a translation.

Proof. First of all, notice that, by (18), if (35) holds, then the function u_t defined in (11) must be an eigenfunction of Ω_t, that is, it must solve (6) in Ω_t.

For $i = 0, 1$, let u_i be an eigenfunction of Ω_i, that is a solution of problem (6) in Ω_i. We set

$$w_i = -\ln(u_i).$$

Then w_i satisfies

$$\begin{cases} \Delta w_i(x) = \lambda_i + |\nabla w_i(x)|^2 + \langle \nabla w_i(x), x\rangle & \text{in } \Omega_i, \\ \lim_{x\to\partial\Omega_i} w_i(x) = +\infty. \end{cases} \tag{36}$$

Here, for simplicity, we denote by λ_i the eigenvalue of Ω_i for $i = 0, 1, t$. By Theorem 1.7, we know that w_i is convex, and hence the boundary condition implies

$$\nabla w_i(\Omega_i) = \{\nabla w_i(x) : x \in \Omega_i\} = \mathbb{R}^n. \tag{37}$$

By Theorem 4.1, we further know that w_i is strongly convex in Ω_i, i.e., $D^2 w_i(x) > 0$ for every $x \in \Omega_i$. Now we consider the standard Legendre conjugate of w_i,

$$v_i(y) = w_i^*(y) = \sup_{x\in\Omega_i} \langle x, y\rangle - w_i(x)$$

(note that the superscript * for functions has been already used with a different meaning in Section 3).

By (37), the domain of v_i is $\mathbb{R}^n$. The strict positivity of $D^2 w_i$ on Ω_i implies that $v_i \in C^2(\mathbb{R}^n)$; moreover ∇w_i and ∇v_i are diffeomorphisms between Ω_i and $\mathbb{R}^n$, inverse to each other, and finally

$$D^2 w_i(x) = \big(D^2 v_i(\nabla w_i(x))\big)^{-1}. \tag{38}$$

We set

$$w_t = ((1-t)v_0 + t v_1)^* = ((1-t)w_0^* + t w_1^*)^*. \tag{39}$$

The function w_t is defined and convex in Ω_t and, as is well known, can also be defined as the inf-convolution of w_0 and w_1, that is,

$$w_t(x_t) = \inf\big\{(1-t)w_0(x_0) + t w_1(x_1) : x_0 \in \Omega_0,\, x_1 \in \Omega_1,\, x_t = (1-t)x_0 + t x_1\big\} \tag{40}$$

for every $x_t \in \Omega_t := (1-t)\Omega_0 + t\Omega_1$. Then we see that

$$w_t = -\log(u_t), \tag{41}$$

where u_t is the function defined in (11).

By the boundary condition satisfied by w_0 and w_1,

$$\lim_{x\to\partial\Omega_t} w_t(x) = +\infty. \tag{42}$$

Moreover, as $v_i \in C^2(\mathbb{R}^n)$ and $D^2 v_i > 0$ in $\mathbb{R}^n$ for $i = 0, 1$, we have that $v_t := (w_t)^* \in C^2(\mathbb{R}^n)$ and $D^2 v_t > 0$ in $\mathbb{R}^n$.

By (40) and (42), for every $x_t \in \Omega_t$ there exist (unique) $x_0 \in \Omega_0,\ x_1 \in \Omega_1$ such that

$$x_t = (1-t)x_0 + t x_1, \quad w_t(x_t) = (1-t)w_0(x_0) + t w_1(x_1).$$

Using the Lagrange multipliers theorem (see also [Colesanti 2005, pp. 126–127]), it follows that

$$\nabla w_t(x_t) = \nabla w_1(x_1) = \nabla w_0(x_0).$$

Moreover, by (38) and (39) (see also [Colesanti 2005, p. 127]), we have

$$D^2 w_t(x_t) = \big((1-t)(D^2 w_0(x_0))^{-1} + t(D^2 w_1(x_1))^{-1}\big)^{-1}.$$

Next, we consider the function

$$M \to \operatorname{tr}(M^{-1})$$

as defined in the space of symmetric positive definite square matrices of order n. This function is convex; see, for instance, [Colesanti and Salani 2003, Appendix, Lemma 4] or [Alvarez et al. 1997, Appendix]. Hence

$$\Delta w_t(x_t) \geq (1-t)\Delta w_0(x_0) + t\Delta w_1(x_1), \tag{43}$$

that is,

$$\Delta w_t(x_t) \geq (1-t)\lambda_0 + t\lambda_1 + (1-t)|\nabla w_0(x_0)|^2 + t|\nabla w_1(x_1)|^2 + (1-t)\langle \nabla w_0(x_0), x_0\rangle + t\langle \nabla w_1(x_1), x_1\rangle,$$

which, by the previous considerations, implies

$$\Delta w_t(x_t) \geq (1-t)\lambda_0 + t\lambda_1 + |\nabla w_t(x_t)|^2 + \langle \nabla w_t(x_t), x_t\rangle. \tag{44}$$

On the other hand, as observed at the very beginning of the proof, the function

$$u_t = e^{-w_t}$$

satisfies

$$\Delta u_t = -((1-t)\lambda_0 + t\lambda_1)u_t + \langle \nabla u_t, x\rangle \quad \text{in } \Omega_t, \tag{45}$$

which implies the equality in (44) and in turn in (43). Then equality holds in the differential inequality in problem (45). Going backwards through the previous steps of the proof, we find that, by the equality conditions in [Colesanti and Salani 2003, Appendix, Lemma 4],

$$D^2 w_0(x_0) = D^2 w_1(x_1).$$

As a further consequence we deduce

$$D^2 v_0(y) = D^2 v_1(y) \quad \text{for all } y \in \mathbb{R}^n,$$

and then

$$\nabla v_0(y) = \nabla v_1(y) + a$$

for some $a \in \mathbb{R}^n$. On the other hand,

$$\Omega_i = \nabla v_i(\mathbb{R}^n), \quad i = 0, 1,$$

(as ∇v_i and ∇w_i are inverse to each other). We conclude that

$$\Omega_0 = \Omega_1 + a. \qquad \square$$

Corollary 5.2. *In Theorem 5.1 substitute assumption* (34) *with*

$$u_0 \text{ and } u_1 \text{ have maxima at the origin.} \tag{46}$$

Then the same conclusion holds.

Proof. As in Corollary 4.2, we just have to observe that assumption (46), together with the convexity of the involved sets, yields assumption (34). $\square$

Proof of Theorem 1.3. Like in Corollary 4.3, observe that the symmetry of Ω_0 and Ω_1 implies (46). Then we can apply Corollary 5.2 to obtain that there exists a such that $\Omega_1 = \Omega_0 + a$. As Ω_0 and Ω_1 are origin symmetric, we have that $a = 0$. $\square$

Acknowledgements

Colesanti, Francini and Salani are partially supported by INdAM through different GNAMPA projects and by the Italian "Ministero dell'Università e Ricerca" and EU through different PRIN 2022 projects, within the Next Generation EU program. Livshyts was supported by the National Science Foundation grant NSF-BSF DMS-2247834.

References

[Acker et al. 1981] A. Acker, L. E. Payne, and G. Philippin, "On the convexity of level lines of the fundamental mode in the clamped membrane problem, and the existence of convex solutions in a related free boundary problem", *Z. Angew. Math. Phys.* **32**:6 (1981), 683–694. MR

[Alvarez et al. 1997] O. Alvarez, J.-M. Lasry, and P.-L. Lions, "Convex viscosity solutions and state constraints", *J. Math. Pures Appl.* (9) **76**:3 (1997), 265–288. MR

[Artstein-Avidan et al. 2015] S. Artstein-Avidan, A. Giannopoulos, and V. D. Milman, *Asymptotic geometric analysis, I*, Mathematical Surveys and Monographs **202**, Amer. Math. Soc., Providence, RI, 2015. MR

[Artstein-Avidan et al. 2021] S. Artstein-Avidan, A. Giannopoulos, and V. D. Milman, *Asymptotic geometric analysis, II*, Mathematical Surveys and Monographs **261**, Amer. Math. Soc., Providence, RI, 2021. MR

[Bianchini and Salani 2009] C. Bianchini and P. Salani, "Concavity properties for elliptic free boundary problems", *Nonlinear Anal.* **71**:10 (2009), 4461–4470. Correction in **72** (2010), 3351. MR

[Bianchini and Salani 2013] M. Bianchini and P. Salani, "Power concavity for solutions of nonlinear elliptic problems in convex domains", pp. 35–48 in *Geometric properties for parabolic and elliptic PDE's*, edited by R. Magnanini et al., Springer INdAM Ser. **2**, Springer, 2013. MR

[Borell 1975] C. Borell, "The Brunn–Minkowski inequality in Gauss space", *Invent. Math.* **30**:2 (1975), 207–216. MR

[Borell 2000] C. Borell, "Diffusion equations and geometric inequalities", *Potential Anal.* **12**:1 (2000), 49–71. MR

[Borell 2003] C. Borell, "The Ehrhard inequality", *C. R. Math. Acad. Sci. Paris* **337**:10 (2003), 663–666. MR

[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

[Brasco 2014] L. Brasco, "On torsional rigidity and principal frequencies: an invitation to the Kohler–Jobin rearrangement technique", *ESAIM Control Optim. Calc. Var.* **20**:2 (2014), 315–338. MR

[Caffarelli and Friedman 1985] L. A. Caffarelli and A. Friedman, "Convexity of solutions of semilinear elliptic equations", *Duke Math. J.* **52**:2 (1985), 431–456. MR

[Caffarelli and Spruck 1982] L. A. Caffarelli and J. Spruck, "Convexity properties of solutions to some classical variational problems", *Comm. Partial Differential Equations* **7**:11 (1982), 1337–1379. MR

[Carlen and Kerce 2001] E. A. Carlen and C. Kerce, "On the cases of equality in Bobkov's inequality and Gaussian rearrangement", *Calc. Var. Partial Differential Equations* **13**:1 (2001), 1–18. MR

[Colesanti 2005] A. Colesanti, "Brunn–Minkowski inequalities for variational functionals and related problems", *Adv. Math.* **194**:1 (2005), 105–140. MR

[Colesanti and Salani 2003] A. Colesanti and P. Salani, "The Brunn–Minkowski inequality for p-capacity of convex bodies", *Math. Ann.* **327**:3 (2003), 459–479. MR

[Crandall et al. 1992] M. G. Crandall, H. Ishii, and P.-L. Lions, "User's guide to viscosity solutions of second order partial differential equations", *Bull. Amer. Math. Soc.* (*N.S.*) **27**:1 (1992), 1–67. MR

[Ehrhard 1983] A. Ehrhard, "Symétrisation dans l'espace de Gauss", *Math. Scand.* **53**:2 (1983), 281–301. MR

[Ehrhard 1984] A. Ehrhard, "Inégalités isopérimétriques et intégrales de Dirichlet gaussiennes", *Ann. Sci. École Norm. Sup.* (4) **17**:2 (1984), 317–332. MR

[Eskenazis and Moschidis 2021] A. Eskenazis and G. Moschidis, "The dimensional Brunn–Minkowski inequality in Gauss space", *J. Funct. Anal.* **280**:6 (2021), art. id. 108914. MR

[Evans 2010] L. C. Evans, *Partial differential equations*, 2nd ed., Graduate Studies in Mathematics **19**, Amer. Math. Soc., Providence, RI, 2010. MR

[Gardner 2002] R. J. Gardner, "The Brunn–Minkowski inequality", *Bull. Amer. Math. Soc.* (*N.S.*) **39**:3 (2002), 355–405. MR

[Gardner and Zvavitch 2010] R. J. Gardner and A. Zvavitch, "Gaussian Brunn–Minkowski inequalities", *Trans. Amer. Math. Soc.* **362**:10 (2010), 5333–5353. 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. MR

[Guan and Xu 2013] P. Guan and L. Xu, "Convexity estimates for level sets of quasiconcave solutions to fully nonlinear elliptic equations", *J. Reine Angew. Math.* **680** (2013), 41–67. MR

[Herscovici and Livshyts 2024] O. Herscovici and G. V. Livshyts, "Kohler–Jobin meets Ehrhard: the sharp lower bound for the Gaussian principal frequency while the Gaussian torsional rigidity is fixed, via rearrangements", *Proc. Amer. Math. Soc.* **152**:10 (2024), 4437–4450. MR

[Ishii 1995] H. Ishii, "On the equivalence of two notions of weak solutions, viscosity solutions and distribution solutions", *Funkcial. Ekvac.* **38**:1 (1995), 101–120. MR

[Kennington 1985] A. U. Kennington, "Power concavity and boundary value problems", *Indiana Univ. Math. J.* **34**:3 (1985), 687–704. MR

[Kohler-Jobin 1978] M.-T. Kohler-Jobin, "Une méthode de comparaison isopérimétrique de fonctionnelles de domaines de la physique mathématique, I: Une démonstration de la conjecture isopérimétrique $P\lambda^2 \geq \pi j_0^4/2$ de Pólya et Szegő", *Z. Angew. Math. Phys.* **29**:5 (1978), 757–766. MR

[Kohler-Jobin 1982] M.-T. Kohler-Jobin, "Symmetrization with equal Dirichlet integrals", *SIAM J. Math. Anal.* **13**:1 (1982), 153–161. MR

[Kolesnikov and Livshyts 2021] A. V. Kolesnikov and G. V. Livshyts, "On the Gardner–Zvavitch conjecture: symmetry in inequalities of Brunn–Minkowski type", *Adv. Math.* **384** (2021), art. id. 107689. MR

[Korevaar 1983] N. J. Korevaar, "Convex solutions to nonlinear elliptic and parabolic boundary value problems", *Indiana Univ. Math. J.* **32**:4 (1983), 603–614. MR

[Korevaar and Lewis 1987] N. J. Korevaar and J. L. Lewis, "Convex solutions of certain elliptic equations have constant rank Hessians", *Arch. Rational Mech. Anal.* **97**:1 (1987), 19–32. MR

[Leindler 1972] L. Leindler, "On a certain converse of Hölder's inequality, II", *Acta Sci. Math.* (*Szeged*) **33**:3-4 (1972), 217–223. MR

[Livshyts 2024a] G. V. Livshyts, "Lecture notes on concentration of measure phenomena and convexity", lecture notes, Georgia Tech, 2024, available at https://glivshyts6.math.gatech.edu/Conc-lecture-notes.pdf.

[Livshyts 2024b] G. V. Livshyts, "On a conjectural symmetric version of Ehrhard's inequality", *Trans. Amer. Math. Soc.* **377**:7 (2024), 5027–5085. MR

[Ma and Ou 2010] X.-N. Ma and Q. Ou, "The convexity of level sets for solutions to partial differential equations", pp. 295–322 in *Trends in partial differential equations*, edited by B. Bian et al., Adv. Lect. Math. (ALM) **10**, International Press, Somerville, MA, 2010. MR

[Pólya and Szegö 1951] G. Pólya and G. Szegö, *Isoperimetric Inequalities in Mathematical Physics*, Annals of Mathematics Studies **27**, Princeton University Press, Princeton, NJ, 1951. MR

[Prékopa 1971] A. Prékopa, "Logarithmic concave measures with application to stochastic programming", *Acta Sci. Math.* (*Szeged*) **32** (1971), 301–316. MR

[Prékopa 1973] A. Prékopa, "On logarithmic concave measures and functions", *Acta Sci. Math.* (*Szeged*) **34** (1973), 335–343. MR

[Salani 2015] P. Salani, "Combination and mean width rearrangements of solutions to elliptic equations in convex sets", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **32**:4 (2015), 763–783. MR

[Schneider 1993] R. Schneider, *Convex bodies*: *the Brunn–Minkowski theory*, Encyclopedia of Mathematics and its Applications **44**, Cambridge University Press, Cambridge, 1993. MR

[Schneider 2014] R. Schneider, *Convex bodies*: *the Brunn–Minkowski theory*, 2nd expanded ed., Encyclopedia of Mathematics and its Applications **151**, Cambridge University Press, Cambridge, 2014. MR

[Sudakov and Tsirel'son 1974] V. N. Sudakov and B. S. Tsirel'son, "Extremal properties of half-spaces for spherically invariant measures", pp. 14–24 in *Problems of the theory of probability distributions*, *II*, edited by V. N. Sudakov, Zap. Nauchn. Sem. LOMI **41**, 1974. In Russian; translated in *J. Sov. Math.* **9**, 9–18 (1978). MR

Received 6 Aug 2024. Revised 14 Jul 2025. Accepted 13 Oct 2025.

ANDREA COLESANTI: andrea.colesanti@unifi.it
Dipartimento di Matematica e Informatica "Ulisse Dini", Università di Firenze, Firenze, Italy

ELISA FRANCINI: elisa.francini@unifi.it
Dipartimento di Matematica e Informatica "Ulisse Dini", Università di Firenze, Firenze, Italy

GALYNA LIVSHYTS: glivshyts6@math.gatech.edu
Georgia Institute of Technology, Atlanta, GA, United States

PAOLO SALANI: paolo.salani@unifi.it
Dipartimento di Matematica e Informatica "Ulisse Dini", Università di Firenze, Firenze, Italy

ANALYSIS AND PDE
Vol. 19 (2026), No. 7, pp. 1361–1394
DOI: 10.2140/apde.2026.19.1361

ALMOST-FUCHSIAN STRUCTURES ON DISC BUNDLES OVER A SURFACE

SAMUEL BRONSTEIN

Considering an integer $d > 0$, we show the existence of convex-cocompact representations of surface groups into SO(4, 1) admitting an embedded minimal map with small second fundamental form, and whose associated hyperbolic 4-manifolds are disc bundles of degree d over the surface, provided its genus is large enough. We also show that we can realize these representations as complex variation of Hodge structures. This gives examples of quasicircles in $\mathbb{S}^3$ bounding superminimal discs in $\mathbb{H}^4$ of arbitrarily small second fundamental form. Those are examples of generalized almost-fuchsian representations which are not deformations of fuchsian representations by discrete and faithful representations.

1. Introduction 1361
2. The curvature equations 1366
3. Balanced holomorphic sections of positive line bundles 1372
4. Estimates on the solutions to the flat curvature equations 1380
5. Proof of the main theorem 1390
Acknowledgment 1393
References 1393

1. Introduction

1.1. *Context.*

1.1.1. *Minimal surfaces and convex-cocompact representations.* In this paper, we consider representations of a closed surface group into $\mathrm{SO}(4,1) \approx \mathrm{Isom}(\mathbb{H}^4)$ and minimal surfaces in the hyperbolic 4-space equivariant under the group action. We prove the existence of "exotic" minimal surfaces of small curvature, that is, minimal surfaces with arbitrarily small second fundamental form while the quotient of the hyperbolic space by the group action is a nontrivial disc bundle over the surface considered.

Minimal surfaces are surfaces which are locally critical points of the area functional. Considering an immersion from a Riemann surface Σ to $\mathbb{H}^4$, the image of an immersion is a minimal surface if the immersion is conformal and harmonic. The study of minimal surfaces in hyperbolic manifolds is a rich and diverse topic. Schoen and Yau [1979] and Sacks and Uhlenbeck [1982] showed that in any isotopy class of an incompressible surface in a compact manifold, there is a weakly branched immersion whose image is minimal. See [Meeks and Yau 1982] for the case when the target manifold has mean convex boundary. In the case of a hyperbolic three-dimensional manifold with convex boundary, there is an immersion, which will be an embedding if we're in the isotopy class of an incompressible surface.

MSC2020: 51M10, 53A10, 53A35, 58J90.
Keywords: minimal surface, almost-fuchsian, hyperbolic, disc bundles.

© 2026 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.

Considering a faithful and discrete representation ρ in $\mathrm{Isom}(\mathbb{H}^n)$, it is then natural to ask whether the quotient manifold $\rho\backslash\mathbb{H}^n$ contains a nonempty compact convex set. When that is the case, we say that ρ is *convex-cocompact*. In particular, for a convex-cocompact representation ρ from a closed surface group in $\mathrm{Isom}(\mathbb{H}^n)$, there is always an equivariant weak immersion $\mathbb{D}^2 \to \mathbb{H}^4$ whose image is minimal (but has branch points in general). While branch points can be ruled out for least area minimal surfaces in hyperbolic 3-manifolds with convex boundary, it is not true from dimension 4 on. In this regard, our theorem gives examples of minimal maps into a disc bundle over a surface which are embeddings, so without branch points and self-intersections.

1.1.2. *Hyperbolic structures on disc bundles.* Among the family of hyperbolic 4-manifolds whose fundamental group is a surface group, one family is of specific interest to us: disc bundles over a Riemann surface. Considering a fuchsian representation of a surface group into $\mathrm{SO}(4, 1)$, it preserves a totally geodesic disc in $\mathbb{H}^4$ and the quotient manifold identifies with the normal bundle to the quotient of the disc. Topologically the quotient manifold is the product of a disc with the Riemann surface. However, these are not the only examples of hyperbolic disc bundles. Gromov, Lawson and Thurston [Gromov et al. 1988] were the first to give examples of hyperbolic structures on nontrivial disc bundles over a surface, considering equivariant piecewise-linear embeddings of the disc into $\mathbb{H}^4$. Kuiper [1988] also gave examples of those, using tessellations of the hyperbolic 4-space.

Topologically, disc bundles over a Riemann surface are classified by the Euler characteristic of the surface and the Euler class of the bundle. The precise list of disc bundles over a closed surface admitting a hyperbolic structure is still an open problem. Gromov, Lawson and Thurston conjectured that a necessary condition to admit a hyperbolic structure would be that the degree d is smaller in absolute value than the Euler characteristic $2g - 2$. Kapovich [1989] showed that for $0 \leq d \leq (g - 1)/11$, there is indeed a hyperbolic structure on the corresponding disc bundle.

Our result gives a new method of finding some of those hyperbolic structures, with more geometric properties, although we don't have an explicit condition on how small the degree has to be with regard to g. Check [Ho 2014] for the latest results concerning the Gromov–Lawson–Thurston conjecture.

1.1.3. *Almost-fuchsian representations.* Another important question regarding minimal surfaces is the potential uniqueness in a fixed isotopy class. Already in dimension 3, there are examples of hyperbolic 3-manifolds containing arbitrarily many isotopic minimal surfaces [Huang and Wang 2015]. However, we may get uniqueness if we add some conditions on the manifold: Uhlenbeck [1983] introduced the notion of *almost-fuchsian representations* which are faithful and discrete representations of a surface group into $\mathrm{PSL}(2, \mathbb{C})$ admitting an equivariant immersion $\mathbb{D}^2 \to \mathbb{H}^3$ minimal and with principal curvatures in $(-1, 1)$.

Note that asking the principal curvatures to be zero is equivalent to asking that our immersion is totally geodesic. Here as we relaxed this closed condition into asking the principal curvatures to be in $(-1, 1)$, the notion is open in the character variety. This differential geometric assumption has several consequences: the representation is then convex-cocompact, the image of f is only ρ-invariant minimal disc, and f is an embedding.

Since then, the notion of almost-fuchsian representation has been extended to almost-fuchsian immersions [Krasnov and Schlenker 2007]. Seppi [2016] studied almost-fuchsian discs in $\mathbb{H}^3$, Donaldson [2003], Hodge [2005] and Trautwein [2019] considered the hyperkähler structure on the moduli space of almost-fuchsian representations extending the hyperkähler structure on Teichmüller space. With Smith [Bronstein and Smith 2023], we have exhibited a parametrization of the set of almost-fuchsian discs in $\mathbb{H}^3$ by a convex set of bounded holomorphic quadratic differentials on the disc. There are important questions related to foliations by surfaces of almost-fuchsian manifolds [Guo et al. 2010; Choudhury et al. 2024].

Jiang [2025] considered almost-fuchsian discs in $\mathbb{H}^n$. For faithful and discrete representations into $\mathrm{SO}(n, 1)$, the existence of an equivariant minimally immersed disc whose scalar second fundamental forms have curvatures in $(-1, 1)$ actually implies the uniqueness of the minimal surface in its isotopy class, its embeddedness, and the convex-cocompactness of the representation. Recently, Davalo [2025] extended the notion of almost-fuchsian immersion into nearly geodesic immersion into a higher rank symmetric space of noncompact type. The only known examples of nearly geodesic immersions are deformations of totally geodesic immersions. While in $\mathrm{PSL}(2, \mathbb{C})$ those are indeed the only examples, we bring here examples of such immersions which cannot be deformed to totally geodesic ones. In particular, the limit set of our representations is a quasicircle, and there is a quasiconformal homeomorphism of $\mathbb{S}^3$ sending our limit set to a round curve, but this homeomorphism cannot be made equivariant.

1.1.4. *Superminimal maps and complex variations of Hodge structures.* The construction we will exhibit will naturally give examples of complex variations of Hodge structures in $\mathrm{SO}(4, 1)$. The notion of complex variation of Hodge structures for representations of surface groups come from the nonabelian Hodge correspondence; see for instance the survey [Wentworth 2016]. The nonabelian Hodge correspondence is a one-to-one correspondence between the moduli space of representations of surface groups with the moduli space of an object called the Higgs bundle, on which there is an action of $\mathbb{C}^*$. The fixed points of this action are called complex variations of Hodge structures. Loftin and McIntosh [2019] give an extensive description of those representations in their article about equivariant minimal surfaces in $\mathbb{H}^4$. For a representation ρ with an equivariant minimal immersion f, asking that ρ is a complex variation of Hodge structure is equivalent to asking that f is superminimal, i.e., that it has circular ellipse of curvature (see Definition 2.5).

The term superminimal was coined in [Kommerell 1897]. In particular, it implies the minimality of the image of the map. Regarding the study of superminimal surfaces in the sphere $\mathbb{S}^4$, there are some impressive results: see for instance [Bryant 1982], which exhibited a duality between superminimal surfaces in $\mathbb{S}^4$ and holomorphic legendrian curves in $\mathbb{C}P^3$. Check [Forstnerič 2021] for the analogous statement in $\mathbb{H}^4$. More precisely, Loftin and McIntosh showed that for any pair of integers d and $g \geq 2$ such that $|d| < 2g-2$, there are representations from $\pi_1 \Sigma_g$ into $\mathrm{SO}(4, 1)$ with an equivariant superminimal immersion whose normal bundle is of degree d over Σ_g. Our result states that for d very small compared to g, there are such pairs (ρ, f) where f is actually embedded and ρ is convex-cocompact. We also have, as the normal bundle to f will identify to the quotient manifold $\rho\backslash\mathbb{H}^4$, hyperbolic structures on degree d disc bundles over Σ_g.

1.1.5. *Solving "mixed" Toda systems.* Finally, the method used to build our examples is to solve a nonlinear PDE system which we call a mixed Toda system. Toda systems, and Toda systems with opposite signs, are in [Guest and Lin 2014] and designate PDE systems where the unknown is (w_i) and satisfy an equation of the type

$$\Delta w_i = e^{w_i - w_{i-1}} - e^{w_i - w_{i+1}}. \tag{1-1}$$

While our system definitely cannot be written in this form, it belongs to a class of equations arising frequently when dealing with flat curvature equations.

This kind of system appears quite easily in the theory of flat bundles. Li and Mochizuki [2025] encounter Toda systems (of the opposite type) when dealing with cyclic Higgs bundles associated to representations in rank 2 Lie groups. Malchiodi and Ruiz [2013] and Battaglia, Jevnikar, Malchiodi and Ruiz [Battaglia et al. 2015] made use of the Moser–Trudinger inequality to solve some Toda systems motivated by the nonabelian Chern–Simons theory. Here in rank 1 with the complex Hodge variation structure assumption, we get something alike, but with tweaked signs:

$$\begin{cases} \Delta u = -1 + e^{2u} + e^{-2u} e^{2v} |\alpha|^2, \\ \Delta v = \dfrac{1}{2g-2} - e^{-2u} e^{2v} |\alpha|^2. \end{cases} \tag{1-2}$$

We think the method used in this article to solve this system could be applied to a broader class of PDE systems, yet to be described.

1.2. ***Statement of the results.*** First, we state our result in dimension 4 concerning degree 1-disc bundles over Riemann surfaces.

Theorem 1.1. *Fix $r \in (0, 1)$. There is $g_0 \geq 2$ such that, for any $g \geq g_0$ and Σ a closed Riemann surface of genus g, there is a representation $\rho : \pi_1 \Sigma \to \mathrm{Isom}(\mathbb{H}^4)$ and an immersion $f : \mathbb{D}^2 \to \mathbb{H}^4$ satisfying*

(a) *ρ is faithful, discrete, convex-cocompact, and almost-fuchsian;*

(b) *f is a ρ-equivariant embedding, superminimal;*

(c) *the second fundamental form of f satisfies $|\mathbb{I}_f| < r$ everywhere;*

(d) *the quotient hyperbolic 4-manifold $\rho \backslash \mathbb{H}^4$ is a degree 1 disc bundle over Σ.*

As a corollary, we have the existence of almost-fuchsian representations uniformizing degree d disc bundles over Riemann surfaces of high genus.

Corollary 1.2. *Let $d > 0$. For any $g_0 > 0$, there is $g \geq g_0$, Σ a surface of genus g and $\rho : \pi_1 \Sigma \to \mathrm{Isom}(\mathbb{H}^4)$ and $f : \mathbb{D}^2 \to \mathbb{H}^4$ satisfying*

(a) *ρ is faithful, discrete, convex-cocompact, almost-fuchsian;*

(b) *f is ρ-equivariant embedding, superminimal;*

(c) *the second fundamental form of f satisfies $|\mathbb{I}_f| < r$ everywhere;*

(d) *the quotient manifold $\rho \backslash \mathbb{H}^4$ is a degree d disc bundle over Σ.*

1.3. *Scheme of proof.*

1.3.1. *Writing the flat curvature equations.* In order to build a hyperbolic structure on a degree d disc bundle, we first show the existence of a hyperbolic structure on a degree 1 disc bundle over a surface with the desired properties. Once this example is built, we can take a cover of degree d of this bundle in order to obtain a hyperbolic structure on a degree d disc bundle over a surface (of genus large enough).

Consider a Riemann surface Σ of genus g. Our proof will rely on solving a PDE system of the following kind:

$$\begin{cases} \Delta u = -1 + e^{2u} + e^{-2u}e^{2v}|\alpha|^2, \\ \Delta v = \dfrac{1}{2g-2} - e^{2v}e^{-2u}|\alpha|^2, \end{cases} \tag{1-3}$$

where α has to be a holomorphic section of a degree $4g-3$ line bundle over Σ. The Laplacian is always taken with respect to the Poincaré volume form on the surface, and every line bundle will be considered endowed with a metric whose curvature form is proportional to the Poincaré volume form, hence defining the norm $|\alpha|$.

The second section will be devoted to showing how solving this system yields a minimal disc in $\mathbb{H}^4$ equivariant under a representation ρ. With the good pointwise control on $e^{2v}e^{-2u}|\alpha|^2$, the representation ρ will be almost-fuchsian and the quotient manifold $\rho\backslash\mathbb{H}^4$ will be a degree 1 disc bundle over the Riemann surface.

The associated minimal map will have a holomorphic second fundamental form (Definition 2.2) $\alpha \oplus 0$, hence be superminimal.

1.3.2. *Controlling the geometric data.* The third section will be devoted to finding families (Σ_n) of surfaces with genus going to infinity, with data (α_n) associated, while controlling the behavior of solutions of PDEs on (Σ_n) depending on (α_n). To do so, we need to control the systole and spectral gap of (Σ_n), and the ratio $\left(\fint |\alpha_n|^2\right)/|\alpha_n|^2_\infty$.

Controlling the systole is the easiest part, as the systole is nondecreasing when taking the Riemannian cover of a surface. Controlling the spectral gap is trickier, but can be done by taking (Σ_n) a random cover of a genus 2 surface thanks to a result of Magee, Naud and Petri [Magee et al. 2022]: considering a cover $\Sigma' \to \Sigma$ of closed Riemann surfaces, the relative spectrum of Σ' is defined to be the part of the spectrum of Σ' which is not in the spectrum of Σ. Magee, Naud and Petri states that for a random cover Σ of large degree, the probability that the relative spectral gap is lower than $\frac{3}{16} - \varepsilon$ goes to zero when the degree goes to infinity. Using this result, we are able to find families of coverings of a Riemann surface whose spectral gap is uniformly lower bounded.

Finally, we need to build α_n such that the ratios $|\alpha_n|^2_\infty / \fint |\alpha_n|^2$ are uniformly bounded. This property will be called *balanced* family of sections. We will show that multiplying a cover of a section with a section of a line bundle of degree 1 yields a balanced family of sections of the desired type.

1.3.3. *Building the maps Φ and Ψ.* Once this is done, we need to solve our PDE system. For this, we will consider each line of the system and build two maps. Fixing v and solving the first line in u, we will

build a map $\Phi : C^{0,\alpha}(\Sigma) \to C^{2,\alpha}(\Sigma)$ which will be well-defined on the set $\mathcal{K}_v$, where

$$\mathcal{K}_v := \left\{ v \in C^{0,\alpha}(\Sigma) : e^{2v}|\alpha|^2 \leq \tfrac{1}{2} \right\}. \tag{1-4}$$

And it happens that the image of $\mathcal{K}_v$ is in the set $\mathcal{K}_u$, where

$$\mathcal{K}_u := \left\{ u \in C^{0,\alpha}(\Sigma) : -\tfrac{1}{2}\ln 2 \leq u \leq 0 \right\}. \tag{1-5}$$

We would like to do the same work for the second equation, fixing v, solving in u to get a map Ψ. However, the work to build Ψ is considerably harder. The third section will be devoted to the study of Ψ. The method used to solve it is heavily inspired from the analysis of the prescribed curvature equation in the projective plane $\mathbb{R}P^2$: for an even function on the sphere which is positive on an open set, it is the curvature of a metric conformal to the round metric on $\mathbb{S}^2$ (and even). See [Chang and Yang 2003] for a precise explanation of this. With the same tools, we are able to solve the second line and to get the following estimate: there is a constant $C > 0$ such that the solution v satisfies

$$e^{2v}e^{-2u}|\alpha|^2 \leq \frac{C}{2g-2}. \tag{1-6}$$

The constant C will depend on three parameters: the systole δ of the surface, its spectral gap Λ and the ratio between the L^∞-norm of α and its L^2-average. Hence, thanks to the third section it can be bounded on surfaces of sufficiently large genus.

1.3.4. *Piecing everything together.* Considering a family of covers (Σ_n) with a balanced family of sections (α_n) and controlled systole and spectral gap, we are able, provided n is large enough, to build a continuous map $\Psi : \mathcal{K}_u \to \mathcal{K}_v$.

With Ψ constructed and well-defined, we will show that the composed map $\Phi \circ \Psi$ defined on $\mathcal{K}_u$ has image in a compact convex subset of $\mathcal{K}_u$. Hence, by Schauder's fixed-point theorem, $C^{0,\alpha}(\Sigma)$ being a Banach space, $\Phi \circ \Psi$ has a fixed point $u \in \mathcal{K}_u$. Write $v = \Psi(u)$.

It is clear that (u, v) is a solution of the PDE system considered, and so it corresponds to an almost-fuchsian representation ρ whose quotient manifold is a degree 1 disc bundle over the Riemann surface.

2. The curvature equations

2.1. ***Minimal equivariant immersions in*** $\mathbb{H}^4$**.** Let Σ be a closed oriented surface of genus $g \geq 2$.

Consider a representation

$$\rho : \pi_1\Sigma \to \mathrm{SO}_0(4,1) \approx \mathrm{Isom}_+(\mathbb{H}^4).$$

We will always assume that ρ is faithful and discrete. Let $f : \mathbb{D}^2 = \widetilde{\Sigma} \to \mathbb{H}^4$ be an equivariant immersion with minimal image.

Endowing Σ with the induced metric by f, we get a conformal and harmonic map.

Associated to f, there is a flat $\mathbb{R}^{4,1}$ vector bundle E, which decomposes as follows:

Let $\mathbb{I}$ be the second fundamental form of f, and B its shape operator. The vector bundle (E, ∇_E) splits as $E = O_\mathbb{R} \oplus T_\mathbb{R} \oplus N_\mathbb{R}$, where $O_\mathbb{R}$ is the trivial vector bundle of rank 1, $T_\mathbb{R}$ the tangent bundle of rank 2,

and $N_{\mathbb{R}}$ is the normal bundle to f, again of rank 2. In this splitting, the connection has the expression

$$\nabla_E = \begin{pmatrix} \nabla & a^\dagger & 0 \\ a & \nabla_T & B \\ 0 & \mathbb{I} & \nabla_N \end{pmatrix}. \tag{2-1}$$

Denote by q the quadratic form in $\mathbb{R}^{4,1}$ given by

$$q(x_1, \ldots, x_5) := x_1^2 + x_2^2 + x_3^2 + x_4^2 - x_5^2. \tag{2-2}$$

As the map f is valued in the quadric $q = -1$ in $\mathbb{R}^{4,1}$, the image of f gives a reduction of E into an $S(O(1) \times O(4))$ bundle. The differential of f being valued, by definition, in the tangent bundle, we denote by $a = df$ the nonvanishing section of $T^*\Sigma \otimes \mathrm{Hom}(O_{\mathbb{R}}, T)$. The definition of $\mathbb{I}$ and of B gives the desired decomposition of ∇_E.

Remark. This splitting is orthogonal with respect to b the signature (4, 1) bilinear form, with $O_{\mathbb{R}}$ being a negative line. As the surface is oriented and our representation is into orientation-preserving isometries, $\mathrm{SO}_0(4, 1) \approx \mathrm{Isom}_+(\mathbb{H}^4)$, and $T_{\mathbb{R}}$ and $N_{\mathbb{R}}$ are endowed with an SO(2) structure corresponding to the induced metrics on the tangent and normal bundle to f. The nondegenerate metric b also implies that $\mathbb{I}$ and B are antiadjoint with respect to b:

$$b(B(w)u, v) + b(\mathbb{I}(u, v), w) = 0 \quad \text{for all } (u, v) \in T_{\mathbb{R}},\ w \in N_{\mathbb{R}}. \tag{2-3}$$

We let $B = -\mathbb{I}^\dagger$, where $\dagger$ denotes the transposition, identifying $\mathrm{Hom}(E, F)$ with $\mathrm{Hom}(F^*, E^*)$.

Complexify now the bundles $T_{\mathbb{R}}$ and $N_{\mathbb{R}}$. As each of these are equipped with an SO(2)-structure, their complexification splits into the sum of two line bundles, the restriction of B giving a pairing between those. Now for each line subbundle of E coming into play, we consider ∇_L the induced connection of ∇_E and we take its (0, 1)-part to get $\bar\partial_L$ a holomorphic structure on L. Remark that the (0, 1)-part is well-defined because the induced metric endows the surface with a Riemann surface structure.

The diagonalization of the SO(2) structure with regard to the metrics yields:

Proposition 2.1. *Viewing* $(\mathbb{I}_{\mathbb{C}})$ *as a* 1*-form valued in* $\mathrm{Hom}(K^{-1}, N)$, *the complexification of the second fundamental form can be written*

$$(T_{\mathbb{R}})_{\mathbb{C}} = \overline{K} \oplus K = K^{-1} \oplus K, \tag{2-4}$$

$$(N_{\mathbb{R}})_{\mathbb{C}} = N \oplus \overline{N} = N \oplus N^{-1}, \tag{2-5}$$

$$(\mathbb{I})_{\mathbb{C}} = \begin{pmatrix} \alpha & \beta^* \\ \beta & \alpha^* \end{pmatrix}, \tag{2-6}$$

with α *and* β $(1, 0)$*-forms.*

Proof. The splitting of the complexification of an SO(2)-bundle is given by the eigenspaces of J the fiberwise rotation of angle $\frac{\pi}{2}$. Remark that this splitting diagonalizes ∇_T and ∇_N. In this splitting, write down the second fundamental form

$$(\mathbb{I})_{\mathbb{C}} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}. \tag{2-7}$$

As $\mathbb{I}$ is real-valued, we have

$$\gamma = \bar{\beta} = \beta^*, \tag{2-8}$$

$$\delta = \bar{\alpha} = \alpha^*. \tag{2-9}$$

Finally, as f is minimal, $\mathbb{I}$ is traceless, and so the $(1,1)$-part of the second fundamental form is zero. This implies that both $\alpha \in \Gamma(K \otimes \mathrm{Hom}(K^{-1}, N))$ and $\beta \in \Gamma(K \otimes \mathrm{Hom}(K^{-1}, N^{-1}))$ are $(1,0)$-forms. □

Remark. In the process, we endowed the bundle $E \otimes \mathbb{C}$ with the following holomorphic structure:

$$E \otimes \mathbb{C} = \mathcal{O} \oplus K^{-1} \oplus K \oplus N \oplus N^{-1}, \tag{2-10}$$

$$\bar{\partial}_E = \bar{\partial}_{\mathcal{O}} \oplus \bar{\partial}_{K^{-1}} \oplus \bar{\partial}_K \oplus \bar{\partial}_N \oplus \bar{\partial}_{N^{-1}}. \tag{2-11}$$

It is clear that the sections α, β characterize wholly the second fundamental form. Hence we define the *holomorphic second fundamental form*:

Definition 2.2. Let $f : \mathbb{D}^2 \hookrightarrow \mathbb{H}^4$ be a minimal immersion. The *holomorphic second fundamental form* of f is the $(2,0)$-part of the complexification of the second fundamental form

$$\mathbb{I}^{2,0} = \alpha \oplus \beta \in \Gamma(K^2(N \oplus N^{-1})). \tag{2-12}$$

2.2. ***Writing the flat curvature equations as a scalar PDE system.*** In order to write this system as a scalar PDE system, we need to write down each metric as a conformal change of a preferred one. For the metric on K, we denote by $e^{2u}\sigma$ the induced metric, where σ is the unique complete hyperbolic metric in the conformal class.

Let ω be the Poincaré $(1,1)$-form associated to σ. Also, we denote by Δ the Laplace–Beltrami operator on (Σ, σ). Note that Δ is defined to be locally $\partial_x^2 + \partial_y^2$, so it is a nonpositive operator satisfying

$$\int_\Sigma (\Delta f) g\, \omega + \int_\Sigma \nabla f \cdot \nabla g\, \omega = 0 \quad \text{for all } (f, g) \in C(\Sigma). \tag{2-13}$$

Generally, on a holomorphic line bundle L of degree d, there is a hermitian metric h_L whose curvature form is

$$iF_{L,h_L} = \frac{d}{2g-2}\omega. \tag{2-14}$$

The metric h_L is uniquely defined up to multiplicative constant. From now on, every line bundle will be endowed with a metric of this kind, and a given fiberwise norm $|\cdot|_L$.

Let's write down the flat curvature condition as a scalar PDE system:

Theorem 2.3. *Let $f : \Sigma \hookrightarrow M$ be a conformal harmonic immersion into a complete hyperbolic 4-manifold. Consider $\alpha \oplus \beta$ its holomorphic second fundamental form. Let $e^{2u}\sigma$ be the induced metric on Σ and $e^{2v}h_N$ the induced metric on the normal bundle. Then the flat curvature equations are equivalent to the system*

$$\begin{cases} \bar{\partial}\alpha = 0, \\ \bar{\partial}\beta = 0, \\ \Delta u = e^{2u} - 1 + e^{-2u}(e^{2v}|\alpha|^2 + e^{-2v}|\beta|^2), \\ \Delta v = c - e^{-2u}(e^{2v}|\alpha|^2 - e^{-2v}|\beta|^2). \end{cases} \tag{2-15}$$

Proof. First, as $B = -\mathbb{I}^\dagger$, we have the explicit description

$$B = \begin{pmatrix} -\alpha^{\dagger,*} & -\beta^{\dagger,*} \\ -\beta^\dagger & -\alpha^\dagger \end{pmatrix}. \tag{2-16}$$

Now the flat curvature equation is $F_{\nabla_E} = 0$, so coordinatewise,

$$\begin{cases} \nabla\alpha = 0, \\ \nabla\beta = 0, \\ F_{K^{-1}} + aa^* - \alpha^{*,\dagger}\alpha - \beta^{*,\dagger}\beta = 0, \\ F_N - \alpha\alpha^{*,\dagger} - \beta^*\beta^\dagger = 0, \end{cases} \tag{2-17}$$

where we wrote $\alpha\alpha^*$ for the wedge product $\alpha \wedge \alpha^*$. Also, the notation $\nabla\alpha = 0$ means that α is parallel for the connection on $\mathrm{Hom}(K^{-1}, N)$, namely $\nabla_N \wedge \alpha + \alpha \wedge \nabla_{K^{-1}} = 0$.

The standard formula for the curvature under a conformal change of metric is

$$\begin{cases} iF_{K^{-1}} = e^{-2u}(-\Delta u - 1)\omega, \\ iF_N = e^{-2u}(-\Delta v + c)\omega. \end{cases} \tag{2-18}$$

Now recall that as α is a $(1,0)$-form, and locally $dz \wedge d\bar{z} = -2i\,dx \wedge dy$, we have

$$\begin{cases} iaa^* = \omega, \\ i\alpha\alpha^* = e^{-4u}e^{2v}|\alpha|^2\omega, \\ i\beta\beta^* = e^{-4u}e^{-2v}|\beta|^2\omega. \end{cases} \tag{2-19}$$

So the flat curvature equation is equivalent to

$$\begin{cases} e^{-2u}(\Delta u + 1) = 1 + e^{-4u}(e^{2v}|\alpha|^2 + e^{-2v}|\beta|^2), \\ e^{-2u}(\Delta v - c) = e^{-4u}(-e^{2v}|\alpha|^2 + e^{-2v}|\beta|^2). \end{cases} \tag{2-20}$$

Also, as α is of type $(1,0)$, necessarily $\partial\alpha = 0$. Hence, $\nabla\alpha = 0$ rewrites as $\bar{\partial}\alpha = 0$, and the same applies to β. Finally, rearranging the right- and left-hand sides, we get the scalar PDE system

$$\begin{cases} \bar{\partial}\alpha = 0, \\ \bar{\partial}\beta = 0, \\ \Delta u = -1 + e^{2u} + e^{-2u}(e^{2v}|\alpha|^2 + e^{-2v}|\beta|^2), \\ \Delta v = c - e^{-2u}(e^{2v}|\alpha|^2 - e^{-2v}|\beta|^2), \end{cases} \tag{2-21}$$

as desired. □

Remark. The Higgs bundle associated to f can be read with our notation. It will be

$$\mathcal{E} = \mathcal{O} \oplus K^{-1} \oplus K \oplus N \oplus N^{-1},$$

and the holomorphic structure with the Higgs field has the explicit description

$$\bar\partial_{\mathcal{E}} = \begin{pmatrix} \bar\partial & 0 & 0 & 0 & 0 \\ 0 & \bar\partial & 0 & -\alpha^* & -\beta^* \\ 0 & 0 & \bar\partial & 0 & 0 \\ 0 & 0 & \beta^* & \bar\partial & 0 \\ 0 & 0 & \alpha^* & 0 & \bar\partial \end{pmatrix}, \quad \Phi_{\mathcal{E}} = \begin{pmatrix} 0 & a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ a^* & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2-22}$$

Recall that we chose N so that $\deg N \geq 0$. Observing that $\mathcal{O} \oplus K^{-1} \oplus N$ is an invariant subbundle, the stability condition implies

$$0 \leq \deg N \leq 2g - 2. \tag{2-23}$$

Proposition 2.4. *Let f be a minimal map with $\alpha = 0$. Then* $\deg N = 0$ *and f is totally geodesic.*

Proof. As $\alpha = 0$, N is an invariant subbundle of the corresponding Higgs bundle to f. Then, by semistability, $\deg N = 0$. Now if $\beta \neq 0$ then v is solution of

$$\Delta v = e^{-2u} e^{-2v} |\beta|^2. \tag{2-24}$$

Integrating this identity implies

$$\int_\Sigma e^{-2u} e^{-2v} |\beta|^2 = 0. \tag{2-25}$$

Hence β is the zero section and f is totally geodesic, as asserted. □

We will be interested in the specific case when the second fundamental form, viewed as a map $T_{\mathbb{R}} \to (T_{\mathbb{R}})^*$, has *circular ellipse of curvature*, i.e., the image of the unit circle in $T_{\mathbb{R}}$ by the map $X \mapsto \mathbb{I}(X, X)$ is a circle. This corresponds to asking the $(2, 0)$ and $(0, 2)$ parts of $\mathbb{I}_{\mathbb{C}}$ to be orthogonal.

Definition 2.5 [Loftin and McIntosh 2019]. We say that f is "superminimal" if f is harmonic, conformal and its holomorphic second fundamental form $\alpha \oplus \beta$ satisfies

$$\alpha \cdot \beta = 0 \in H^0(K^4). \tag{2-26}$$

Remark. The notion of superminimal maps goes back to [Kommerell 1897], which studied superminimal immersions in $\mathbb{S}^4$. See [Forstnerič 2021] for a recent review of the topic. Another interpretation of the superminimality is that it is equivalent to asking that the normal Gauss map, valued in the Grassmannian of geodesic discs in $\mathbb{H}^4$, is conformal.

Depending on the degree of N, superminimal maps behave differently:

Proposition 2.6. *Let f be a superminimal map with* $\deg N = 0$, *equivariant under a faithful and discrete representation $\rho : \pi_1\Sigma \to \mathrm{SO}(4, 1)$, Σ being a closed surface. Then f is totally geodesic and ρ is a fuchsian representation.*

Proof. While this proposition can be found in [Loftin and McIntosh 2019], we reproduce it here as it can be seen as a consequence of our PDE system: As the normal bundle is trivial, we have $c = 0$ and the

superminimality implies $\alpha = 0$ or $\beta = 0$. For convenience's sake, assume $\beta = 0$. Then (u, v) must be a solution of the system

$$\begin{cases} \Delta u = -1 + e^{2u} + e^{-2u} e^{2v} |\alpha|^2, \\ \Delta v = -e^{2u} e^{2v} |\alpha|^2. \end{cases} \tag{2-27}$$

Fix u, and observe that, on a closed Riemann surface, we must have $\int \Delta v = 0$. Hence, $\alpha = 0$, the immersion f is totally geodesic and ρ is fuchsian. □

2.3. ***Almost-fuchsian discs in*** $\mathbb{H}^4$**.** First introduced in [Uhlenbeck 1983] for immersions into $\mathbb{H}^3$, the notion of almost-fuchsian immersion naturally generalizes to immersions into $\mathbb{H}^4$.

Definition 2.7. Let $f : \mathbb{D} \hookrightarrow \mathbb{H}^4$ be a proper immersion. Then f is said to be *almost-fuchsian* if it is minimal and there is $\delta > 0$ such that the second fundamental form satisfies

$$|\mathbb{I}(u, u)| \leq (1 - \delta)|u|^2 \quad \text{for all } u \in T_f. \tag{2-28}$$

Proposition 2.8. *Let $f : \mathbb{D} \to \mathbb{H}^4$ be a proper almost-fuchsian immersion. Then the exponential map $N_f \to \mathbb{H}^4$ is a diffeomorphism.*

As this constitutes, with the following result, [Jiang 2025, Lemma 4], we won't reproduce a proof here. Note that Jiang's convention for the second fundamental form differs from ours by a factor one half, hence the difference in the condition. More specifically, almost-fuchsian discs in $\mathbb{H}^4$ are embeddings, bound by a quasicircle at infinity, and if they are ρ-equivariant for some $\rho : \Gamma \to \mathrm{SO}(4, 1)$ discrete and faithful, ρ is convex-cocompact.

Proposition 2.9. *Let Σ be a closed hyperbolic surface, and $\rho : \pi_1 \Sigma \to \mathrm{SO}(4, 1)$ a faithful and discrete representation such that there exists $f : \mathbb{D} \hookrightarrow \mathbb{H}^4$ a ρ-equivariant proper almost-fuchsian immersion. Then ρ is convex-cocompact, f is an embedding and the hyperbolic 4-manifold $M = \rho \backslash \mathbb{H}^4$ is diffeomorphic via $\exp_f$ to a disc bundle over Σ. Also, f is the unique ρ-equivariant minimal immersion.*

Definition 2.10. A representation $\rho : \Gamma \to \mathrm{SO}(4, 1)$ is said to be *almost-fuchsian* if it is discrete, faithful and it admits a proper almost-fuchsian ρ-equivariant immersion $f : \widetilde{\Sigma} \to \mathbb{H}^4$.

It is not known which disc bundles can be uniformized by almost-fuchsian representations. Fuchsian representations uniformize the trivial bundle, and we will show in this paper that for a surface of large enough genus, there are almost-fuchsian representations uniformizing vector bundles with positive degree over the surface.

Proposition 2.11. *Let Σ be a closed hyperbolic surface of genus g. Denote its volume form by w. Consider N a line bundle of degree $d \geq 0$, endowed with a metric h of curvature form*

$$c\omega = \frac{d}{2g-2}\,\omega.$$

Let $\alpha \in H^0(K^2N)$ be such that there exist smooth functions u, v on Σ satisfying

$$\Delta u = -1 + e^{2u} + e^{-2u}e^{2v}|\alpha|^2, \tag{2-29}$$

$$\Delta v = c - e^{-2u}e^{2v}|\alpha|^2, \tag{2-30}$$

$$\sup e^{-4u}e^{2v}|\alpha|^2 < 1. \tag{2-31}$$

Then there is a convex-cocompact representation $\rho : \pi_1\Sigma \to \mathrm{SO}(4,1)$ and a minimal, superminimal, ρ-equivariant and almost-fuchsian immersion $f : \mathbb{D} \to \mathbb{H}^4$ such that

(1) *the induced metric by f is $e^{2u}w$,*

(2) *the hyperbolic manifold $\rho\backslash\mathbb{H}^4$ identifies with the normal bundle to f and has degree d,*

(3) *the metric induced on the normal bundle is $e^{2v}h$,*

(4) *the holomorphic second fundamental form of f is α.*

Proof. With the given data, consider the bundle $\mathcal{E} = \mathcal{O} \oplus K^{-1} \oplus K \oplus N \oplus N^{-1}$, with the metric achieved by product of the metrics on each line. Adding α and its h-dual α^*, the two equations are exactly the flat curvature equation, corresponding to a representation ρ and a map $f : \mathbb{D} \to \mathbb{H}^4$ whose holomorphic second fundamental form satisfies

$$(\mathbb{I}_f)_{\mathbb{C}} = \alpha \oplus 0 \in \Gamma(K^2(N \oplus N^{-1})). \tag{2-32}$$

It is clear that such an f is minimal and superminimal. The last hypothesis

$$\sup e^{-4u}e^{2v}|\alpha|^2 < 1 \tag{2-33}$$

ensures that f and ρ are almost-fuchsian, as asserted. □

3. Balanced holomorphic sections of positive line bundles

In this section, we show that there are families of holomorphic sections of line bundles of degree $4g-3$ which behave roughly like covers in the following sense:

Definition 3.1 (balanced family). Let (Σ_g) be a family of genus g Riemann surfaces, equipped with their hyperbolic metric. Let N_g be a family of line bundles of degree $n_g \geq 0$. Let (α_g) be a family of holomorphic sections of N_g.

On each Σ_g, denote by ω_g the Poincaré volume form. On each N_g, denote by h_g the metric on N_g with curvature form

$$\frac{n_g}{2g-2}\omega_g.$$

We say that (α_g, N_g) is *balanced* if there is $C > 0$ such that

$$\sup_g \frac{\sup_{x \in X_g} |\alpha_g|^2_{\omega_g, h_g}}{\fint_{X_g} |\alpha_g|^2_{h_g} d\omega_g} \leq C. \tag{3-1}$$

The main idea behind this control is that when solving elliptic PDEs depending on α_g, as α_g is coarsely periodic, the solutions should be too, and we will show that the solutions satisfy estimates which are independent of the genus g.

We will make extensive use of the following notation:

Notation 3.2. In this section, Σ denotes a closed hyperbolic surface of genus g. We denote by δ and Λ the systole and spectral gap of Σ, respectively. Any N positive line bundle over Σ is endowed with a metric of curvature form proportional to the Poincaré volume form ω, i.e., whose curvature form satisfies

$$iF(N, h_N) = \frac{\deg(N)}{2g-2}\omega. \tag{3-2}$$

Remark. It is easy to exhibit examples of balanced sections for n_g proportional to $g-1$. Indeed, if $n_g = a(g-1)$, consider N_2 a degree a line bundle over Σ_2, and α a holomorphic section of N_2. and consider covers $\Sigma_g \to \Sigma_2$. Then the family of lifts of N_2 and α yields a balanced family.

Here we will prove the existence of balanced families for line bundles of degree $4g-3$:

Theorem 3.3. *Let Σ be a closed hyperbolic surface of genus $g \geq 2$. Consider (Σ_n) a family of covers of Σ of degree n. Then there is a balanced family (α_n, N_n), with N_n a line bundle of degree $(4g-4)n+1$ over Σ_n.*

The main idea will be to consider covers to get a balanced section of a degree $4n(g-1)$ line bundle, and then to tensorize with a well-chosen degree 1 line bundle to get the desired family of sections. This requires a description of the oscillatory behavior of a degree 1 bundle section, which we provide here.

Before all, we will make use of the continuous embedding $W^{2,2} \hookrightarrow L^\infty$.

Lemma 3.4. *Let Σ be a hyperbolic surface with systole δ. Let M be the constant of continuity of the embedding $W^{2,2}(\mathbb{H}^2) \hookrightarrow L^\infty(\mathbb{H}^2)$. Then, for any $u \in W^{2,2}(\Sigma)$,*

$$|u|_\infty \leq \frac{M}{\delta}|u|_{2,2}. \tag{3-3}$$

Proof. As smooth maps are dense in $W^{2,2}(\Sigma)$, consider $u \in W^{2,2}(\Sigma)$ smooth. Let $z_0 \in \Sigma$. By definition of the systole, there is an isometric embedding from a hyperbolic disc of radius δ around z_0. Fix χ a smooth radial map on $D(z_0, \delta)$, $1/\delta$ Lipschitz, between 0 and 1, such that $\chi(z_0) = 1$ and $\chi(\partial D(z_0, \delta)) = 0$. Choose for χ a rescaling of a cutoff map on $D(0,1) \subset \mathbb{H}^2$, so that

$$\sup D\chi = \sup D^2\chi = \frac{1}{\delta}. \tag{3-4}$$

Then we can extend by zero to consider $u\chi \in W^{2,2}(\mathbb{H}^2)$. Using our cutoff function, we get

$$|u(z_0)| = |u\chi(z_0)| \leq |u\chi|_\infty \leq M|u\chi|_{2,2} \leq \frac{M}{\delta}|u|_{2,2}, \tag{3-5}$$

as asserted. □

Proposition 3.5. *Let $z_0 \in \Sigma$. Denote by $D = D(z_0, \delta)$ the disc of radius δ around z_0. Denote by a the constant $a = 2\pi/\mathrm{Vol}(D)$. Consider $k : \Sigma \to \mathbb{R}$ the C^1 function satisfying, in the weak sense,*

$$\begin{cases} \displaystyle\int_\Sigma k = 0, \\ \Delta k = \dfrac{1}{2g-2} - a\mathbf{1}_D. \end{cases} \tag{3-6}$$

Then there is a bound $C(\delta, \Lambda)$ such that

$$|\sup k - \inf k| \le C(\delta, \Lambda). \tag{3-7}$$

Proof. As $a = 2\pi/\mathrm{Vol}\, D$,

$$\int_\Sigma \frac{1}{2g-2} - a\mathbf{1}_D = 0. \tag{3-8}$$

It is clear that $1/(2g-2) - a\mathbf{1}_D \in L^2(\Sigma)$. Hence there exists a unique $k \in W^{2,2}(\Sigma)$ satisfying the stated conditions. From the Sobolev embedding $W^{2,2} \hookrightarrow C^{0,\alpha}(\Sigma)$, we get the k is Hölder-continuous, and hence in particular it belongs to $L^p(\Sigma)$ for any $p > 0$. From the L^p-interior estimates (see [Gilbarg and Trudinger 1977, Theorem 9.11]) it follows that k belongs to $W^{2,p}(\Sigma)$ for any $p > 1$. From the Sobolev embedding $W^{2,p} \hookrightarrow C^{1,\alpha}$, provided $\alpha < (p-2)/p$, we deduce that k is of regularity $C^{1,\alpha}$ for any $\alpha < 1$.

We now compute an estimate of the sup norm of k. By construction $k \in W^{2,2}(\Sigma)$, so it admits a Hessian in the distributional sense. First, we have $\int k\Delta k = -|\nabla k|_2^2$, so

$$|\nabla k|_2^2 = a\int_D k \le 2\pi \sup g. \tag{3-9}$$

Recall Bochner's identity

$$\int |\nabla^2 k|^2 = \int (\Delta k)^2 + \int |\nabla k|^2, \tag{3-10}$$

which translates to give the estimate

$$|\nabla^2 k|_2^2 \le \frac{4\pi^2}{\mathrm{Vol}(D)} - \frac{2\pi}{2g-2} + 2\pi \sup k. \tag{3-11}$$

As $\int k = 0$, by definition of the spectral gap we have

$$|k|_2^2 \le \Lambda |\nabla k|_2^2 \le 2\pi\Lambda \sup k. \tag{3-12}$$

Combining all this, we bound the $W^{2,2}$-norm

$$|k|_{2,2}^2 = |\nabla^2 k|_2^2 + |\nabla k|_2^2 + |k|_2^2 \le \frac{4\pi^2}{\mathrm{Vol}(D)} - \frac{2\pi}{2g-2} + 2\pi(2+\Lambda)\sup k. \tag{3-13}$$

By Lemma 3.4, this implies

$$|\sup k - \inf k| \le \frac{2M}{\delta}\left(\frac{4\pi^2}{\mathrm{Vol}(D)} - \frac{2\pi}{2g-2} + 2\pi(2+\Lambda)\sup k\right)^{1/2}. \tag{3-14}$$

Because $\int k = 0$, we know that $\inf k \leq 0$ and $\sup k \geq 0$. Thus,

$$0 \leq (\sup k)^2 \leq \frac{4M^2}{\delta^2}\left(\frac{4\pi^2}{\mathrm{Vol}(D)} - \frac{\pi}{g-1} + 2\pi(2+\Lambda)\sup k\right). \tag{3-15}$$

Going to the squares, this is a control of the type $X^2 \leq AX + B$, with $A, B > 0$. Explicitly, A and B are given by

$$\begin{cases} A = \dfrac{8\pi(2+\Lambda)M^2}{\delta^2}, \\ B = \dfrac{4M^2}{\delta^1}\left(\dfrac{4\pi^2}{\mathrm{Vol}(D)} - \dfrac{\pi}{g-1}\right), \end{cases} \tag{3-16}$$

so it implies the existence of $C(\delta, \Lambda) = \frac{1}{2}\left(A + \sqrt{A+4B}\right)$ such that

$$\sup k \leq C(\delta, \Lambda)s. \tag{3-17}$$

Plugging this back into the inequality (3-14), we get the existence of $\widetilde{C}(\delta, \Lambda)$ satisfying

$$|\sup k - \inf k| \leq \widetilde{C}(\delta, \Lambda), \tag{3-18}$$

as asserted. □

It happens that the oscillations of k away from $D(z_0, r)$ are actually the same as those of the log of the section s associated to the line bundle of divisor $\{z_0\}$.

Proposition 3.6. *Let N be a degree 1 line bundle over Σ and $s \in H^0(N)$ a section with a unique zero z_0. Denote by $D = D(z_0, \delta)$ the disc of radius δ around z_0. Set $a = 2\pi/\mathrm{Vol}(D)$. Consider $k : \Sigma \to \mathbb{R}$ the function associated to z_0 defined in Proposition 3.5. Denote by f the function $f = -\frac{1}{2}\log|s|^2$. Then $f - k$ is bounded on Σ and constant on $\Sigma - D(z_0, \delta)$.*

Remark. As s is well-defined up to a multiplicative constant, we can choose the value of $f - k$ on $\Sigma - D(z_0, \delta)$ at will.

Proof. Let r be the radial geodesic coordinate on $D(z_0, \delta)$. With the notation above, define the map $h : D \to \mathbb{R}$ by

$$h(r) = -2a \ln\cosh\left(\frac{r}{2}\right) + b + \ln\left(\tanh\frac{r}{2}\right) \tag{3-19}$$

for some b to be specified later in the proof. As h is radial, we can explicitly compute the Laplacian of h:

$$\Delta h = \partial_r^2 h + \frac{1}{\tanh r}\partial_r h = -a. \tag{3-20}$$

We also have

$$|z - z_0| = \tanh\left(\frac{r}{2}\right). \tag{3-21}$$

Because z_0 is a simple zero of s, locally

$$|s(z)|^2 \approx K|z - z_0|^2, \quad f(z) = -\ln\tanh\frac{r}{2} + \mathcal{O}(1). \tag{3-22}$$

And so $f + h$ is bounded.

Also, we compute the radial derivative at the boundary of the disc of h:

$$\partial_r h(\delta) = \Big(\frac{1}{2} - a\Big) \tanh \frac{\delta}{2} + \frac{1}{2\tanh(\delta/2)}. \tag{3-23}$$

But by definition of a,

$$a = \frac{2\pi}{\mathrm{Vol}(D)} = \frac{2\pi}{4\pi \sinh^2(r/2)} = \frac{1}{2\sinh^2(r/2)}, \tag{3-24}$$

which implies

$$\partial_r h(\delta) = 0. \tag{3-25}$$

Hence, with the appropriate choice of b, we can extend h by zero to get a C^1 function on Σ such that $f + h$ is bounded and satisfies weakly

$$\Delta(f + h) = \Delta f - a\mathbf{1}_D. \tag{3-26}$$

It remains to make Δf explicit. Since s is a holomorphic section, f can be expressed from the curvature form of the metric h_N on N. Denoting $H = h(s, s)$ and using local holomorphic coordinates on S, one has the formula (cf. [Huybrechts 2005, Chapter 4, Proposition 4.2.14 and Example 4.2.15]) for the Chern connection which is locally $d + \bar\partial \ln H$. From [Huybrechts 2005, Lemma 4.3.4], we then compute the curvature of that connection:

$$F(h) = d\partial \ln(H) + \partial \ln(H) \wedge \partial \ln(H) = \tfrac{1}{4}\Delta^e \ln|s|^2 \, d\bar z \wedge dz, \tag{3-27}$$

where Δ^e is the Laplacian with regard to the metric $dz\,d\bar z$. Note that the second term is zero since it is a $(2, 0)$-form on a Riemann surface. Keeping in mind that $d\bar z \wedge dz = 2i\, dx \wedge dy$, we get the formula

$$F(h_N) = \tfrac{1}{2}\Delta^e \ln|s|^2 \cdot i\, dx \wedge dy. \tag{3-28}$$

Finally, with regard to the hyperbolic metric on the Riemann surface which is locally $e^\varphi\, dz\, d\bar z$, the hyperbolic Laplacian is $\Delta = e^{-\varphi}\Delta^e$ while the volume form is $\omega = e^\varphi\, dx \wedge dy$. All in all, this shows that

$$F(h_N) = \tfrac{1}{2}\Delta \ln|s|^2 \cdot i\omega. \tag{3-29}$$

But we chose the Hermitian metric h_N so that its curvature form satisfies

$$iF(h_N) = \frac{\deg(N)}{2g-2}\,\omega. \tag{3-30}$$

Since $f = -\frac{1}{2}\ln|s|^2$, we deduce that $\Delta f = 1/(2g-2)$, and $f + h$ satisfies

$$\Delta(f + h) = \frac{1}{2g-2} - a\mathbf{1}_D. \tag{3-31}$$

Satisfied weakly, it is in particular true in $L^2(\Sigma)$. This implies that $f + h = k + A$ for some constant A. As h is constant out of D, this means that $f - k$ is constant on $\Sigma - D$, as asserted. □

Corollary 3.7. *Let Σ be a hyperbolic surface with systole δ and spectral gap Λ. Let N be a degree 1 line bundle over Σ and $\alpha \in H^0(N)$ section, with a unique zero z_0. Let $D = D(z_0, \delta)$ be the disc of radius δ around z_0. Then there is some $C(\delta, \Lambda)$ and $\lambda > 0$ such that*

$$\frac{1}{C(\delta, \Lambda)} \leq \lambda|\alpha|^2 \leq C(\delta, \Lambda) \quad \textit{for all } z \in \Sigma - D(z_0, \delta). \tag{3-32}$$

Proof. With the g defined in Proposition 3.5, we know

$$|\sup g - \inf g| \leq C(\delta, \Lambda). \tag{3-33}$$

Let $f = -\frac{1}{2}\ln|\alpha|^2$. By Proposition 3.6, $f - g$ is constant on $\Sigma - D$. Thus

$$\big|\sup_{\Sigma - D} f - \inf_{\Sigma - D} f\big| = \big|\sup_{\Sigma - D} g - \inf_{\Sigma - D} g\big| \leq C(\delta, \Lambda), \tag{3-34}$$

which directly gives, on $\Sigma - D$,

$$\frac{\sup|\alpha|^2}{\inf|\alpha|^2} \leq e^{2C(\delta,\Lambda)}. \tag{3-35}$$

Multiply by λ such that $(\sup\lambda|\alpha|^2)(\inf\lambda|\alpha|^2) = 1$ to get

$$e^{-C(\delta,\Lambda)} \leq \lambda|\alpha|^2(z) \leq e^{C(\delta,\Lambda)} \quad \text{for all } z \in \Sigma - D, \tag{3-36}$$

as desired. □

It remains to check that we can also have an upper bound on D. We can do this by some version of the Schwarz lemma adapted to sections of positive line bundles:

Proposition 3.8. *Let Σ be a hyperbolic surface with systole δ and spectral gap Λ. Let N be a degree 1 line bundle over Σ and $\alpha \in H^0(N)$ a section with a unique zero z_0. Let $D = D(z_0, \delta)$, the disc of radius δ around z_0. Then there is $C(\delta) > 0$ satisfying*

$$|\alpha|^2(z) \leq C(\delta)|z|^2 \sup_{\partial D}|\alpha|^2 \quad \textit{for all } z \in D. \tag{3-37}$$

Proof. Let $f = -\frac{1}{2}\ln|\alpha|^2$. Then f satisfies

$$\Delta f = \frac{1}{2g-2}. \tag{3-38}$$

Let r be the radial geodesic coordinate on D, and consider the map

$$h = f + \ln\tanh\frac{r}{2} - \frac{2}{2g-2}\ln\cosh\frac{r}{2}. \tag{3-39}$$

Then h is bounded on D and satisfies

$$\Delta h = 0. \tag{3-40}$$

By the minimum principle,

$$\inf_D h = \inf_{\partial D} h = \inf_{\partial D} f + \ln\tanh\frac{\delta}{2} - \frac{1}{g-1}\ln\cosh\frac{\delta}{2}. \tag{3-41}$$

Exponentiating, we get the upper bound on $|\alpha|$:

$$\frac{|\alpha|^2}{|z|^2}\cosh\Big(\frac{r}{2}\Big)^{2/(g-1)} \leq \sup_{\partial D}|\alpha|^2\frac{(\cosh(\delta/2))^{2/(g-1)}}{(\tanh(\delta/2))^2} \quad \text{for all } z\in D. \tag{3-42}$$

Because $g>1$ and $\cosh(t)\geq 1$, this implies

$$|\alpha|^2(z)\leq |z|^2\sup_{\partial D}|\alpha|^2\frac{(\cosh(\delta/2))^2}{(\tanh(\delta/2))^2}, \tag{3-43}$$

as required. □

Finally, we combine these results to show our estimate of oscillations of sections of degree 1 line bundles:

Proposition 3.9. *Let Σ be a hyperbolic surface. Denote by δ its systole, and by Λ its spectral gap. Consider $r<\delta/2$. There are constants $C_i(\delta,\Lambda,r)$ such that, for any holomorphic section s of a degree* 1 *line bundle with a unique zero $z_0\in\Sigma$, there is a $\lambda>0$ such that*

$$|\lambda s|^2\leq C_2 \qquad \text{on } B(z_0,r), \tag{3-44}$$

$$\frac{1}{C_1}\leq|\lambda s|^2\leq C_1 \quad \text{on } \Sigma - B(z_0,r). \tag{3-45}$$

Remark. Up to renormalizing our section, one can always assume the first condition to be satisfied on the ball of radius r, and then the result states that the oscillations away from the zero of the section are controlled, essentially by the spectral gap and the systole of the surface.

Proof. Let Σ be a hyperbolic surface with systole δ and spectral gap Λ. Let N be a degree 1 line bundle over Σ, with $\alpha\in H^0(N)$ a holomorphic section having one unique zero z_0. Consider $D=D(z_0,\delta)$. Thanks to Corollary 3.7, up to renormalizing α, there is $C(\delta,\Lambda)>0$ such that

$$\frac{1}{C(\delta,\Lambda)}\leq|\alpha|^2(z)\leq C(\delta,\Lambda) \quad \text{for all } z\in\Sigma-D. \tag{3-46}$$

Now because of Proposition 3.8, there is $C(\delta)$ such that

$$|\alpha|^2\leq C(\delta)C(\delta,\Lambda) \quad \text{for all } z\in D, \tag{3-47}$$

as asserted. □

Multiplying those sections with covers of a section of a degree $4(g-1)$ line bundle will result in a balanced section of a bundle of the desired degree $4n(g-1)+1$.

Proof of Theorem 3.3. Let Σ be a closed hyperbolic surface of genus $g\geq 2$. Consider Σ_n a family of degree n covers of Σ such that the spectral gap of Σ_n remains lower bounded by $\Lambda>0$. This kind of cover exists; for instance we can use the result of [Magee et al. 2022] which states that for a random cover of Σ, the probability that the relative spectral gap (i.e., the part of the spectral gap which is not in Σ) of the cover is bigger than $\frac{3}{16}-\varepsilon$ goes to 1 when the degree is large. Hence, for any $\eta<\min\{\Lambda,\frac{3}{16}\}$, we can choose a family of covers whose spectral gap is larger than η.

Of course, the systole is nondecreasing by taking a finite cover, and hence we can consider δ the systole of Σ.

Consider s a holomorphic section of K^2 over Σ which is of degree $4g-4$. Denote by s_n the lift of s to Σ_n, which is a section of a degree $4n(g-1)$ line bundle. Obviously, we have

$$\int |s_n|^2 = n \int |s|^2, \quad |s_n|_\infty = |s|_\infty. \tag{3-48}$$

Also, consider (z_n) a sequence of points $z_n \in \Sigma_n$. Define $L_n = \mathcal{O}(z_n)$ the degree 1 line bundle associated to the divisor $\{z_n\}$.

Thanks to Proposition 3.9, there are sections τ_n of L_n, and constants $C_i(\delta, \Lambda)$ such that

$$\begin{cases} \dfrac{1}{C_1(\delta,\Lambda)} \le |\tau_n|^2 \le C_1(\delta,\Lambda) & \text{out of } D(z_n,\delta), \\ |\tau_n|^2 \le C_2(\delta,\Lambda) & \text{on } D(z_n,\delta). \end{cases} \tag{3-49}$$

Consider now $N_n = K^2 L_n$ and $\alpha_n = s_n \tau_n$ a section of N_n line bundle of degree $4n(g-1)+1$. By construction, we have the estimates

$$\frac{1}{C_1(\delta,\Lambda)} |s_n|_\infty^2 \le |\alpha_n|_\infty^2 \le C_2(\delta,\Lambda) |s_n|_\infty^2, \tag{3-50}$$

$$\frac{1}{C_1(\delta,\Lambda)} \int_{\Sigma_n - D(z_n,\delta)} |s_n|^2 \le \int |\alpha_n|^2. \tag{3-51}$$

Considering that

$$|s_n|_\infty = |s|_\infty, \tag{3-52}$$

$$\int_{\Sigma_n - D(z_n,\delta)} |s_n|^2 \ge (n-1) \int_\Sigma |s|^2, \tag{3-53}$$

we eventually obtain

$$\frac{|\alpha_n|_\infty^2}{\fint |\alpha_n|^2} \le \frac{n}{n-1} \frac{|s|_\infty^2}{\int |s|^2} C_1(\delta,\Lambda) C_2(\delta,\Lambda) \operatorname{Vol}(\Sigma), \tag{3-54}$$

which is bounded in n, so the family (N_n, α_n) is balanced. □

Remark. In a broader setting, given a section s of a degree d line bundle, we expect the ratio between the L^2 average and the L^∞ norm to be a quantitative indicator of how much s looks like a lift of a section on a smaller surface. With this in mind, we expect this ratio to be controlled by the lowest distance between the zeroes of s. This would have the advantage of being an open property, and that when it is actually possible, the minimal ratio is achieved by a lift.

Remark. The same proof can be adapted to show that for any $a, d \ge 0$, we can build a balanced family $(\Sigma_g, N_g, \alpha_g)$ with N_g of degree $a(g-1)+d$.

4. Estimates on the solutions to the flat curvature equations

Throughout this section, Σ will be a closed hyperbolic surface of genus $g \geq 2$. Unless otherwise stated, N will be a degree 1 line bundle over Σ and α will be a holomorphic 2-form valued in N, i.e., $\alpha \in H^0(K^2N)$. Fix $c = 1/(2g-2)$, and consider the PDE system

$$\begin{cases} \Delta u = e^{2u} - 1 + e^{-2u}e^{2v}|\alpha|^2, \\ \Delta v = c - e^{-2u}e^{2v}|\alpha|^2. \end{cases} \tag{4-1}$$

In this part, we consider each equation separately, considering the first as giving u depending on v, and vice versa for the second.

4.1. ***Dealing with Gauss's equation.*** Fixing a map u, the first equation is the famous Gauss equation, and is well-behaved. As the data v and α are fixed, we set $f = e^{2v}|\alpha|^2$, and our problem boils down to solving

$$\Delta u = -1 + e^{2u} + e^{-2u} f. \tag{4-2}$$

Denote by $C^{0,\alpha}(\Sigma)$ the set of continuous α-Hölder functions on Σ. Looking at the results of [Bronstein and Smith 2023], we can apply the same sub-supersolution method to get:

Theorem 4.1. *Let $\eta \in (0, 1]$. Consider the set*

$$\mathcal{C} := \left\{ f \in C^{0,\alpha}(\Sigma) : 0 \leq f \leq \frac{\eta}{(1+\eta)^2} \right\}. \tag{4-3}$$

Then for any $f \in \mathcal{C}$, there is a unique $u \in C^{2,\alpha}(\Sigma)$ satisfying

$$-\tfrac{1}{2}\ln 2 \leq u \leq 0, \tag{4-4}$$

$$\Delta u = e^{2u} - 1 + e^{-2u} f. \tag{4-5}$$

Moreover, u satisfies the upper bound

$$e^{-2u} f \leq \eta. \tag{4-6}$$

Proof. We proceed with the sub-supersolution method; see, for instance, [Evans 2010, Section 9.3]. A subsolution to the equation $\Delta u = F(u)$ is a function v_- satisfying $\Delta v_- \geq F(v_-)$. Analogously, a supersolution is a function v_+ satisfying $\Delta v_+ \leq F(v_+)$.

The sub-supersolution principle states that the existence of a subsolution v_- and a supersolution v_+ satisfying $v_- \leq v_+$ ensures the existence of a solution v satisfying $v_- \leq v \leq v_+$.

Freely, 0 is a supersolution to the equation (4-5). As $f \in \mathcal{C}$, the constant map $-\frac{1}{2}\ln(1+\eta)$ is a subsolution. Hence, by the sub-supersolution principle, there is a solution u satisfying

$$-\tfrac{1}{2}\ln 2 \leq u \leq 0. \tag{4-7}$$

To prove uniqueness, assume u_1 and u_2 are both solutions in $\left[-\frac{1}{2}\ln 2, 0\right]$ with

$$\Delta u_i = e^{2u_i} - 1 + e^{-2u_i} f. \tag{4-8}$$

Then the difference $w = u_1 - u_2$ satisfies

$$\Delta w = (e^{2u_2} - e^{-2u_2} f)(e^{2w} - 1) + f e^{-2u_2}(e^{2w} + e^{-2w} - 2). \tag{4-9}$$

We observe that the second term is nonnegative, and thus

$$\Delta w \geq (e^{2u_2} - e^{-2u_2} f)(e^{2w} - 1). \tag{4-10}$$

But as $f \in \mathcal{C}$ and $u_2 \geq -\frac{1}{2}\ln 2$, we get

$$e^{-4u_2} f \leq 1, \tag{4-11}$$

which implies

$$e^{2u_2} - e^{-2u_2} f \geq 0. \tag{4-12}$$

Hence, by the maximum principle, at a maximum of w,

$$(e^{2u_2} - e^{-2u_2} f)(e^{2w} - 1) \leq 0. \tag{4-13}$$

Hence

$$w \leq 0. \tag{4-14}$$

Exchanging the roles of u_1 and u_2 shows

$$u_1 = u_2. \tag{4-15}$$

So there is a unique solution in $\left[-\frac{1}{2}\ln 2, 0\right]$.

Finally, because $-\frac{1}{2}\ln(1+\eta) \leq u \leq 0$, we deduce

$$e^{-2u} f \leq \frac{\eta}{1+\eta} \leq \eta, \tag{4-16}$$

as asserted. □

Proposition 4.2. *The map $\Phi : C^{0,\alpha}(\Sigma) \to C^{2,\alpha}(\Sigma)$ which to f associates the unique solution u of*

$$\Delta u = e^{2u} - 1 + e^{-2u} f \tag{4-17}$$

satisfying $-\frac{1}{2}\ln 2 \leq u \leq 0$ is continuous on the set $\left\{f \in C^{0,\alpha}(\Sigma) : 0 \leq f < \frac{1}{2}\right\}$.

Proof. Thanks to Theorem 4.1, Φ is well-defined. For f given, consider the linearized equation

$$\Delta \dot{u} = 2(e^{2u} - e^{-2u}) f \dot{u}. \tag{4-18}$$

The assumptions on u imply

$$e^{2u} - e^{-2u} f > 0. \tag{4-19}$$

Denote by L the operator $Lu = -\Delta u + 2(e^{2u} - e^{-2u} f)u$, which is a bounded linear operator $C^{2,\alpha} \to C^{0,\alpha}$. We claim that L is an isomorphism. First, L is surjective. This is a consequence of the positivity of $e^{2u} - e^{-2u}$ and the ellipticity of L; see [Gilbarg and Trudinger 1977, Section 6.3]. Also, L is injective: if $Lv = 0$, then at a maximum of v, we have

$$2(e^{2u} - e^{-2u})v = \Delta v \leq 0. \tag{4-20}$$

Hence we deduce that $v \leq 0$. But since $L(-v) = 0$ too, we get $v = 0$ and L is an isomorphism. By a local inversion principle, we get that the map Φ is continuous. □

4.2. ***Dealing with Ricci's equation.*** Considering the second equation, we will fix v and express a solution v as the maximizer of the following functional:

Notation 4.3. For $u \in C^{1,\alpha}(\Sigma)$, consider $W_0^{1,2}(\Sigma)$ the Sobolev space of zero average functions that are square integrable with square integrable gradient. Define the functional $J_u : W_0^{1,2}(\Sigma) \to \mathbb{R}$ as

$$J_u(w) := \log\left(⨍ e^{-2u}|\alpha|^2 e^{2w}\right) - \frac{1}{c \operatorname{Vol} \Sigma} \int |\nabla w|^2 \tag{4-21}$$

with $c = 1/(2g-2)$. The notation $⨍$ is used for the average of the considered function.

We write $|w|_2^2$ for the squared L^2-norm:

$$|w|_2^2 = \int_\Sigma |w|^2. \tag{4-22}$$

When no confusion is possible, we write J for J_u. We prove that J attains its maximum at a function w which is, up to an additive constant, the solution v that we are looking for. This proof is heavily inspired from the method used to solve the prescribed curvature equation on the sphere when the prescribed data is even. Chang and Yang [2003] used this method in this precise case. Keep in mind that our equation is not a prescribed curvature equation on the surface, but rather a prescribed curvature equation on a disc bundle on it.

Lemma 4.4. *J is well-defined and weakly upper semicontinuous on $W_0^{1,2}(\Sigma)$.*

Proof. We rely on the Moser–Trudinger inequality proved in [Li 2001] on any compact Riemannian surface. There is a constant $C(\Sigma)$ such that, for any $w \in W^{1,2}(\Sigma)$ with zero average,

$$\int_\Sigma \exp\left(\frac{4\pi w^2}{\|\nabla w\|_2^2}\right) \le C(\Sigma). \tag{4-23}$$

To prove that J is well-defined, we dominate e^{2w} in the following way:

$$e^{-2u}|\alpha|^2 e^{2w} \le \max(e^{-2u}|\alpha|^2) \exp\left(\frac{4\pi w^2}{\|\nabla w\|_2^2}\right) \exp\left(\frac{\|\nabla w\|_2^2}{4\pi}\right). \tag{4-24}$$

We deduce that $e^{-2u}|\alpha|^2 e^{2w}$ is integrable, and J_u is well-defined. It remains to prove the weak semicontinuity of J_u.

First, since closed convex sets are weakly closed, the L^2-norm squared of the gradient, being a continuous convex functional on $W^{1,2}(\Sigma)$, is weakly lower semicontinuous, and hence its opposite is upper semicontinuous. We finish this proof by proving that the functional

$$H(w) = \int e^{-2u}|\alpha|^2 e^{2w} \tag{4-25}$$

is weakly continuous. Let $w \in W^{1,2}(\Sigma)$, and $w_n \rightharpoonup w$. Since it is weakly converging, it is bounded, and there is a uniform upper bound $A > 0$ on the sequence $\|\nabla(w_n - w)\|_2^2$. For $t > 0$, $e^{t|w|}$ is integrable since we can control

$$t|w| \le \frac{4\pi w^2}{\|\nabla w\|_2^2} + \frac{t\|\nabla w\|_2^2}{4\pi}, \tag{4-26}$$

and apply the Moser–Trudinger inequality to obtain the upper bound

$$\int_\Sigma e^{t|w|} \leq C(\Sigma) \exp \frac{t\|\nabla w\|_2^2}{4\pi}. \tag{4-27}$$

This yields a constant $C(t)$ such that

$$\int_\Sigma e^{t|w_n - w|} \leq C(t). \tag{4-28}$$

To deduce the weak continuity, we use the bound

$$|e^x - 1| \leq |x| e^{|x|} \tag{4-29}$$

to control the convergence:

$$\int_\Sigma (e^{-2u}|\alpha|^2(e^{2w_n} - e^{2w})) \leq \max e^{-2u}|\alpha|^2 \int_\Sigma e^{2w}|e^{2(w_n - w)} - 1| \tag{4-30}$$

$$\leq \max e^{-2u}|\alpha|^2 \int_\Sigma e^{2w} e^{2|w_n - w|}|w_n - w| \tag{4-31}$$

$$\leq \max e^{-2u}|\alpha|^2 \left(\int e^{4w} e^{4|w_n - w|}\right)^{1/2} \|w_n - w\|_2 \tag{4-32}$$

$$\leq \max e^{-2u}|\alpha|^2 \left(\int e^{8w}\right)^{1/4} C(8)^{1/4} \|w_n - w\|_2. \tag{4-33}$$

In this last inequality all terms are bounded in n apart from $\|w_n - w\|_2$, which goes to zero since weak convergence in $W^{1,2}$ implies strong convergence in $L^2(\Sigma)$. This shows that H is weakly continuous. Hence, the functional J_u is weakly upper semicontinuous, as desired. □

In order to study explicitly the behavior of J, we will need a more precise control on the Sobolev constants of the functions involved, namely, we will use a precise Moser–Trudinger inequality obtained in [Bronstein 2023]. Recall that the space $W_0^{1,2}$ is the space of functions in $W^{1,2}$ whose average is zero.

Theorem 4.5 (Moser–Trudinger, Bronstein). *Denote by δ and Λ the systole and spectral gap of Σ, respectively. There is a constant $C(\delta, \Lambda)$ independent of the genus, depending only on δ, Λ and continuously in those parameters such that*

$$\int |\nabla u|^2 \leq 1 \;\Rightarrow\; \int e^{4\pi u^2} \leq \mathrm{Vol}(\Sigma) + C(\delta, \Lambda) \quad \textit{for all } u \in W_0^{1,2}(\Sigma). \tag{4-34}$$

Corollary 4.6. *For any $u \in W^{1,2}(\Sigma)$, $k \geq 1$ and $t > 0$, $e^{tu} \in L^k(\Sigma)$.*

Proof. Since $(e^{tu})^k = e^{tku}$ and $tku \in W^{1,2}(\Sigma)$, it is enough to prove it for $t = k = 1$. Also, since $e^{x+u} = e^x e^u$, it is sufficient to prove that for any $u \in W_0^{1,2}(\Sigma)$, e^u is integrable. Fixing now $u \in W_0^{1,2}(\Sigma)$ we know from the Moser–Trudinger inequality that for some $r > 0$, e^{ru^2} is integrable. Finally, since $0 < e^u \leq e^{ru^2} e^{1/4r}$, we deduce that e^u is integrable, as claimed. □

Corollary 4.7. *The functional J is upper bounded on $W_0^{1,2}(\Sigma)$, and we have the estimate*

$$\log\left(⨍ e^{-2u}|\alpha|^2\right) \le \sup J \le \log\max e^{-2u}|\alpha|^2 + \log\left(1 + \frac{C(\delta,\Lambda)}{\mathrm{Vol}(\Sigma)}\right). \tag{4-35}$$

Also, J is proper on $W_0^{1,2}(\Sigma)$, since

$$J(w) + \frac{1}{4\pi}|\nabla w|_2^2 \le \log\max e^{-2u}|\alpha|^2 + \log\left(1 + \frac{C(\delta,\Lambda)}{\mathrm{Vol}(\Sigma)}\right) \quad \textit{for all } w \in W_0^{1,2}(\Sigma). \tag{4-36}$$

Remark. The fact that the functions considered here have zero average is crucial, so that as a consequence of the Poincaré inequality the energy is proper on that space.

Proof. Let $w \in W_0^{1,2}(\Sigma)$. Consider that

$$2w \le 4\pi\frac{w^2}{|\nabla w|_2^2} + \frac{|\nabla w|_2^2}{4\pi}. \tag{4-37}$$

Exponentiating this inequality, we get

$$e^{2w} \le e^{|\nabla w|_2^2/4\pi} e^{4\pi(w/|\nabla w|_2)^2}. \tag{4-38}$$

Using Theorem 4.5 to bound the integral of the right-hand side,

$$\int e^{2w} \le (C(\delta,\Lambda) + \mathrm{Vol}(\Sigma)) e^{|\nabla w|_2^2/4\pi}. \tag{4-39}$$

We deduce

$$J(w) \le \log\max e^{-2u}|\alpha|^2 + \Big(\frac{1}{4\pi} - \frac{1}{c\,\mathrm{Vol}\,\Sigma}\Big)|\nabla w|_2^2 + \log\Big(1 + \frac{C(\delta,\Lambda)}{\mathrm{Vol}\,\Sigma}\Big). \tag{4-40}$$

As $\mathrm{Vol}(\Sigma) = 4\pi(g-1)$ and $c = 1/(2g-2)$, it follows that

$$\frac{1}{4\pi} - \frac{1}{c\,\mathrm{Vol}(\Sigma)} = -\frac{1}{4\pi}. \tag{4-41}$$

Inserting it on the left-hand side, we get

$$J(w) + \frac{1}{4\pi}|\nabla w|_2^2 \le \log\max e^{-2u}|\alpha|^2 + \log\Big(1 + \frac{C(\delta,\Lambda)}{\mathrm{Vol}(\Sigma)}\Big). \tag{4-42}$$

So J is indeed proper and upper bounded with the desired explicit upper bound.

The lower estimate for $\sup J$ is achieved by explicitly calculating $J(0) \le \sup J$. □

Proposition 4.8. *The functional J attains its maximum at $w \in W_0^{1,2}(\Sigma)$ satisfying*

$$\frac{1}{4\pi}|\nabla w|_2^2 \le \log\frac{\max e^{-2u}|\alpha|^2}{⨍ e^{-2u}|\alpha|^2} + \log\Big(1 + \frac{C(\delta,\Lambda)}{\mathrm{Vol}\,\Sigma}\Big). \tag{4-43}$$

Proof. Thanks to Corollary 4.7, we know that J is upper bounded. Consider a maximizing sequence (w_n) of J in $W_0^{1,2}(\Sigma)$. We have the estimate

$$J(w_n) + \frac{1}{4\pi}|\nabla w_n|_2^2 \le \log\max e^{-2u}|\alpha|^2 + \log\Big(1 + \frac{C(\delta,\Lambda)}{\mathrm{Vol}\,\Sigma}\Big). \tag{4-44}$$

As $(J(w_n))$ is upper bounded by $\sup J$, so is $(|\nabla w_n|_2^2)$. We deduce that $(|\nabla w_n|_2^2)$ is bounded. The explicit bound, assuming $J(w_n) \geq J(0)$, is

$$\frac{1}{4\pi}|\nabla w_n|_2^2 \leq \log \frac{\max e^{-2u}|\alpha|^2}{⨍ e^{-2u}|\alpha|^2} + \log\Big(1 + \frac{C(\delta, \Lambda)}{\operatorname{Vol}\Sigma}\Big). \tag{4-45}$$

As (w_n) is bounded in $W_0^{1,2}(\Sigma)$, extracting a subsequence we can assume it converges weakly towards some $w \in W_0^{1,2}(\Sigma)$. By Lemma 4.4, J is weakly continuous, so w is a maximizer of J. As $|\nabla w|_2^2 \leq \limsup |\nabla w_n|_2^2$, we get the desired bound on w. □

Theorem 4.9. *Let N be of degree 1, $\alpha \in H^0(K^2 N)$, $c = 1/(2g-2)$ and $u \in C^{0,\alpha}\big(\Sigma, \big[-\frac{1}{2}\ln 2, 0\big]\big)$. Then there is a solution v to*

$$\Delta v = c - e^{2v} e^{-2u}|\alpha|^2 \tag{4-46}$$

satisfying

$$\frac{1}{4\pi}|\nabla v|_2^2 \leq \log \frac{\max e^{-2u}|\alpha|^2}{⨍ e^{-2u}|\alpha|^2} + \log\Big(1 + \frac{C(\delta, \Lambda)}{\operatorname{Vol}\Sigma}\Big). \tag{4-47}$$

Proof. Thanks to Proposition 4.8, there exists $w \in W_0^{1,2}(\Sigma)$ a maximizer of the functional J. For any $s \in W_0^{1,2}$, we have $D_w J \cdot s = 0$ and so

$$\frac{2}{\operatorname{Vol}(\Sigma)} \frac{\int e^{-2u}|\alpha|^2 e^{2w} s}{⨍ e^{-2u}|\alpha|^2 e^{2w}} - \frac{2}{c \operatorname{Vol}(\Sigma)} \int \nabla w \cdot \nabla s = 0. \tag{4-48}$$

Weakly, this means

$$\frac{1}{c}\Delta w + \frac{e^{-2u}|\alpha|^2 e^{2w}}{⨍ e^{-2u}|\alpha|^2 e^{2w}} \in (L_0^2)^{\perp}. \tag{4-49}$$

As the orthogonal of the zero average functions are the constant maps, there is a constant A such that

$$\Delta w = A - c\frac{e^{-2u}|\alpha|^2 e^{2w}}{⨍ e^{-2u}|\alpha|^2 e^{2w}}. \tag{4-50}$$

Note that since $w \in W_0^{1,2}(\Sigma)$, applying Corollary 4.6, $e^{2w} \in L^2(\Sigma)$. Hence, by [Gilbarg and Trudinger 1977, Theorem 8.8], it follows that $w \in W^{2,2}(\Sigma)$ and the equation (4-50) is satisfied strongly. Via the Morrey embedding $W^{2,2} \subset C^{0,\alpha}$ for any $\alpha < 1$, it follows from (4-50) that Δw is Hölder continuous. By elliptic regularity, we deduce that $w \in C^{2,\alpha}(\Sigma)$, and the equality (4-50) is satisfied in $C^{0,\alpha}(\Sigma)$.

Writing $\int \Delta w = 0$, we deduce

$$A = c. \tag{4-51}$$

Consider the element

$$v = w + \frac{1}{2}\ln\left(\frac{c}{⨍ e^{-2u}|\alpha|^2 e^{2w}}\right) \in W^{1,2}(\Sigma). \tag{4-52}$$

As $\Delta v = \Delta w$, we get

$$\Delta v = c - e^{-2u}|\alpha|^2 e^{2v}. \tag{4-53}$$

As $|\nabla w|_2^2 = |\nabla v|_2^2$, the estimate of Proposition 4.8 ensures the desired bound on $|\nabla v|_2^2$. □

Remark. In contrast with the Gauss equation, we cannot in this case establish the uniqueness of the solution v. We will address this problem by showing stability of the solutions under some conditions after stating some precise estimates on the solution w and v.

4.3. ***Estimating the regularity of the solution.*** In order to apply later a fixed point theorem, we need to understand the behavior of the solutions v. Particularly, we need an L^∞ bound on $e^{2v}|\alpha|^2$. Because of the sign difference with Gauss's equation, applying a maximum principle won't yield bounds on v. We get our upper bound via an estimate of the Hessian of a maximizer w of J_u.

Proposition 4.10. *Assume $u \in C^{0,\alpha}(\Sigma)$. Then there is a constant $C(\delta, \Lambda)$ such that, for any maximizer w of the functional J_u, it satisfies*

$$|\nabla^2 w|_2^2 \le c^2\left((\operatorname{Vol}\Sigma + C(\delta,\Lambda))\left(\frac{\max|\alpha|^2}{⨍|\alpha|^2}\right)^2 - \operatorname{Vol}\Sigma\right) + 4\pi\left(\log\frac{\max e^{-2u}|\alpha|^2}{⨍ e^{-2u}|\alpha|^2} + \log\left(1 + \frac{C(\delta,\Lambda)}{\operatorname{Vol}\Sigma}\right)\right). \tag{4-54}$$

Proof. Let w be a maximizer of J. First, we use the Bochner identity to estimate the Hessian of w; see [Jost 2017, Theorem 4.5.1]. We have

$$-\tfrac{1}{2}\Delta(|\nabla w|^2) = -\sigma(\Delta\nabla w, \nabla w) + |\nabla^2 w|^2 + \operatorname{Ric}(\nabla \mathrm{w}, \nabla \mathrm{w}), \tag{4-55}$$

where σ is the hyperbolic metric and Ric is its Ricci tensor. As the metric is hyperbolic, we have $\operatorname{Ric} = -\sigma$. Integrating this formula yields the identity

$$|\nabla^2 w|_2^2 = \int (\Delta w)^2 + \int |\nabla w|^2. \tag{4-56}$$

Since Δw satisfies

$$\Delta w = c - \frac{c|\alpha|^2 e^{-2u} e^{2w}}{⨍ e^{-2u}|\alpha|^2 e^{2w}}, \tag{4-57}$$

we deduce that

$$\int (\Delta w)^2 = -c^2 \operatorname{Vol}\Sigma + c^2(\operatorname{Vol}\Sigma)^2 \frac{\int e^{-4u}|\alpha|^4 e^{4w}}{\left(\int e^{-2u}|\alpha|^2 e^{2w}\right)^2}. \tag{4-58}$$

To control the right-hand side, we estimate

$$4w \le \frac{4\pi w^2}{|\nabla w|_2^2} + \frac{|\nabla w|_2^2}{\pi} \tag{4-59}$$

and apply the Moser–Trudinger inequality of Theorem 4.5 to get

$$\int e^{4w} \le (\operatorname{Vol}\Sigma + C(\delta,\Lambda)) e^{|\nabla w|_2^2/\pi}. \tag{4-60}$$

This ensures

$$\int |\alpha|^4 e^{-4u} e^{4w} \le (\operatorname{Vol}\Sigma + C(\delta,\Lambda)) \max e^{-4u}|\alpha|^4 \exp\left(\frac{|\nabla w|_2^2}{\pi}\right). \tag{4-61}$$

It remains to estimate the denominator of the right-hand side of (4-58). As $c \operatorname{Vol} \Sigma = 2\pi$ and by maximality $J(w) \geq J(0)$, we have

$$\log \fint |\alpha|^2 e^{-2u} \leq \log \fint e^{-2u}|\alpha|^2 e^{2w} - \frac{1}{2\pi}|\nabla w|_2^2. \tag{4-62}$$

Exponentiating this,

$$\int e^{-2u}|\alpha|^2 \exp\Big(\frac{1}{2\pi}|\nabla w|_2^2\Big) \leq \int e^{-2u}|\alpha|^2 e^{2w}. \tag{4-63}$$

Combining this with the upper bound on the numerator, we get

$$\frac{\int e^{-4u}|\alpha|^4 e^{4w}}{\big(\int e^{-2u}|\alpha|^2 e^{2w}\big)^2} \leq \frac{(\operatorname{Vol}\Sigma + C(\delta, \Lambda)) \max e^{-4u}|\alpha|^4 \exp(|\nabla w|_2^2/\pi)}{\big(\int e^{-2u}|\alpha|^2 \exp(|\nabla w|_2^2/2\pi)\big)^2}. \tag{4-64}$$

Rearranging the right-hand side, we obtain

$$\int (\Delta w)^2 \leq -c^2 \operatorname{Vol}\Sigma + c^2(\operatorname{Vol}\Sigma + C(\delta, \Lambda)) \frac{(\max e^{-2u}|\alpha|^2)^2}{\big(\fint e^{-2u}|\alpha|^2\big)^2}. \tag{4-65}$$

Combining with the bound on $|\nabla w|_2^2$ from Proposition 4.8, we get the desired result. □

We sum up with the local upper bound on the solution, which will be key to building solutions of the PDE system:

Proposition 4.11. *Let $u : \Sigma \to \mathbb{R}$ be a map in $C^{0,\alpha}(\Sigma)$ such that*

$$1 \leq e^{-2u} \leq 2. \tag{4-66}$$

Then there is a constant $C_1 = C_1\big(\delta, \Lambda, |\alpha|_\infty^2/\fint |\alpha|^2\big)$, and $v \in C^{2,\alpha}(\Sigma)$ a solution of

$$\Delta v = c - e^{-2u} e^{2v}|\alpha|^2 \tag{4-67}$$

satisfying

$$\sup e^{2v}|\alpha|^2 \leq \frac{C_1\big(\delta, \Lambda, |\alpha|_\infty^2/\fint |\alpha|^2\big)}{2g-2}. \tag{4-68}$$

The constant C_1 is uniformly bounded if the parameters δ, Λ are bounded away from zero and $|\alpha|_\infty^2/\fint |\alpha|^2$ is bounded away from infinity.

Proof. Once again, consider $w \in W_0^{1,2}$ a maximizer of J and build the solution v as a translate of w so that

$$w = v - \bar{v}. \tag{4-69}$$

Thanks to Proposition 4.10, there is a bound $C_1\big(\delta, \Lambda, |\alpha|_\infty^2/\fint |\alpha|^2\big)$ such that

$$\|\nabla^2 w\|_2 \leq C_1\bigg(\delta, \Lambda, \frac{|\alpha|_\infty^2}{\fint |\alpha|^2}\bigg). \tag{4-70}$$

But from Proposition 4.8, there is a bound $C_2\big(\delta, \Lambda, |\alpha|_\infty^2/\fint |\alpha|^2\big)$ such that the gradient norm satisfies

$$\|\nabla w\|_2 \leq C_2\bigg(\delta, \Lambda, \frac{|\alpha|_\infty^2}{\fint |\alpha|^2}\bigg). \tag{4-71}$$

Also, since w has zero average, its L^2-norm can be bounded with the help of the Poincaré inequality:

$$\|w\|_2^2 \leq \Lambda C_2\left(\delta, \Lambda, \frac{|\alpha|_\infty^2}{⨍ |\alpha|^2}\right). \tag{4-72}$$

Hence, there is a constant $C_3\big(\delta, \Lambda, |\alpha|_\infty^2/⨍ |\alpha|^2\big)$ such that

$$\|w\|_{2,2} \leq C_3\left(\delta, \Lambda, \frac{|\alpha|_\infty^2}{⨍ |\alpha|^2}\right). \tag{4-73}$$

As the continuous embedding $W^{2,2} \hookrightarrow L^\infty$ depends only on the systole δ, this implies that we have $C_4 > 0$ with the same dependencies such that

$$|w|_\infty \leq C_4\left(\delta, \Lambda, \frac{|\alpha|_\infty^2}{⨍ |\alpha|^2}\right). \tag{4-74}$$

We now compute the average of $e^{2v}|\alpha|^2$:

$$c = ⨍ e^{-2u}e^{2v}|\alpha|^2 = e^{2\bar{v}} ⨍ e^{-2u}e^{2w}|\alpha|^2, \tag{4-75}$$

from which we get the estimate (recalling that $c = 1/(2g-2)$)

$$e^{2\bar{v}}e^{-2C_4} ⨍ |\alpha|^2 \leq \frac{1}{2g-2} \leq 2e^{2\bar{v}}e^{2C_4} ⨍ |\alpha|^2. \tag{4-76}$$

Combining this, we get

$$e^{2v}|\alpha|^2 = e^{2\bar{v}}e^{2w}|\alpha|^2 \leq \frac{e^{4C_4}}{2g-2}\frac{|\alpha|_\infty^2}{⨍ |\alpha|^2}, \tag{4-77}$$

as asserted. As the constant C_1 from Proposition 4.10 is continuous in the parameters, so is the exhibited constant C_4. □

It remains to show the stability of the solutions under suitable perturbations.

Lemma 4.12. *Consider a function $f \geq 0$ in $C^{0,\alpha}(\Sigma)$ such that there is a solution v of*

$$\Delta v = c - e^{2v} f \tag{4-78}$$

satisfying $2e^{2v} f < \Lambda$. Consider the operator $H = \Delta + 2e^{2v} f : W^{2,2}(\Sigma) \to L^2(\Sigma)$. Then H is invertible, and we have the upper bound on its inverse:

$$\|H^{-1}\| \leq \max\{(\Lambda - 2|e^{2v} f|_\infty)^{-1}, c^{-1}\}. \tag{4-79}$$

Proof. H is a self-adjoint operator between Hilbert spaces. Moreover, for λ large enough, we have the estimate

$$|(\lambda I - H)w|_2^2 \, |(\lambda - e^{2v} f)w - \Delta w|_2^2 \tag{4-80}$$

$$\leq \big(\lambda - \tfrac{1}{2}\Lambda\big)^2 \int w^2 + \int (\Delta w)^2 - 2\int (\lambda - e^{2v} f) w \Delta w \tag{4-81}$$

$$\leq C|w|_{2,2}^2. \tag{4-82}$$

Hence, for λ large enough, $(\lambda I - H)$ has compact inverse, so it is a Fredholm operator, and H has discrete spectrum.

To conclude the lemma, it remains to check that H has no eigenvalue in the range $(-\Lambda + |2e^{2v} f|_\infty, 2c)$.

Let t be in this range. If $\dot{v}$ is an eigenvector for the value t, we have

$$\Delta \dot{v} + (2e^{2v} f - t)\dot{v} = 0. \tag{4-83}$$

Decompose $\dot{v} = m(\dot{v}) + \check{v}$ with $m(\dot{v})$ being the average of $\dot{v}$. Then it becomes

$$m(\dot{v}) \int 2e^{2v} f - t = -\int (2e^{2v} f - t)\check{v}, \tag{4-84}$$

$$\int |\nabla \check{v}|^2 - (2e^{2v} f - t)\check{v}^2 = m(\dot{v}) \int (2e^{2v} f - t)\check{v}. \tag{4-85}$$

Combining these two equations with $\int e^{2v} f = c$, we get

$$\int |\nabla \check{v}|^2 - (2e^{2v} f - t)\check{v}^2 = -2m(v)^2 (2c - t) \operatorname{Vol} \Sigma. \tag{4-86}$$

As $t < 2c$, the right-hand side is nonpositive, while because $-\Lambda + 2e^{2v} f < t$, the left-hand side is nonnegative. So both must be zero, and $\dot{v} = 0$, as desired. □

A consequence of this lemma is that, for a varying family of functions f, we are locally able to lift branches of solutions to $\Delta v = c - e^{2v} f$. But we will need to be able to lift to a continuous branch for all f obtained from solving the Gauss–Codazzi equation. To this aim, we give a more precise description of the open set on which we are able to do so:

Theorem 4.13. *Let Σ be a closed hyperbolic surface. We denote by (C_i) constants depending on the systole, the spectral gap, and the ratio $|\alpha|_\infty^2 / \int |\alpha|^2$. Consider v_0 a solution of*

$$\Delta v = c - e^{2v} |\alpha|^2 \tag{4-87}$$

satisfying

$$e^{2v_0} |\alpha|^2 \leq \frac{C_1}{2g - 2} < \Lambda. \tag{4-88}$$

Fix $\varepsilon > 0$. Consider the functional $\mathcal{H} : C^{0,\alpha}(\Sigma) \times W_{2,2}(\Sigma) \to L^2(\Sigma)$ such that

$$\mathcal{H}(f, v) = \Delta v - c + e^{2v} f |\alpha|^2. \tag{4-89}$$

Then there is $C_2 > 0$ such that for $|f - 1|_\infty \leq C_2 / \sqrt{2g - 2}$, there is a unique $v \in W^{2,2}$ satisfying

$$\begin{cases} |v - v_0|_{2,2} \leq \varepsilon, \\ \qquad \Delta v = c - e^{2v} f |\alpha|^2. \end{cases} \tag{4-90}$$

In particular, this v also satisfies

$$e^{2v} |\alpha|^2 \leq \frac{C_3}{2g - 2}. \tag{4-91}$$

Proof. This explicit lift comes from an estimation of the bounds appearing in the implicit function theorem. We will use the version in [Holtzman 1970]. With the notation of the theorem, we need to compute $k_1 = |(D_v\mathcal{H}(f_0, v_0))^{-1}|$. The Lemma 4.12 shows that

$$k_1 \leq \tfrac{1}{2}(2g-2), \tag{4-92}$$

provided that $1/(2(2g-2)) \leq \Lambda$. Now we need to find $\delta > 0$ and g_1 such that, for $s \in [0, \delta]$, $v \in [0, \varepsilon]$,

$$|f-1| \leq \delta,\ |v-v_0|_{2,2} \leq \varepsilon \quad \Rightarrow \quad \|D_v\mathcal{H}(f, v) - D_v\mathcal{H}(1, v_0)\| \leq g_1(|f-1|_\infty, |v-v_0|_{2,2}). \tag{4-93}$$

But, as

$$|D_v\mathcal{H}(f, v)\dot{v} - D_v\mathcal{H}(1, v_0)\dot{v}|_2^2 = \int (2(e^{2v}f - e^{2v_0})|\alpha|^2\dot{v})^2 \leq \frac{4C_1^2}{(2g-2)^2} e^{4|v-v_0|_\infty} |1-f|_\infty |\dot{v}|_2^2, \tag{4-94}$$

we deduce

$$\|D_v\mathcal{H}(f, v) - D_v\mathcal{H}(1, v_0)\| \leq \frac{C_4}{2g-2} e^{2C_5|v-v_0|_{2,2}} |1-f|_\infty. \tag{4-95}$$

Finally, we must find g_2 such that

$$|f-1|_\infty \leq \delta \ \Rightarrow\ \|\mathcal{H}(f, v_0)\| \leq g_2(|f-1|_\infty). \tag{4-96}$$

Explicitly,

$$|\mathcal{H}(f, v_0)|_2^2 = \int (e^{2v_0}|\alpha|^2(f-1)^2)^2 \leq \frac{C_6}{2g-2}, \tag{4-97}$$

from which we get

$$g_2(s) = \frac{C_7 s}{\sqrt{2g-2}}. \tag{4-98}$$

Now we can lift a continuous branch satisfying $\mathcal{H}(f, \Psi(f)) = 0$ provided that δ satisfies

$$k_1 g_1(\delta, \varepsilon) \leq \eta < 1 \quad \text{and} \quad k_1 g_2(\delta) \leq \varepsilon(1-\eta). \tag{4-99}$$

This is verified for δ of the form

$$\delta = \frac{C_8}{\sqrt{2g-2}}, \tag{4-100}$$

and the solutions $v = \Psi(f)$ satisfy

$$|v - v_0|_{2,2} \leq \varepsilon, \tag{4-101}$$

from which we deduce

$$e^{2v}|\alpha|^2 \leq e^{2C_9|v-v_0|_{2,2}} e^{2v_0}|\alpha|^2 \leq \frac{C_{10}}{2g-2}, \tag{4-102}$$

as desired. □

5. Proof of the main theorem

In this section we prove the geometric meaning of all this study:

Theorem 5.1. *Let Σ be a closed hyperbolic surface of genus g. Fix $\eta \in (0, 1)$ and consider a family $(\Sigma_n, N_n, \alpha_n)_{n\geq 2}$ satisfying*

(a) *Σ_n is a degree n cover of Σ, equipped with the lift of the hyperbolic metric on Σ;*

(b) *N_n is a degree $4n(g-1)+1$ line bundle over Σ_n, equipped with a metric of constant sectional curvature*;

(c) *$\alpha_n \in H^0(N_n)$ is a section of N_n*;

(d) *there is $\Lambda_0 > 0$ such that for all n the spectral gap of Σ_n is bigger than Λ_0*;

(e) *the family of sections is balanced, i.e.*,

$$\sup_{n\geq 2} \frac{|\alpha_n|_\infty^2}{\fint |\alpha_n|^2} < \infty. \tag{5-1}$$

Then, for n large enough, there are $u, v : \Sigma_n \to \mathbb{R}$ such that

$$\sup_{z\in\Sigma_n} e^{-4u}e^{2v}|\alpha|^2(z) < \eta, \tag{5-2}$$

$$\begin{cases} \Delta u = e^{2u} - 1 + e^{-2u}e^{2v}|\alpha_n|^2, \\ \Delta v = \dfrac{1}{n(2g-2)} - e^{-2u}e^{2v}|\alpha_n|^2. \end{cases} \tag{5-3}$$

As a corollary, we get the desired existence theorem for almost-fuchsian representations with nontrivial normal bundle.

Corollary 5.2. *There is a genus g_0 such that for any $g \geq g_0$, there is a representation $\rho : \pi_1\Sigma_g \to \mathrm{SO}(4,1)$ which is almost-fuchsian and a complex variation of Hodge structure, and for which the hyperbolic 4-manifold $\rho\backslash\mathbb{H}^4$ is a degree 1 disc bundle over Σ_g.*

Proof. Assuming Theorem 5.1, let Σ be a closed hyperbolic surface of genus 2. Thanks to the result of [Magee et al. 2022], we know the existence of a family of covers (Σ_n) satisfying condition (d) of Theorem 5.1. Thanks to Theorem 3.3, there is a balanced sequence (α_n, N_n) satisfying conditions (b), (c) and (e). Thus for n large enough, we have (u, v) satisfying

$$\sup_{z\in\Sigma_n} e^{-4u}e^{2v}|\alpha|^2(z) < 1, \tag{5-4}$$

$$\begin{cases} \Delta u = e^{2u} - 1 + e^{-2u}e^{2v}|\alpha_n|^2, \\ \Delta v = \dfrac{1}{n(2g-2)} - e^{-2u}e^{2v}|\alpha_n|^2. \end{cases} \tag{5-5}$$

Thanks to Proposition 2.11, this corresponds to an almost-fuchsian representation and complex variation of Hodge structure, with holomorphic data α, and the hyperbolic 4-manifold is a degree 1 disc bundle over Σ_n. □

The main argument of Theorem 5.1 will be Schauder's fixed point theorem: a continuous map which preserves a compact convex set has a fixed point.

Proof of Theorem 5.1. Denote by $\mathcal{K}_u^n$ the convex set

$$\mathcal{K}_u^n := \left\{u \in C^{0,\alpha}(\Sigma_n) : -\tfrac{1}{2}\ln(1+\eta) \leq u \leq 0\right\}. \tag{5-6}$$

Thanks to Theorem 4.9, for $u = 0$, there is a function v_0 which is a solution of

$$\Delta v_0 = \frac{1}{|\chi(\Sigma_n)|} - e^{2v_0}|\alpha_n|^2. \tag{5-7}$$

By Proposition 4.11, there is $C_0 > 0$ depending continuously on δ, Λ and the ratio $|\alpha_n|^2_\infty / \int |\alpha_n|^2$ such that

$$e^{2v_0}|\alpha_n|^2 \le \frac{C_0}{2n(g-1)}. \tag{5-8}$$

The assumptions (a), (d) and (e) ensure that the constant C can be chosen to be independent of n, as it must be for our proof. Recall Λ_0 defined in (d) as a common lower bound to the spectral gaps of (Σ_n). Consider now n large enough that

$$\frac{C_0}{2n(g-1)} < \min\left\{\Lambda_0, \frac{1}{2}\right\}. \tag{5-9}$$

Applying Theorem 4.13 for some fixed $\varepsilon > 0$, there are constants C_1, C_2, C_3 independent of n and a continuous map

$$\Psi : \left\{ u \in C^{0,\alpha}(\Sigma_n) : |e^{2u} - 1|_\infty \le \frac{C_1}{\sqrt{2n(g-1)}} \right\} \to \{ v \in C^{2,\beta}(\Sigma_n) : |v - v_0|_\infty \le C_2\varepsilon \} \tag{5-10}$$

such that $v = \Psi(u)$ is a solution of

$$\Delta v = \frac{1}{|\chi(\Sigma_n)|} - e^{2v}e^{-2u}|\alpha_n|^2, \tag{5-11}$$

and $v = \Psi(u)$ satisfies

$$e^{2v}|\alpha_n|^2 \le \frac{C_3}{2n(g-1)}. \tag{5-12}$$

Here β is any exponent $0 < \beta < \alpha$. Provided $C_3/(2n(g-1)) \le \eta/(1+\eta)^2$, Theorem 4.1 gives us a map

$$\Phi : \left\{ v \in C^{0,\alpha}(\Sigma_n) : e^{2v}|\alpha_n|^2 \le \frac{C_3}{2n(g-1)} \right\} \to \left\{ u \in C^{2,\beta}(\Sigma_n) : |e^{2u} - 1|_\infty \le \frac{C_4}{2n(g-1)} \right\} \tag{5-13}$$

for an n-independent constant C_4 such that $u = \Phi(v)$ is a solution of

$$\Delta u = e^{2u} - 1 + e^{-2u}e^{2v}|\alpha_n|^2 \tag{5-14}$$

and u satisfies

$$\sup e^{-4u}e^{2v}|\alpha_n|^2 < \eta. \tag{5-15}$$

Finally, for n sufficiently large,

$$\frac{C_4}{2n(g-1)} \le \frac{C_1}{\sqrt{2n(g-1)}}, \tag{5-16}$$

and $\Phi \circ \Psi$ is a well-defined self-map of $\mathcal{K} = \{u : |e^{2u} - 1| \le C_4/(2n(g-1))\} \subset \mathcal{K}^u_n$. We also point out that, by ellipticity, every element in the image of Φ is in $C^{2,\alpha}$ and satisfies

$$\begin{cases} -\frac{1}{2}\ln 2 \le u \le 0, \\ \quad -1 \le \Delta u \le 1. \end{cases} \tag{5-17}$$

And so the image of Φ is in the set

$$\mathcal{Q} = \left\{ u \in C^{2,\alpha}(\Sigma_n) : -1 \leq \Delta u \leq 1 \text{ and } -\tfrac{1}{2}\ln 2 \leq u \leq 0 \right\}, \tag{5-18}$$

whose closure in $C^{0,\alpha}(\Sigma_n)$ is convex and compact, since the uniform bounds on u and Δu induce uniform bounds on the L^2-norms of u and ∇u, the Bochner formula and the bound on Δu induce a uniform bound on $\nabla^2 u$, and the embedding $W^{2,2} \hookrightarrow C^{0,\alpha}$ is compact.

Finally, $\Phi \circ \Psi$ is a continuous self-map of $\mathcal{K}$, which is convex, and its image is contained in a compact convex set. Hence, by the Schauder fixed point theorem, it has a fixed point u. Let $v = \Psi(u)$. Then (u, v) is solution of the desired PDE system, as asserted. By construction, it satisfies the estimate

$$e^{-4u} e^{2v} |\alpha_n|^2 \leq \eta, \tag{5-19}$$

completing the proof. □

Acknowledgment

The author thanks Nicolas Tholozan (DMA) for his help and guidance during this work.

References

[Battaglia et al. 2015] L. Battaglia, A. Jevnikar, A. Malchiodi, and D. Ruiz, "A general existence result for the Toda system on compact surfaces", *Adv. Math.* **285** (2015), 937–979. MR

[Bronstein 2023] S. Bronstein, "On the Moser–Trudinger inequality for surfaces", preprint, 2023. arXiv 2305.06662

[Bronstein and Smith 2023] S. Bronstein and G. A. Smith, "A geometric description of the set of almost-fuchsian representations", preprint, 2023. arXiv 2301.00715v1

[Bryant 1982] R. L. Bryant, "Conformal and minimal immersions of compact surfaces into the 4-sphere", *J. Differential Geometry* **17**:3 (1982), 455–473. MR

[Chang and Yang 2003] S.-Y. A. Chang and P. C. Yang, "The inequality of Moser and Trudinger and applications to conformal geometry", *Comm. Pure Appl. Math.* **56**:8 (2003), 1135–1150. MR

[Choudhury et al. 2024] D. Choudhury, F. Mazzoli, and A. Seppi, "Quasi-Fuchsian manifolds close to the Fuchsian locus are foliated by constant mean curvature surfaces", *Math. Ann.* **388**:4 (2024), 3981–4010. MR

[Davalo 2025] C. Davalo, "Nearly geodesic immersions and domains of discontinuity", *Geom. Topol.* **29**:5 (2025), 2391–2461. MR

[Donaldson 2003] S. K. Donaldson, "Moment maps in differential geometry", pp. 171–189 in *Surveys in differential geometry* (Boston, 2002), vol. VIII, edited by S.-T. Yau, Surv. Differ. Geom. **8**, International Press, Somerville, MA, 2003. MR

[Evans 2010] L. C. Evans, *Partial differential equations*, 2nd ed., Grad. Studies in Math. **19**, Amer. Math. Soc., Providence, RI, 2010. MR

[Forstnerič 2021] F. Forstnerič, "The Calabi–Yau property of superminimal surfaces in self-dual Einstein four-manifolds", *J. Geom. Anal.* **31**:5 (2021), 4754–4780. MR

[Gilbarg and Trudinger 1977] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Grundl. Math. Wissen. **224**, Springer, 1977. MR

[Gromov et al. 1988] M. Gromov, H. B. Lawson, Jr., and W. Thurston, "Hyperbolic 4-manifolds and conformally flat 3-manifolds", *Inst. Hautes Études Sci. Publ. Math.* 68 (1988), 27–45. MR

[Guest and Lin 2014] M. A. Guest and C.-S. Lin, "Nonlinear PDE aspects of the tt* equations of Cecotti and Vafa", *J. Reine Angew. Math.* **689** (2014), 1–32. MR

[Guo et al. 2010] R. Guo, Z. Huang, and B. Wang, "Quasi-Fuchsian 3-manifolds and metrics on Teichmüller space", *Asian J. Math.* **14**:2 (2010), 243–256. MR

[Ho 2014] S. L. Ho, *On conformally flat circle bundles over surfaces*, Ph.D. thesis, University of Maryland, College Park, 2014, available at http://hdl.handle.net/1903/15825. MR

[Hodge 2005] T. Hodge, *Hyperkähler geometry and Teichmüller space*, Ph.D. thesis, Imperial College London, 2005.

[Holtzman 1970] J. M. Holtzman, "Explicit ε and δ for the implicit function theorem", *SIAM Rev.* **12** (1970), 284–286. MR

[Huang and Wang 2015] Z. Huang and B. Wang, "Counting minimal surfaces in quasi-Fuchsian three-manifolds", *Trans. Amer. Math. Soc.* **367**:9 (2015), 6063–6083. MR

[Huybrechts 2005] D. Huybrechts, *Complex geometry*, Springer, 2005. MR

[Jiang 2025] R. Jiang, "Counting essential minimal surfaces in closed negatively curved n-manifolds", *J. Geom. Anal.* **35**:1 (2025), art. id. 5. MR

[Jost 2017] J. Jost, *Riemannian geometry and geometric analysis*, 7th ed., Springer, 2017. MR

[Kapovich 1989] M. E. Kapovich, "Flat conformal structures on three-dimensional manifolds: existence problem", *Sib. Math. J.* **30**:5 (1989), 712–722. MR

[Kommerell 1897] K. Kommerell, *Die Krümmung der zweidimensionalen Gebilde im ebenen Raum von vier Dimensionen*, H. Laupp, Tübingen, 1897.

[Krasnov and Schlenker 2007] K. Krasnov and J.-M. Schlenker, "Minimal surfaces and particles in 3-manifolds", *Geom. Dedicata* **126** (2007), 187–254. MR

[Kuiper 1988] N. H. Kuiper, "Hyperbolic 4-manifolds and tesselations", *Inst. Hautes Études Sci. Publ. Math.* **68** (1988), 47–76. MR

[Li 2001] Y. Li, "Moser–Trudinger inequality on compact Riemannian manifolds of dimension two", *J. Partial Differential Equations* **14**:2 (2001), 163–192. MR

[Li and Mochizuki 2025] Q. Li and T. Mochizuki, "Complete solutions of Toda equations and cyclic Higgs bundles over non-compact surfaces", *Int. Math. Res. Not.* **2025**:7 (2025), art. id. rnaf081. MR

[Loftin and McIntosh 2019] J. Loftin and I. McIntosh, "The moduli spaces of equivariant minimal surfaces in $\mathbb{RH}^3$ and $\mathbb{RH}^4$ via Higgs bundles", *Geom. Dedicata* **201** (2019), 325–351. MR

[Magee et al. 2022] M. Magee, F. Naud, and D. Puder, "A random cover of a compact hyperbolic surface has relative spectral gap $\frac{3}{16} - \varepsilon$", *Geom. Funct. Anal.* **32**:3 (2022), 595–661. MR

[Malchiodi and Ruiz 2013] A. Malchiodi and D. Ruiz, "A variational analysis of the Toda system on compact surfaces", *Comm. Pure Appl. Math.* **66**:3 (2013), 332–371. MR

[Meeks and Yau 1982] W. W. Meeks, III and S.-T. Yau, "The existence of embedded minimal surfaces and the problem of uniqueness", *Math. Z.* **179**:2 (1982), 151–168. MR

[Sacks and Uhlenbeck 1982] J. Sacks and K. Uhlenbeck, "Minimal immersions of closed Riemann surfaces", *Trans. Amer. Math. Soc.* **271**:2 (1982), 639–652. MR

[Schoen and Yau 1979] R. Schoen and S.-T. Yau, "Existence of incompressible minimal surfaces and the topology of three-dimensional manifolds with nonnegative scalar curvature", *Ann. of Math.* (2) **110**:1 (1979), 127–142. MR

[Seppi 2016] A. Seppi, "Minimal discs in hyperbolic space bounded by a quasicircle at infinity", *Comment. Math. Helv.* **91**:4 (2016), 807–839. MR

[Trautwein 2019] S. Trautwein, "The hyperkähler metric on the almost-Fuchsian moduli space", *EMS Surv. Math. Sci.* **6**:1-2 (2019), 83–131. MR

[Uhlenbeck 1983] K. K. Uhlenbeck, "Closed minimal surfaces in hyperbolic 3-manifolds", pp. 147–168 in *Seminar on minimal submanifolds*, edited by E. Bombieri, Ann. of Math. Stud. **103**, Princeton Univ. Press, 1983. MR

[Wentworth 2016] R. A. Wentworth, "Higgs bundles and local systems on Riemann surfaces", pp. 165–219 in *Geometry and quantization of moduli spaces*, Springer, 2016. MR

Received 10 Aug 2024. Revised 14 Jul 2025. Accepted 13 Oct 2025.

SAMUEL BRONSTEIN: bronstein@mis.mpg.de
Max Planck Institut für Mathematik in der Wissenschaften, Leipzig, Germany

ANALYSIS AND PDE
Vol. 19 (2026), No. 7, pp. 1395–1416
DOI: 10.2140/apde.2026.19.1395

ON THE OHSAWA–TAKEGOSHI L^2 EXTENSION THEOREM AND REMOVABLE SINGULARITIES OF PLURISUBHARMONIC FUNCTIONS

XIEPING WANG

To Li and Rui'an

The celebrated Ohsawa–Takegoshi extension theorem for L^2 holomorphic functions on bounded pseudoconvex domains in $\mathbb{C}^n$ is a fundamental result in the fields of complex analysis and algebraic geometry. In 1995, Ohsawa conjectured that the theorem holds more generally on bounded complete Kähler domains in $\mathbb{C}^n$. Recently, Chen, Wu and Wang confirmed this conjecture in a special case. We extend their result to the case of holomorphic sections of twisted canonical bundles over relatively compact complete Kähler domains in Stein manifolds. As an application, we establish a Hartogs-type extension theorem for plurisubharmonic functions across compact complete pluripolar sets. This result complements a classical theorem of Shiffman and may be regarded as a plurisubharmonic analogue of the Skoda–El Mir extension theorem, thereby filling a gap that appears to have remained open in the literature since at least 1985.

1. Introduction

One of the central results in the L^2 theory of several complex variables is the Ohsawa–Takegoshi L^2 extension theorem, a classical version of which can be stated as follows:

Theorem A [Ohsawa and Takegoshi 1987]. *Let Ω be a bounded pseudoconvex domain in $\mathbb{C}^n$ and H be a complex affine subspace of $\mathbb{C}^n$ intersecting Ω. Then, for every plurisubharmonic function φ on Ω and every holomorphic function f on $\Omega \cap H$ satisfying*

$$\int_{\Omega\cap H} |f|^2 e^{-\varphi}\, \mathrm{d}V_H < \infty,$$

there exists a holomorphic function F on Ω such that $F|_{\Omega\cap H} = f$ and

$$\int_{\Omega} |F|^2 e^{-\varphi}\, \mathrm{d}V \le C \int_{\Omega\cap H} |f|^2 e^{-\varphi}\, \mathrm{d}V_H,$$

where $C > 0$ is a constant depending only on the diameter of Ω and the codimension of H in $\mathbb{C}^n$, and $\mathrm{d}V$ (resp. $\mathrm{d}V_H$) denotes the Lebesgue measure on $\mathbb{C}^n$ (resp. H).

Since the groundbreaking work of Ohsawa and Takegoshi [1987], an extensive literature has developed around the above theorem, encompassing its far-reaching generalizations and applications in various contexts. The Ohsawa–Takegoshi L^2 extension theorem has proved to be of fundamental importance in the fields of several complex variables and complex algebraic geometry, serving as a highly effective tool for producing, by induction on dimension, holomorphic functions (or, more generally, holomorphic sections)

MSC2020: primary 32D15, 32D20, 32U05, 32U30; secondary 32Q15, 32W05.
Keywords: L^2 extension, $\bar\partial$-equation, complete Kähler manifolds, removable singularities, plurisubharmonic functions, complete pluripolar sets.

© 2026 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.

with uniformly controlled L^2 norms. We refer the reader to [Berndtsson 2010; Demailly 2012a; Ohsawa 2018; Siu 2002] for comprehensive treatments of this subject, and also to the surveys [Błocki 2014; McNeal and Varolin 2015; Berndtsson 2018; Ohsawa 2020] for brief introductions. Significant recent advances include work by Błocki [2013] and by Guan and Zhou [2015a; 2015b]. In particular, [Guan and Zhou 2015a] established an optimal-constant version of the Ohsawa–Takegoshi L^2 extension theorem in a very general setting, with many impressive applications. Their work provided further impetus to research in this area; see, for example, [Berndtsson and Lempert 2016; Cao et al. 2024; Chan and Choi 2022; Deng et al. 2023; 2024; Guan et al. 2025; 2026; Hosono 2020; Kim 2023; Li et al. 2023; Liu and Xu 2024; Rao and Zhang 2021; Xu and Zhou 2024; Zhou and Zhu 2018; 2020].

Despite the above achievements, several questions about the L^2 extension of holomorphic functions remain open, including the following long-standing one posed by Ohsawa himself (compare the first paragraph of [Berndtsson 2010, Section 6.2.1]):

Question B [Ohsawa 1995][1]. Does Theorem A hold more generally for bounded *complete Kähler* domains in $\mathbb{C}^n$?

For background on and motivation for this question, we refer the reader to the original paper of Ohsawa [1995] (and his recent book [Ohsawa 2018]). We only note here that — as has been amply demonstrated by the work of Demailly [1982] — Hörmander's weighted L^2 theory for the $\bar{\partial}$-operator on Stein manifolds extends equally well to more general complete Kähler manifolds; it is therefore natural to expect that the L^2 extension theory of Ohsawa–Takegoshi, which builds on the former, does as well. At present, however, Question B appears difficult to resolve in full generality. To the best of our knowledge, the only result known to date is due to Chen, Wu and Wang [Chen et al. 2015], who gave an affirmative answer in the special case in which H consists of a single point. Their result has found a few unexpected applications; see [Chen et al. 2015] for details.

In this paper we first extend the above result of Chen, Wu and Wang to the case of holomorphic sections of twisted canonical bundles over relatively compact complete Kähler domains in Stein manifolds, and then we present some new applications. Our first main result reads as follows:

Theorem 1.1. *Let X be an n-dimensional Stein manifold with a continuous volume form $\mathrm{d}V_X$ and Ω be a relatively compact complete Kähler domain in X. Suppose L is a holomorphic line bundle over Ω and h is a singular Hermitian metric on L with semipositive curvature current. Then, for every $x_0 \in \Omega$ and $s_{x_0} \in \Lambda^n T^*_{x_0}\Omega \otimes L_{x_0}$, there is an L-valued holomorphic n-form s on Ω such that $s(x_0) = s_{x_0}$ and*

$$\int_\Omega \sqrt{-1}^{n^2}\{s, s\}_h \le C\frac{\sqrt{-1}^{n^2}\{s_{x_0}, s_{x_0}\}_h}{\mathrm{d}V_X(x_0)}$$

provided the right-hand side is finite, where $C > 0$ is a constant depending only on n, Ω and $\mathrm{d}V_X$.

Here $\{\cdot, \cdot\}_h$ denotes the sesquilinear product combining the wedge product of scalar-valued forms with the Hermitian metric h on L (see Section 4 below for a precise definition).

[1] Although formulated slightly differently, this question is actually equivalent to Ohsawa's original question, as can be easily seen by induction on n.

We now sketch the idea of the proof of Theorem 1.1. Following the strategy of [Ohsawa 1995] and [Chen et al. 2015], we reduce Theorem 1.1 to an Ohsawa–Takegoshi-type L^2 existence theorem for the $\bar{\partial}$-operator on Hermitian holomorphic line bundles, subject to certain curvature constraints, over *complete Kähler* manifolds (see Theorem 3.2 below). Since the Hermitian metric on the line bundle under consideration is allowed to be *singular*, the proof of Theorem 3.2 relies heavily on the approximation theory of Demailly [1982; 1992; 1994] (see also [Błocki and Kołodziej 2007]), in which there are mainly two approximation theorems for singular Hermitian metrics:

- one is a local approximation by Hermitian metrics that are smooth on the complement of suitably chosen complex subvarieties in the manifold, with arbitrarily small loss of positivity in curvature;
- the other is a global approximation by Hermitian metrics that are smooth on the entire manifold but with less controllable loss of positivity in curvature.

The first one has been used by Zhou and Zhu [2018] (see also [Zhou and Zhu 2020]) to establish a singular-metric (and optimal-constant) version of the Ohsawa–Takegoshi–Manivel L^2 extension theorem on weakly pseudoconvex Kähler manifolds [Demailly 2000]. In the more general case of complete Kähler manifolds, the lack of pseudoconvexity however makes it hard to apply the first approximation theorem. Instead we make use of the second one, which combined with the a priori inequality for the twisted $\bar{\partial}$-operator due to Ohsawa and Takegoshi [1987] enables us to solve the $\bar{\partial}$-equation with two error terms via an induction argument, ultimately leading to Theorem 3.2. This idea of arguing by induction, inspired by [Demailly 1982], differs from that of [Chen et al. 2015] even when applied to the case they considered (see Remark 3.3(iv) for further explanation).

Next we present an application of Theorem 1.1 to the study of removable singularities of plurisubharmonic (psh for short) functions. Specifically, we prove the following somewhat surprising result.

Theorem 1.2. *Let Ω be a domain in a Stein manifold of dimension greater than or equal to 2 and K be a compact complete pluripolar subset of Ω. Then every psh function on $\Omega \setminus K$ admits a unique psh extension to Ω.*

It is worth emphasizing here that the full strength of Theorem 1.1 is required for the proof of this theorem, while its usual version in which Ω is Stein is not sufficient, as will be seen in Section 5.

Recently, Chen [2018] showed that the well-known Hartogs extension theorem for holomorphic functions is also valid for pluriharmonic functions; see also [Wang 2024a]. However, it has long been known that this is not the case for psh functions; see, e.g., [Bedford and Taylor 1988]. Given this situation, it is worthwhile to establish a Hartogs extension theorem of a slightly different nature for psh functions, as stated above. Our motivation for this result also stems from the renewed interest in the original Hartogs extension theorem in recent years; for details see the most recent work [Boggess et al. 2022; Vîjîitu 2022] and the references therein.

To put Theorem 1.2 in perspective, we recall that there are two landmark results in the extension theory of closed positive currents, namely the Harvey extension theorem [1974] and the Skoda–El Mir extension theorem [Skoda 1982; El Mir 1984]. These two theorems complement each other, and each unifies and generalizes many important earlier results due to Bishop, Remmert and Stein, Shiffman, Siu and others;

for more information see the original papers just mentioned and also [Harvey and Polking 1975; Harvey 1977; Sibony 1985]. Corresponding to Harvey's extension theorem there is an extension theorem for psh functions due to Shiffman [1972], which states that every psh function on an n-dimensional complex manifold extends plurisubharmonically across a closed set of Hausdorff $(2n-2)$-measure zero. Theorem 1.2 may be regarded, in a broad sense, as a plurisubharmonic analogue of the Skoda–El Mir extension theorem. To the best of our knowledge, no such analogue has previously appeared in the literature.

To further clarify the scope of Theorem 1.2, we recall the standard fact from pluripotential theory that pluripolar sets in an n-dimensional complex manifold have Hausdorff dimension at most $2n-2$. We also note that there exist compact complete pluripolar sets in $\mathbb{C}^n$ whose Hausdorff $(2n-2)$-measure is positive and may be either finite or infinite; see Section 2.1 for precise references. Together, these facts show that Theorem 1.2 is complementary to Shiffman's theorem described above, in the sense that each covers cases not covered by the other. Furthermore, Theorem 1.2 is global in nature, whereas Shiffman's theorem, like many other related results in the literature (see, for example, [Siu 1974; Harvey and Polking 1975; Chen et al. 2015]), is local in nature.

Interestingly, the proof of Theorem 1.2 involves a *combination* of the L^2 theory for the $\bar{\partial}$-operator (more specifically, Theorem 1.1), pluripotential theory, and some basic geometric measure theory; see Section 5 for details. By contrast, in much of the earlier work on extension problems for closed positive currents (particularly, complex varieties) and psh functions, the L^2 method and the geometric measure-theoretic method were generally used *separately*, notably by [Siu 1974] and by [Shiffman 1968; Harvey 1974; Harvey and Polking 1975], respectively. The use of Theorem 1.1 in the proof of Theorem 1.2 was inspired by [Chen et al. 2015] and [Demailly 1992]. It should also be mentioned that a weaker version of Theorem 1.2, under some additional assumptions, was already established by the author in his previous work [Wang 2024b] by quite a different method.

We now pose the following fairly natural question concerning Theorem 1.2:

Question 1.3. Does Theorem 1.2 still hold when K is only assumed to be a compact pluripolar subset of Ω?

The answer is likely to be positive, although it remains out of reach at the time of writing. We also refer the interested reader to [Chen et al. 2015, Conjecture 4.1] for a different but closely related problem.

Finally, we hope that the results presented in this paper (as well as in [Chen et al. 2015]) will, to some extent, serve as an incentive to investigate the Ohsawa problem and related topics.

The paper is organized as follows. In Section 2 we review some necessary notions and results that will be used in this paper. We then prove in Section 3 an Ohsawa–Takegoshi-type L^2 existence theorem for the $\bar{\partial}$-operator on complete Kähler manifolds, which is in turn used in Section 4 to prove Theorem 1.1. Finally, the proof of Theorem 1.2 is carried out in Section 5.

2. Preliminaries

In this section we fix some necessary notation and collect some fundamental results from pluripotential theory and the L^2 theory of the $\bar{\partial}$-operator that will be used in subsequent sections.

2.1. ***Complete pluripolar sets and their defining functions.*** We begin by recalling the definition of complete pluripolar sets. Let X be a complex manifold, and let $\mathrm{Psh}(X)$ denote the set of all psh functions on X.

Definition 2.1. A set $E \subset X$ is said to be *complete pluripolar* if for every point $x \in E$ there exists a neighborhood U of x and a function $\varphi \in \mathrm{Psh}(U)$ such that

$$E \cap U = \varphi^{-1}(-\infty).$$

The collection of all complex subvarieties of X constitutes a particularly important class of (closed) complete pluripolar sets, but complete pluripolar sets are much more general: for instance, the Cartesian product of finitely many (possibly distinct) Cantor-type sets of logarithmic capacity zero in the complex plane is a *compact* complete pluripolar set in the corresponding complex Euclidean space (see, e.g., [Ransford 1995]) but is obviously not complex-analytic. One may also consult [El Mir 1984; Levenberg et al. 1992; Edlund 2004; Dieu and Manh 2014] for many other nontrivial, constructive examples of *compact* complete pluripolar sets in $\mathbb{C}^n$, especially those of minimal Hausdorff codimension (i.e., of Hausdorff codimension 2). Using these examples one can easily construct a large number of compact complete pluripolar sets with positive (finite or infinite) Hausdorff $(2n-2)$-measure in general Stein manifolds of dimension n, where n is an arbitrary positive integer.

On Stein manifolds, Colţoiu [1990] proved the following important result concerning the existence of a global defining function for a closed complete pluripolar set.

Theorem 2.2 [Colţoiu 1990]. *Let X be a Stein manifold and $E \subset X$ be a closed complete pluripolar set. Then there is a function $\rho \in \mathrm{Psh}(X) \cap C^\infty(X \setminus E)$ such that*

$$\rho^{-1}(-\infty) = E \quad \textit{and} \quad \sqrt{-1}\partial\bar{\partial}\rho > 0 \ \textit{ on } X \setminus E.$$

As we shall see later, this result is crucial in the proof of Theorem 1.2.

2.2. ***Twisted Bochner–Kodaira–Nakano inequality.*** Let (X, ω) be an n-dimensional Kähler manifold and (L, h) be a Hermitian holomorphic line bundle over X. Given integers $0 \le p, q \le n$, we let $L^2(X, \Lambda^{p,q}T^*X \otimes L; \omega, h)$ denote the space of L-valued (p, q)-forms u with measurable complex coefficients on X satisfying

$$\|u\|_{\omega,h} := \left(\int_X |u|^2_{\omega,h}\, \mathrm{d}V_\omega\right)^{1/2} < \infty,$$

where $|\cdot|^2_{\omega,h}$ is the pointwise norm on $\Lambda^{p,q}T^*X \otimes L$ induced by ω and h, and $\mathrm{d}V_\omega = \omega^n/n!$ is the volume form on X induced by ω. We also denote by $\langle \cdot, \cdot \rangle_{\omega,h}$ (resp. $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\omega,h}$) the pointwise (resp. global) inner product corresponding to $|\cdot|^2_{\omega,h}$ (resp. $\|\cdot\|^2_{\omega,h}$). Obviously $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\omega,h}$ makes $L^2(X, \Lambda^{p,q}T^*X \otimes L; \omega, h)$ a Hilbert space, which contains as a dense subspace the space $\mathcal{D}(X, \Lambda^{p,q}T^*X \otimes L)$ of L-valued C^∞ (p, q)-forms with compact support in X. For our later purposes, we also need to consider the maximal weak extension of the $\bar{\partial}$-operator to $L^2(X, \Lambda^{p,q}T^*X \otimes L; \omega, h)$, which we still denote by $\bar{\partial}$. Then it is a closed, densely defined operator, and so is its Hilbert adjoint $\bar{\partial}^*_{\omega,h}$ with respect to $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\omega,h}$.

Sometimes it is necessary to consider L-valued (p, q)-forms that are locally square integrable on X, rather than (globally) square integrable. The space consisting of all such forms is obviously independent of the choice of the C^∞ metrics ω and h and is simply denoted by $L^2_{\mathrm{loc}}(X, \Lambda^{p,q}T^*X \otimes L)$.

We now record the following a priori inequality due to Ohsawa and Takegoshi [1987] (see also [Ohsawa 1995; Demailly 2000; 2012a]), which plays a fundamental role in the whole L^2 theory of the $\bar\partial$-operator.

Lemma 2.3. *Let (X, ω) be an n-dimensional Kähler manifold and (L, h) be a Hermitian holomorphic line bundle over X. Given two C^∞ functions $\eta, \lambda > 0$ on X, we then have*

$$\big\|\sqrt{\eta+\lambda}\,\bar\partial^*_{\omega,h}u\big\|^2_{\omega,h} + \big\|\sqrt{\eta}\,\bar\partial u\big\|^2_{\omega,h} \geq \big\langle\!\big\langle\big[\eta\sqrt{-1}\Theta_h(L) - \sqrt{-1}\partial\bar\partial\eta - \lambda^{-1}\sqrt{-1}\partial\eta\wedge\bar\partial\eta,\, \Lambda_\omega\big]u, u\big\rangle\!\big\rangle_{\omega,h} \quad (2\text{-}1)$$

for all $u \in \mathcal{D}(X, \Lambda^{n,q}T^*X \otimes L)$, $q \geq 1$, where $\sqrt{-1}\Theta_h(L)$ denotes the Chern curvature form of (L, h) and Λ_ω denotes the pointwise adjoint of the Lefschetz operator $L_\omega = \omega \wedge \cdot\,$.

Inequality (2-1) also applies to the case where (L, h) is a Hermitian holomorphic *vector* bundle over X. Since we only need the case of line bundles in this paper, we restrict ourselves to this particular case.

3. An Ohsawa–Takegoshi-type L^2 existence theorem

In this section we present an Ohsawa–Takegoshi-type L^2 existence theorem for the $\bar\partial$-operator on complete Kähler manifolds, which includes as a special case Demailly's far-reaching generalization of the L^2 existence theorems due to Hörmander and Andreotti–Vesentini; see [Demailly 1982, théorème 5.1] and also Remark 3.3(iii). Such an existence result is more or less known to experts, but we were unable to locate a good reference for its proof and will include a detailed argument for the reader's convenience. (We do not claim any originality here.)

To begin with, we recall the notion of singular Hermitian metrics. Let L be a holomorphic line bundle over a complex manifold X.

Definition 3.1. A (possibly) *singular Hermitian metric* h on L is a measurable Hermitian metric such that, with respect to any smooth local frame e_L of L over an open set $U \subset X$, one has

$$|e_L|^2_h = e^{-\varphi}$$

for some $\varphi \in L^1_{\mathrm{loc}}(U)$.

It is easy to see that, in terms of a given smooth Hermitian metric h_0, every singular Hermitian metric h on L can be written exactly in the form $h = h_0 e^{-\varphi}$, where $\varphi \in L^1_{\mathrm{loc}}(X)$; and the curvature current

$$\sqrt{-1}\Theta_h(L) := -\sqrt{-1}\partial\bar\partial \log h$$

of (L, h) is semipositive exactly when the weight $\varphi = -\log(h/h_0)$ of h (with respect to h_0) satisfies

$$\sqrt{-1}\Theta_{h_0}(L) + \sqrt{-1}\partial\bar\partial\varphi \geq 0 \quad \text{on } X$$

in the sense of currents.

Given a singular Hermitian metric h on L, we can also define $L^2(X, \Lambda^{p,q}T^*X \otimes L; \omega, h)$ and $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\omega,h}$ in the same way as in Section 2.2. We now state and prove the following:

Theorem 3.2. *Let X be an n-dimensional complete Kähler manifold with a (not necessarily complete) Kähler metric ω and L be a holomorphic line bundle over X with a singular Hermitian metric h. Suppose there are bounded C^∞ functions η, $\lambda > 0$ on X such that*

$$\eta\sqrt{-1}\Theta_h(L) - \sqrt{-1}\partial\bar\partial\eta - \lambda^{-1}\sqrt{-1}\partial\eta\wedge\bar\partial\eta \geq \Theta \tag{3-1}$$

in the sense of currents, where Θ is a continuous (strictly) positive $(1,1)$-form on X. Then, for every $\bar\partial$-closed $v \in L^2_{\mathrm{loc}}(X, \Lambda^{n,q}T^*X\otimes L)$, $q \geq 1$, satisfying

$$\int_X \langle [\Theta, \Lambda_\omega]^{-1} v, v\rangle_{\omega,h}\, \mathrm{d}V_\omega < \infty,$$

there exists $u \in L^2(X, \Lambda^{n,q-1}T^*X\otimes L; \omega, h)$ such that $\bar\partial u = v$ and

$$\int_X (\eta+\lambda)^{-1}|u|^2_{\omega,h}\, \mathrm{d}V_\omega \leq \int_X \langle [\Theta, \Lambda_\omega]^{-1} v, v\rangle_{\omega,h}\, \mathrm{d}V_\omega. \tag{3-2}$$

We are mainly interested in the case $q = 1$, but the general statement is needed in the proof.

Proof. First of all, a standard monotonicity argument as in [Demailly 1982] allows us to reduce the theorem to the case when ω is complete, which we shall henceforward assume.

We next make use of the approximation theorem of Demailly [1982, théorème 9.1], according to which there is an increasing sequence $\{h_j\}_{j\in\mathbb{N}}$ of C^∞ Hermitian metrics on L such that

(i) $h_j \nearrow h$ on X as $j\to\infty$, and

(ii) the curvature of each h_j satisfies

$$\eta\sqrt{-1}\Theta_{h_j}(L) - \sqrt{-1}\partial\bar\partial\eta - \lambda^{-1}\sqrt{-1}\partial\eta\wedge\bar\partial\eta \geq \Theta - \delta_j\omega, \tag{3-3}$$

where $\{\delta_j\}_{j\in\mathbb{N}}$ is a decreasing sequence of continuous functions on X that converges to 0 *almost everywhere* on X as $j\to\infty$.

Since (X,ω) is complete, it is easy to construct a C^∞ exhaustion function ρ for X with $|\mathrm{d}\rho|_\omega \leq 1$ on X. Choose also a C^∞ function $\chi : \mathbb{R}\to[0,1]$ such that $\chi|_{(-\infty,1)} = 1$, $\chi|_{(2,\infty)} = 0$ and $|\chi'| \leq 2$ on $\mathbb{R}$, and define

$$\chi_j := \chi\Big(\frac{\rho}{j}\Big), \quad j\in\mathbb{N}.$$

We now proceed to establish several necessary estimates. For notational simplicity, we set

$$L^2(X, \Lambda^{n,q}T^*X\otimes L) = L^2(X, \Lambda^{n,q}T^*X\otimes L; \omega, h),$$
$$L^2(X, \Lambda^{n,q}T^*X\otimes L)_k = L^2(X, \Lambda^{n,q}T^*X\otimes L; \omega, h_k)$$

and

$$\langle\!\langle\, \cdot, \cdot\,\rangle\!\rangle = \langle\!\langle\, \cdot, \cdot\,\rangle\!\rangle_{\omega,h}, \quad \langle\!\langle\, \cdot, \cdot\,\rangle\!\rangle_k = \langle\!\langle\, \cdot, \cdot\,\rangle\!\rangle_{\omega,h_k}$$

as well as $\bar\partial_k^* = \bar\partial^*_{\omega,h_k}$ for $k\in\mathbb{N}$. Similar simplifications apply to other notation, and we will omit subscripts when no confusion arises. Given v as in the theorem, since the Hermitian operator $[\Theta,\Lambda]$

on $\Lambda^{n,q}T^*X\otimes L$ is positive definite (as follows from the positivity of Θ; see, e.g., [Demailly 2012b, Chapter VI, Proposition 5.8]), we have

$$\begin{aligned}|\langle\!\langle \chi_j v, w\rangle\!\rangle_k|^2 &\le \langle\!\langle [\Theta,\Lambda]^{-1}v, v\rangle\!\rangle_k \langle\!\langle [\Theta,\Lambda](\chi_j w), \chi_j w\rangle\!\rangle_k \\ &\le \langle\!\langle [\Theta,\Lambda]^{-1}v, v\rangle\!\rangle \langle\!\langle [\Theta,\Lambda](\chi_j w), \chi_j w\rangle\!\rangle_k \end{aligned} \tag{3-4}$$

for all $w\in \operatorname{Dom}\bar\partial\cap\operatorname{Dom}\bar\partial_k^*$ and $j,k\in\mathbb{N}$, in view of the Cauchy–Schwarz inequality and the fact that $h_k\le h$. Since $\chi_j w\in\operatorname{Dom}\bar\partial\cap\operatorname{Dom}\bar\partial_k^*$ (as is easily checked by definition), Ohsawa–Takegoshi's a priori inequality (2-1) together with (3-3) and the density theorem of Andreotti and Vesentini (see, e.g., [Demailly 2012b, Chapter VIII]) implies

$$\begin{aligned}\big\|\sqrt{\eta+\lambda}\bar\partial_k^*(\chi_j w)\big\|_k^2+\big\|\sqrt{\eta}\bar\partial(\chi_j w)\big\|_k^2 &\ge \langle\!\langle [\Theta-\delta_k\omega,\Lambda](\chi_j w),\chi_j w\rangle\!\rangle_k \\ &= \langle\!\langle [\Theta,\Lambda](\chi_j w),\chi_j w\rangle\!\rangle_k - q\|\chi_j\delta_k^{1/2}w\|_k^2,\end{aligned} \tag{3-5}$$

where the last equality follows from the fact that $[\omega,\Lambda]=q\,\mathrm{Id}$ on $\Lambda^{n,q}T^*X$. Observe also that

$$\bar\partial(\chi_j w)=\chi_j\bar\partial w+\bar\partial\chi_j\wedge w$$

and

$$\bar\partial_k^*(\chi_j w)=\chi_j\bar\partial_k^* w-(\operatorname{grad}\chi_j)^{0,1}\lrcorner\, w,$$

where $\lrcorner$ denotes the contraction operator and $(\operatorname{grad}\chi_j)^{0,1}$ denotes the $(0,1)$-part of the gradient of χ_j under the metric corresponding to ω. Substituting these two identities into (3-5) and applying the Cauchy–Schwarz inequality, we get

$$\begin{aligned}\langle\!\langle [\Theta,\Lambda](\chi_j w),\chi_j w\rangle\!\rangle_k &\le q\|\chi_j\delta_k^{1/2}w\|_k^2+\Big(1+\frac{1}{j}\Big)\big(\big\|\sqrt{\eta+\lambda}\bar\partial_k^* w\big\|_k^2+\big\|\sqrt{\eta}\bar\partial w\big\|_k^2\big) \\ &\quad+(1+j)\big(\big\|\sqrt{\eta+\lambda}(\operatorname{grad}\chi_j)^{0,1}\lrcorner\, w\big\|_k^2+\big\|\sqrt{\eta}\bar\partial\chi_j\wedge w\big\|_k^2\big) \\ &\le \Big(1+\frac{1}{j}\Big)\Big(\big\|\sqrt{\eta+\lambda}\bar\partial_k^* w\big\|_k^2+\big\|\sqrt{\eta}\bar\partial w\big\|_k^2+\frac{C}{j}\|w\|_k^2\Big)+q\|\chi_j\delta_k^{1/2}w\|_k^2\end{aligned} \tag{3-6}$$

with $C=2\sup_X(2\eta+\lambda)$, where in the last inequality we have used the facts that

$$|\bar\partial\chi_j\wedge w|_k\le|\bar\partial\chi_j||w|_k\le\frac{\sqrt{2}}{j}|w|_k$$

and

$$|(\operatorname{grad}\chi_j)^{0,1}\lrcorner\, w|_k\le|\bar\partial\chi_j||w|_k\le\frac{\sqrt{2}}{j}|w|_k.$$

Now combining (3-4) with (3-6) we conclude that for $j,k\in\mathbb{N}$ the estimate

$$|\langle\!\langle \chi_j v,w\rangle\!\rangle_k|^2\le C_j\Big(\big\|\sqrt{\eta+\lambda}\bar\partial_k^* w\big\|_k^2+\big\|\sqrt{\eta}\bar\partial w\big\|_k^2+q\|\chi_j\delta_k^{1/2}w\|_k^2+\frac{C}{j}\|w\|_k^2\Big) \tag{3-7}$$

holds for all $w\in\operatorname{Dom}\bar\partial\cap\operatorname{Dom}\bar\partial_k^*$, with

$$C_j:=\Big(1+\frac{1}{j}\Big)\langle\!\langle [\Theta,\Lambda]^{-1}v,v\rangle\!\rangle.$$

This estimate enables us to prove the theorem by descending induction on q, as we shall now explain. The idea is due to Demailly [1982], as mentioned in the Introduction.

When $q = n$, (3-7) reads

$$|\langle\!\langle \chi_j v, w\rangle\!\rangle_k|^2 \le C_j\Big(\big\|\sqrt{\eta+\lambda}\,\bar\partial_k^* w\big\|_k^2 + q\|\chi_j\delta_k^{1/2} w\|_k^2 + \frac{C}{j}\|w\|_k^2\Big)$$

for all $w \in \operatorname{Dom}\bar\partial_k^*$. This shows that the antilinear functional $w \mapsto \langle\!\langle \chi_j v, w\rangle\!\rangle_k$ is continuous on the Hilbert space

$$\Big(\operatorname{Dom}\bar\partial_k^*, \big\|\sqrt{\eta+\lambda}\,\bar\partial_k^* \cdot\big\|_k^2 + q\|\chi_j\delta_k^{1/2}\cdot\|_k^2 + \frac{C}{j}\|\cdot\|_k^2\Big)$$

with norm at most $\sqrt{C_j}$. Hence, by the Riesz representation theorem, there exist

$$u_{jk} \in L^2(X, \Lambda^{n,n-1}T^*X \otimes L)_k \quad \text{and} \quad v_{jk}, w_{jk} \in L^2(X, \Lambda^{n,n}T^*X \otimes L)_k$$

such that

$$\|(\eta+\lambda)^{-1/2}u_{jk}\|_k^2 + \|v_{jk}\|_k^2 + \|w_{jk}\|_k^2 \le C_j \tag{3-8}$$

and

$$\langle\!\langle u_{jk}, \bar\partial_k^* w\rangle\!\rangle_k + \sqrt{q}\,\langle\!\langle v_{jk}, \chi_j\delta_k^{1/2} w\rangle\!\rangle_k + \sqrt{\frac{C}{j}}\,\langle\!\langle w_{jk}, w\rangle\!\rangle_k = \langle\!\langle \chi_j v, w\rangle\!\rangle_k$$

for all $w \in \operatorname{Dom}\bar\partial_k^*$. This means that $u_{jk} \in \operatorname{Dom}\bar\partial_k^{**} = \operatorname{Dom}\bar\partial$ and

$$\bar\partial u_{jk} + \sqrt{q}\,\chi_j\delta_k^{1/2} v_{jk} + \sqrt{\frac{C}{j}}\,w_{jk} = \chi_j v$$

for all $j, k \in \mathbb{N}$. Moreover, since $\{h_k\}_{k\in\mathbb{N}}$ is increasing, inequality (3-8) and the Banach–Alaoglu theorem imply that $\{u_{jk}\}_{k\in\mathbb{N}}$ (resp. $\{w_{jk}\}_{k\in\mathbb{N}}$) has a subsequence that converges weakly to some u_j (resp. w_j) in $L^2_{\mathrm{loc}}(X, \Lambda^{n,n-1}T^*X \otimes L)$ (resp. $L^2_{\mathrm{loc}}(X, \Lambda^{n,n}T^*X \otimes L)$). Also since $\{\delta_k\}_{k\in\mathbb{N}}$ decreases to 0 almost everywhere on X, it follows readily from (3-8), the Cauchy–Schwarz inequality and the dominated convergence theorem that $\{\chi_j\delta_k^{1/2}v_{jk}\}_{k\in\mathbb{N}}$ converges to 0 in $L^1_{\mathrm{loc}}(X, \Lambda^{n,n}T^*X \otimes L)$. Consequently

$$\bar\partial u_j + \sqrt{\frac{C}{j}}\,w_j = \chi_j v$$

and

$$\begin{aligned}\int_\Omega ((\eta+\lambda)^{-1}|u_j|_l^2 + |w_j|_l^2)\,\mathrm{d}V &\le \limsup_{k\to\infty}\int_\Omega ((\eta+\lambda)^{-1}|u_{jk}|_l^2 + |w_{jk}|_l^2)\,\mathrm{d}V\\ &\le \limsup_{k\to\infty}\big(\|(\eta+\lambda)^{-1/2}u_{jk}\|_k^2 + \|w_{jk}\|_k^2\big) \le C_j\end{aligned}$$

for all $j, l \in \mathbb{N}$ and open sets $\Omega \Subset X$. Letting first $l \to \infty$ and then $\Omega \to X$ we obtain

$$\|(\eta+\lambda)^{-1/2}u_j\|^2 + \|w_j\|^2 \le C_j$$

for all $j \in \mathbb{N}$, in view of the monotone convergence theorem. Taking a weak limit point u of $\{u_j\}_{j\in\mathbb{N}}$ in $L^2_{\mathrm{loc}}(X, \Lambda^{n,n-1}T^*X \otimes L)$, we then conclude that $\bar\partial u = v$ and

$$\|(\eta+\lambda)^{-1/2}u\|^2 \le \lim_{j\to\infty} C_j = \langle\!\langle [\Theta, \Lambda]^{-1}v, v\rangle\!\rangle.$$

This completes the proof of the theorem in the case $q = n$.

Now assume that $1 \le q \le n-1$ and that the conclusion of the theorem is true for all $\bar\partial$-closed $(n, q+1)$-forms with the prescribed integrability. We proceed to show that the conclusion is also true for all (n, q)-forms with the same property. Given such a form v and $j \in \mathbb{N}$, consider the $\bar\partial$-closed $(n, q+1)$-form $\bar\partial(\chi_j v) = \bar\partial\chi_j \wedge v$ which satisfies

$$\langle\!\langle [\Theta, \Lambda]^{-1}\bar\partial(\chi_j v), \bar\partial(\chi_j v)\rangle\!\rangle \le \sup_X |\bar\partial\chi_j|^2 \langle\!\langle [\Theta, \Lambda]^{-1} v, v\rangle\!\rangle \le \frac{2}{j^2}\langle\!\langle [\Theta, \Lambda]^{-1} v, v\rangle\!\rangle < \infty,$$

where the first inequality follows from [Demailly 1982, lemme 3.2 (3.5)]. By this estimate and the induction hypothesis, there exists $v_j \in L^2(X, \Lambda^{n,q}T^*X \otimes L)$ such that $\bar\partial v_j = \bar\partial(\chi_j v)$ and

$$\|(\eta+\lambda)^{-1/2} v_j\|^2 \le \frac{2}{j^2}\langle\!\langle [\Theta, \Lambda]^{-1} v, v\rangle\!\rangle. \tag{3-9}$$

In particular, $\chi_j v - v_j \in \operatorname{Ker}\bar\partial$ and

$$v_j \in L^2(X, \Lambda^{n,q}T^*X \otimes L) \subset L^2(X, \Lambda^{n,q}T^*X \otimes L)_k.$$

On the other hand, note that every $w \in \operatorname{Dom}\bar\partial_k^*$ admits an orthogonal decomposition $w = w_1 + w_2$, where

$$w_1 \in \operatorname{Ker}\bar\partial \quad \text{and} \quad w_2 \in (\operatorname{Ker}\bar\partial)^\perp \subset \operatorname{Ker}\bar\partial_k^*.$$

It follows that

$$\langle\!\langle \chi_j v - v_j, w\rangle\!\rangle_k = \langle\!\langle \chi_j v - v_j, w_1\rangle\!\rangle_k,$$

and hence

$$\langle\!\langle \chi_j v, w\rangle\!\rangle_k = \langle\!\langle \chi_j v, w_1\rangle\!\rangle_k + \langle\!\langle v_j, w_2\rangle\!\rangle_k.$$

Applying the Cauchy–Schwarz inequality to this and using (3-9), we get

$$\begin{aligned} |\langle\!\langle \chi_j v, w\rangle\!\rangle_k|^2 &\le \Big(1+\frac{1}{j}\Big)|\langle\!\langle \chi_j v, w_1\rangle\!\rangle_k|^2 + (1+j)\|v_j\|^2\|w_2\|_k^2 \\ &\le \Big(1+\frac{1}{j}\Big)\Big(|\langle\!\langle \chi_j v, w_1\rangle\!\rangle_k|^2 + \frac{2}{j}\sup_X(\eta+\lambda)\langle\!\langle [\Theta, \Lambda]^{-1} v, v\rangle\!\rangle\|w\|_k^2\Big). \end{aligned}$$

In addition, since $w_1 \in \operatorname{Ker}\bar\partial \cap \operatorname{Dom}\bar\partial_k^*$ and $w_2 \in \operatorname{Ker}\bar\partial_k^*$, applying (3-7) to w_1 yields

$$|\langle\!\langle \chi_j v, w_1\rangle\!\rangle_k|^2 \le C_j\Big(\big\|\sqrt{\eta+\lambda}\,\bar\partial_k^* w\big\|_k^2 + q\|\chi_j\delta_k^{1/2} w_1\|_k^2 + \frac{C}{j}\|w\|_k^2\Big).$$

Substituting this estimate into the preceding one we arrive at

$$|\langle\!\langle \chi_j v, w\rangle\!\rangle_k|^2 \le \Big(1+\frac{1}{j}\Big)C_j\Big(\big\|\sqrt{\eta+\lambda}\,\bar\partial_k^* w\big\|_k^2 + q\|\chi_j\delta_k^{1/2} P_k w\|_k^2 + \frac{2C}{j}\|w\|_k^2\Big)$$

for all $w \in \operatorname{Dom}\bar\partial_k^*$, where P_k denotes the orthogonal projection from $L^2(X, \Lambda^{n,q}T^*X \otimes L)_k$ onto $\operatorname{Ker}\bar\partial$. Then the Riesz representation theorem implies that there exist

$$u_{jk} \in L^2(X, \Lambda^{n,q-1}T^*X \otimes L)_k \quad \text{and} \quad v_{jk}, w_{jk} \in L^2(X, \Lambda^{n,q}T^*X \otimes L)_k$$

such that

$$\|(\eta+\lambda)^{-1/2} u_{jk}\|_k^2 + \|v_{jk}\|_k^2 + \|w_{jk}\|_k^2 \le \Big(1+\frac{1}{j}\Big)C_j \tag{3-10}$$

and

$$\langle\!\langle u_{jk}, \bar\partial_k^* w\rangle\!\rangle_k + \sqrt{q}\langle\!\langle v_{jk}, \chi_j\delta_k^{1/2}P_k w\rangle\!\rangle_k + \sqrt{\frac{2C}{j}}\langle\!\langle w_{jk}, w\rangle\!\rangle_k = \langle\!\langle \chi_j v, w\rangle\!\rangle_k$$

for all $w \in \operatorname{Dom}\bar\partial_k^*$. It follows that

$$\bar\partial u_{jk} + \sqrt{q}P_k(\chi_j\delta_k^{1/2}v_{jk}) + \sqrt{\frac{2C}{j}}\,w_{jk} = \chi_j v$$

for all $j, k \in \mathbb{N}$. To complete the proof, it remains to show that, for every $j \in \mathbb{N}$,

$$f_{jk} := P_k(\chi_j\delta_k^{1/2}v_{jk}) \to 0 \quad \text{weakly in } L^2_{\mathrm{loc}}(X, \Lambda^{n,q}T^*X \otimes L) \tag{3-11}$$

as $k \to \infty$. Once this is established, the same limiting argument as in the latter half of the preceding paragraph yields the existence of a solution u to equation $\bar\partial u = v$ together with the desired estimate

$$\|(\eta+\lambda)^{-1/2}u\|^2 \le \langle\!\langle [\Theta, \Lambda]^{-1}v, v\rangle\!\rangle,$$

thus completing the induction procedure (and hence the proof of the theorem).

To prove (3-11), we argue as follows. First observe that, by (3-10),

$$\|f_{jk}\|_k \le \|\chi_j\delta_k^{1/2}v_{jk}\|_k \le \sup_X(\chi_j\delta_1^{1/2})\|v_{jk}\|_k \le \mathrm{const}_j \tag{3-12}$$

for all $k \in \mathbb{N}$, where const_j denotes a constant depending only on j. This observation is crucial in the following argument and will be used repeatedly. For the moment, this bound, together with the Banach–Alaoglu theorem, implies that the set of weak limit points of $\{f_{jk}\}_{k\in\mathbb{N}}$ is nonempty. Thus it suffices to show that 0 is the only element of this set. To establish the validity of the latter, we may assume without loss of generality that the sequence $\{f_{jk}\}_{k\in\mathbb{N}}$ itself is weakly convergent in $L^2_{\mathrm{loc}}(X, \Lambda^{n,q}T^*X\otimes L)$, with limit f_j, say. Let H_k (resp. H) denote the only positive definite Hermitian operator on (L, h_1) that satisfies $|H_k\zeta|_1 = |\zeta|_k$ (resp. $|H\zeta|_1 = |\zeta|$) for all $\zeta \in L$. Then

$$\mathrm{Id}_L \le H_k \le H_{k+1}$$

for all $k \in \mathbb{N}$, and $\{H_k\}_{k\in\mathbb{N}}$ converges to H almost everywhere on X. After extending H_k and H naturally to $\Lambda^{n,q}T^*X\otimes L$, we claim that

$$H_k f_{jk} \to Hf_j \quad \text{weakly in } L^2_{\mathrm{loc}}(X, \Lambda^{n,q}T^*X\otimes L)$$

as $k \to \infty$. To show this, it suffices to check that Hf_j is the only weak limit point of $\{H_k f_{jk}\}_{k\in\mathbb{N}}$. Again by (3-12) and the Banach–Alaoglu theorem, we may assume that the sequence $\{H_k f_{jk}\}_{k\in\mathbb{N}}$ itself is weakly convergent in $L^2_{\mathrm{loc}}(X, \Lambda^{n,q}T^*X\otimes L)$, with limit g_j, say. Observe that for every compactly supported $\varphi \in L^2(X, \Lambda^{n,q}T^*X\otimes L)_1$

$$\left|\langle\!\langle H_k f_{jk}, (H_k^{-1} - H^{-1})\varphi\rangle\!\rangle_1\right| \le \|f_{jk}\|_k\|(H_k^{-1}-H^{-1})\varphi\|_1 \le \mathrm{const}_j\|(H_k^{-1}-H^{-1})\varphi\|_1 \to 0$$

as $k \to \infty$, in view of (3-12), the Cauchy–Schwarz inequality and the dominated convergence theorem. Thus

$$\langle\!\langle f_j, \varphi\rangle\!\rangle_1 = \lim_{k\to\infty}\langle\!\langle H_k f_{jk}, H_k^{-1}\varphi\rangle\!\rangle_1 = \lim_{k\to\infty}\langle\!\langle H_k f_{jk}, H^{-1}\varphi\rangle\!\rangle_1 = \langle\!\langle g_j, H^{-1}\varphi\rangle\!\rangle_1,$$

and hence $g_j = Hf_j$ and the claim follows. Moreover, by (3-12) and the definition of H_k, we see that $Hf_j \in L^2(X, \Lambda^{n,q}T^*X \otimes L)_1$. Now using (3-12) again and a standard $\frac{1}{3}\varepsilon$-argument, we further conclude that

$$H_k f_{jk} \to Hf_j \quad \text{weakly in } L^2(X, \Lambda^{n,q}T^*X \otimes L)_1$$

as $k \to \infty$. In particular,

$$\|f_j\|^2 = \|Hf_j\|_1^2 = \lim_{k\to\infty} \langle\!\langle H_k f_{jk}, Hf_j \rangle\!\rangle_1 = \lim_{k\to\infty} \langle\!\langle H_k f_{jk}, H_k f_j \rangle\!\rangle_1,$$

where the last equality follows from the fact that

$$\big|\langle\!\langle H_k f_{jk}, (H_k - H) f_j \rangle\!\rangle_1\big| \le \|f_{jk}\|_k \|(H_k - H) f_j\|_1 \le \mathrm{const}_j \|(H_k - H) f_j\|_1 \to 0$$

as $k \to \infty$. Equivalently,

$$\|f_j\|^2 = \lim_{k\to\infty} \langle\!\langle f_{jk}, f_j \rangle\!\rangle_k.$$

On the other hand, since $\bar\partial f_{jk} = 0$ for all $k \in \mathbb{N}$, it follows that $\bar\partial f_j = 0$ as well; hence

$$\langle\!\langle f_{jk}, f_j \rangle\!\rangle_k = \langle\!\langle \chi_j \delta_k^{1/2} v_{jk}, f_j \rangle\!\rangle_k.$$

Moreover

$$|\langle\!\langle \chi_j \delta_k^{1/2} v_{jk}, f_j \rangle\!\rangle_k| \le \|v_{jk}\|_k \|\chi_j f_j \delta_k^{1/2}\|_k \le \mathrm{const}_j \|\chi_j f_j \delta_k^{1/2}\| \to 0$$

as $k \to \infty$, in view of (3-10) and the dominated convergence theorem. Consequently $\|f_j\| = 0$, i.e., $f_j = 0$ almost everywhere on X. This establishes the desired weak convergence of $\{f_{jk}\}_{k\in\mathbb{N}}$ and hence the theorem. □

Remark 3.3. A few remarks are in order:

(i) When $q = 1$, the estimate in (3-2) is independent of the choice of the Kähler metric ω. Indeed, Demailly [1982] and Ohsawa [1995] have respectively explained this in detail for the right-hand side of (3-2); for the left-hand side, the reason is simpler, as one readily sees that $|\cdot|^2_{\omega,h}\, \mathrm{d}V_\omega$ can be expressed intrinsically in terms of $\{\cdot, \cdot\}_h$ (to be recalled in Section 4) as follows:

$$|u|^2_{\omega,h}\, \mathrm{d}V_\omega = \sqrt{-1}^{\,n^2} \{u, u\}_h. \tag{3-13}$$

Moreover, when $q = 1$, if the form v in Theorem 3.2 is smooth then every solution u of the equation $\bar\partial u = v$ is smooth as well. This follows immediately from the elliptic regularity for the $\bar\partial$-operator on $(n, 0)$-forms.

(ii) Demailly [2000] (see also [Ohsawa 1995; Demailly 2012a]) proved using Ohsawa–Takegoshi's a priori inequality (2-1) an abstract L^2 existence theorem for the $\bar\partial$-operator under the assumption that the metric h on L is *smooth* and v satisfies the additional L^2 integrability

$$\int_X |v|^2_{\omega,h}\, \mathrm{d}V_\omega < \infty.$$

Compared with this result, the point of Theorem 3.2 is that it allows the metric h to be *singular*. This is crucial in proving Theorem 1.1 (and hence Theorem 1.2).

(iii) As the preceding proof indicates, the curvature condition (3-1) can be weakened to a condition analogous to that in [Demailly 1982, théorème 5.1], and L may also be replaced by a higher-rank vector bundle. To avoid technicality, we restrict ourselves to the present case. As we shall see in the next section, Theorem 3.2 is sufficient for the purpose of the present paper.

(iv) In the special case when X is a bounded complete Kähler domain in $\mathbb{C}^n$ and

$$(L, h) = (X \times \mathbb{C}, e^{-\varphi})$$

with $\varphi \in \mathrm{Psh}(X)$, Theorem 3.2 was first proved by Chen, Wu and Wang [Chen et al. 2015][2] using a different method, namely by first solving the twisted Laplace equation with Dirichlet boundary condition on a sequence of smooth subdomains increasing to X and then using these solutions to approximate the desired solution of the twisted $\bar{\partial}$-equation on X with L^2 estimate. Although quite interesting, this approach does not seem to work in the manifold case, since — as far as the author can see — it requires that the *strict positivity* of the curvature term on the left-hand side of (3-1) be preserved during the regularization process of the singular Hermitian metric h; but in general the loss of positivity (i.e., the error term $\delta_j\omega$ in (3-3)) arising from any regularization sequence $\{h_j\}_{j\in\mathbb{N}}$ is inevitable and difficult to control even locally.

4. Ohsawa–Takegoshi L^2 extension theorems: the complete Kähler case

Let Ω be a domain in a complex manifold and L be a holomorphic line bundle over Ω. Denote by $C^\infty(\Omega, \Lambda^{p,q}T^*\Omega \otimes L)$ the set of all L-valued C^∞ (p, q)-forms on Ω. Given a C^∞ Hermitian metric h on L, there is a canonical sesquilinear map

$$\{\cdot, \cdot\}_h : C^\infty(\Omega, \Lambda^{p,q}T^*\Omega \otimes L) \times C^\infty(\Omega, \Lambda^{r,s}T^*\Omega \otimes L) \to C^\infty(\Omega, \Lambda^{p+s,q+r}T^*\Omega)$$

that combines the wedge product of scalar-valued forms with the metric h on L. More explicitly, for any given

$$u \in C^\infty(\Omega, \Lambda^{p,q}T^*\Omega \otimes L) \quad \text{and} \quad v \in C^\infty(\Omega, \Lambda^{r,s}T^*\Omega \otimes L),$$

if we write

$$u|_U = f \otimes \zeta \quad \text{and} \quad v|_U = g \otimes \zeta$$

in terms of a frame ζ for L over an open subset U of Ω, then

$$\{u, v\}_h|_U = f \wedge \bar{g}|\zeta|_h^2.$$

Clearly the definition of $\{\cdot, \cdot\}_h$ is pointwise and hence makes sense for vectors $u \in \Lambda^{p,q}T^*_{x_0}\Omega \otimes L_{x_0}$ and $v \in \Lambda^{r,s}T^*_{x_0}\Omega \otimes L_{x_0}$, $x_0 \in \Omega$, even when h is merely a singular Hermitian metric on L.

With the above necessary notion, we can now turn to the main goal of this section, namely proving Theorem 1.1, whose statement we recall for ease of reference.

[2] They only considered the case when v is smooth and of type $(n, 1)$, but their method actually works equally well for not necessarily smooth (n, q)-forms v, at least under the additional integrability assumption $\int_X |v|^2_\omega e^{-\varphi}\, \mathrm{d}V_\omega < \infty$.

Theorem 4.1. *Let X be an n-dimensional Stein manifold with a continuous volume form $\mathrm{d}V_X$ and Ω be a relatively compact complete Kähler domain in X. Suppose L is a holomorphic line bundle over Ω and h is a singular Hermitian metric on L with semipositive curvature current. Then, for every $x_0 \in \Omega$ and $s_{x_0} \in \Lambda^n T^*_{x_0}\Omega \otimes L_{x_0}$, there is an L-valued holomorphic n-form s on Ω such that $s(x_0) = s_{x_0}$ and*

$$\int_\Omega \sqrt{-1}^{n^2}\{s, s\}_h \le \mathrm{const}_{n,\Omega,\mathrm{d}V_X} \frac{\sqrt{-1}^{n^2}\{s_{x_0}, s_{x_0}\}_h}{\mathrm{d}V_X(x_0)}$$

provided the right-hand side is finite.

Here and henceforward we use $\mathrm{const}_{n,\Omega,\mathrm{d}V_X}$ to denote a generic constant that depends only on n, Ω and $\mathrm{d}V_X$ and may be different in different appearances. We shall also use similar notation (e.g., const_n) with obvious meaning in the context.

We now prove the above theorem. Since $\overline{\Omega}$ is compact, we can assign to every point $x \in \Omega$ a holomorphic local coordinate chart $z_x = (z_x^1, \ldots, z_x^n)$ of X defined on a neighborhood of the closure of $B_x := \{|z_x| < 1\}$ such that $z_x(x) = 0$ and

$$\frac{\sqrt{-1}^{n^2}\,\mathrm{d}z_x \wedge \mathrm{d}\bar{z}_x}{\mathrm{d}V_X} \ge \mathrm{const}_{n,\Omega,\mathrm{d}V_X} \quad \text{on } B_x, \tag{4-1}$$

where $\mathrm{d}z_x := \mathrm{d}z_x^1 \wedge \cdots \wedge \mathrm{d}z_x^n$. Set

$$\varepsilon B_x = \{|z_x| < \varepsilon\}$$

for $x \in \Omega$ and $\varepsilon \in (0, 1)$. We claim the following:

Lemma 4.2. *Given $x_0 \in \Omega$, $f_{x_0} \in \Lambda^n T^*_{x_0}\Omega$ and $\varepsilon \in (0, 1)$, $\varphi \in \mathrm{Psh}(\varepsilon B_{x_0})$, there is a holomorphic n-form f_ε on εB_{x_0} such that $f_\varepsilon(x_0) = f_{x_0}$ and*

$$\varepsilon^{-2n}\int_{\varepsilon B_{x_0}} \sqrt{-1}^{n^2} f_\varepsilon \wedge \bar{f}_\varepsilon e^{-\varphi} \le \mathrm{const}_{n,\Omega,\mathrm{d}V_X} \frac{\sqrt{-1}^{n^2} f_{x_0} \wedge \bar{f}_{x_0}}{\mathrm{d}V_X(x_0)} e^{-\varphi(x_0)}.$$

Proof. Write $f_{x_0} = c\,\mathrm{d}z_{x_0}$ with $c \in \mathbb{C}$. As, e.g., in [Chen et al. 2015], an application of the Ohsawa–Takegoshi extension theorem shows that for every $\varepsilon \in (0, 1)$ there is a holomorphic n-form f_ε on εB_{x_0} such that $f_\varepsilon(x_0) = f_{x_0}$ and

$$\varepsilon^{-2n}\int_{\varepsilon B_{x_0}} \sqrt{-1}^{n^2} f_\varepsilon \wedge \bar{f}_\varepsilon e^{-\varphi} \le \mathrm{const}_n |c|^2 e^{-\varphi(x_0)}.$$

Combining this with inequality (4-1) yields the desired result immediately. □

We now proceed to prove Theorem 4.1. With Theorem 3.2 in hand, we do this by adapting the argument from the proof of [Chen et al. 2015, Theorem 1.3] to the present context.

Without loss of generality we may assume that $\mathrm{d}V_X = \omega^n/n!$, with ω a Kähler form on X. Fix an arbitrary point $x_0 \in \Omega$. Let us write $(B, z) = (B_{x_0}, z_{x_0})$ for the sake of simplicity, and pick a function $\rho \in C^\infty(X, \mathbb{R})$ that is positive, strictly log-psh on $X \setminus \{x_0\}$, and satisfies $\rho < e^{-1}$ on $\overline{\Omega}$ and $\rho = |z|^2$ on a small neighborhood of x_0. Similarly as in [Ohsawa 1995; Chen 2011; Chen et al. 2015], we set

$$\psi_\varepsilon = \log(\rho + \varepsilon^2) \quad \text{and} \quad \eta_\varepsilon = -\psi_\varepsilon + \log(-\psi_\varepsilon)$$

for $0 < \varepsilon \le \varepsilon_0$, where $\varepsilon_0 > 0$ is chosen small enough that $\psi_{\varepsilon_0} < -1$ on $\overline{\Omega}$ and the line bundle L is trivial over $\varepsilon_0 B$. A straightforward calculation shows

$$-\partial\eta_\varepsilon = (1-\psi_\varepsilon^{-1})\partial\psi_\varepsilon \quad \text{and} \quad -\partial\bar\partial\eta_\varepsilon = (1-\psi_\varepsilon^{-1})\partial\bar\partial\psi_\varepsilon + \psi_\varepsilon^{-2}\partial\psi_\varepsilon\wedge\bar\partial\psi_\varepsilon,$$

whence

$$-\sqrt{-1}\partial\bar\partial\eta_\varepsilon - \lambda_\varepsilon^{-1}\sqrt{-1}\partial\eta_\varepsilon\wedge\bar\partial\eta_\varepsilon \ge \sqrt{-1}\partial\bar\partial\psi_\varepsilon \quad \text{on } \Omega,$$

where $\lambda_\varepsilon = (1-\psi_\varepsilon)^2$.

Now let $s_{x_0} \in \Lambda^n T^*_{x_0}\Omega \otimes L_{x_0}$ be given such that

$$\frac{\sqrt{-1}^{n^2}\{s_{x_0}, s_{x_0}\}_h}{\omega^n(x_0)} < \infty.$$

Take a holomorphic frame ζ for L over $\varepsilon_0 B$ and write s_{x_0} as $s_{x_0} = f_{x_0}\otimes\zeta_{x_0}$. Let $\{f_\varepsilon\}_{0<\varepsilon\le\varepsilon_0}$ be a family of holomorphic n-forms as in Lemma 4.2 corresponding to the weight function

$$\varphi = -\log|\zeta|_h^2,$$

which is psh on $\varepsilon_0 B$ because of the semipositivity of $\sqrt{-1}\Theta_h(L)$. Consider

$$\Theta_\varepsilon = \sqrt{-1}\partial\bar\partial(\rho+\psi_\varepsilon) \quad \text{and} \quad v_\varepsilon = \bar\partial\left(\chi\left(\frac{|z|^2}{\varepsilon^2}\right)f_\varepsilon\right)\otimes\zeta,$$

where χ is a C^∞ function on $\mathbb{R}$ satisfying $\chi|_{(-\infty,1/4)} = 1$, $\chi|_{(3/4,\infty)} = 0$ and $|\chi'| \le 4$. Observe that

$$\Theta_\varepsilon \ge \sqrt{-1}\partial\bar\partial\psi_\varepsilon = \sqrt{-1}\partial\bar\partial\log(|z|^2+\varepsilon^2) \ge \frac{\varepsilon^2}{(|z|^2+\varepsilon^2)^2}\sqrt{-1}\partial\bar\partial|z|^2$$

on a neighborhood of x_0. Using this and [Demailly 1982, lemme 3.2] (see also the paragraph following Proposition 1.5 in [Ohsawa 1995]), we can estimate

$$\begin{aligned}
\langle\!\langle [\Theta_\varepsilon, \Lambda_\omega]^{-1} v_\varepsilon, v_\varepsilon\rangle\!\rangle_{\omega, he^{-(\rho+n\log\rho)}} &= \int_{\operatorname{supp} v_\varepsilon} \langle [\Theta_\varepsilon, \Lambda_{\sqrt{-1}\partial\bar\partial|z|^2}]^{-1} v_\varepsilon, v_\varepsilon\rangle_{\sqrt{-1}\partial\bar\partial|z|^2, h} e^{-(\rho+n\log\rho)} \frac{(\sqrt{-1}\partial\bar\partial|z|^2)^n}{n!} \\
&\le \int_{\operatorname{supp} v_\varepsilon} \frac{(|z|^2+\varepsilon^2)^2}{\varepsilon^2} |v_\varepsilon|^2_{\sqrt{-1}\partial\bar\partial|z|^2, h} e^{-(\rho+n\log\rho)} \frac{(\sqrt{-1}\partial\bar\partial|z|^2)^n}{n!} \\
&\le \|\chi'\|^2_{L^\infty(\mathbb{R})} \int_{\frac{\sqrt{3}\varepsilon}{2}B\setminus\frac{\varepsilon}{2}B} \frac{|z|^2}{\varepsilon^4}\frac{(|z|^2+\varepsilon^2)^2}{\varepsilon^2}\sqrt{-1}^{n^2} f_\varepsilon\wedge\bar f_\varepsilon |z|^{-2n} e^{-\varphi-|z|^2} \\
&\le \text{const}_n \varepsilon^{-2n} \int_{\varepsilon B} \sqrt{-1}^{n^2} f_\varepsilon\wedge\bar f_\varepsilon e^{-\varphi} \\
&\le \text{const}_{n,\Omega,\omega} \frac{\sqrt{-1}^{n^2} f_{x_0}\wedge\bar f_{x_0}}{\omega^n(x_0)} e^{-\varphi(x_0)} \\
&= \text{const}_{n,\Omega,\omega} \frac{\sqrt{-1}^{n^2}\{s_{x_0}, s_{x_0}\}_h}{\omega^n(x_0)}
\end{aligned}$$

for all sufficiently small $\varepsilon > 0$. In this way, $(\Omega, \omega, L, he^{-(\rho+n\log\rho)}, \eta_\varepsilon, \lambda_\varepsilon, \Theta_\varepsilon, v_\varepsilon)$ fits into the framework of Theorem 3.2. Therefore, there is a solution u_ε to equation $\bar\partial u = v_\varepsilon$ with the estimate

$$\int_\Omega \frac{\sqrt{-1}^{n^2}\{u_\varepsilon, u_\varepsilon\}_h}{\eta_\varepsilon + \lambda_\varepsilon} e^{-(\rho+n\log\rho)} \le \mathrm{const}_{n,\Omega,\omega} \frac{\sqrt{-1}^{n^2}\{s_{x_0}, s_{x_0}\}_h}{\omega^n(x_0)}. \tag{4-2}$$

Here we have used the identity in (3-13). Define

$$s_\varepsilon := \chi\left(\frac{|z|^2}{\varepsilon^2}\right) f_\varepsilon \otimes \zeta - u_\varepsilon.$$

Then s_ε is an L-valued holomorphic n-form on Ω. Moreover, since u_ε is holomorphic on a small neighborhood of x_0 and $e^{-n\log\rho}$ is not integrable near x_0 (as it is equal to $|z|^{-2n}$ there), the finiteness of the left-hand side of (4-2) forces $u_\varepsilon(x_0) = 0$, and hence

$$s_\varepsilon(x_0) = f_{x_0} \otimes \zeta_{x_0} = s_{x_0}.$$

Also recalling that $\rho < e^{-1}$ on $\overline{\Omega}$ and using the estimate

$$\eta_\varepsilon + \lambda_\varepsilon \le 6\psi_\varepsilon^2 \le 6\log^2\rho,$$

we conclude that

$$(\eta_\varepsilon + \lambda_\varepsilon)^{-1} e^{-(\rho+n\log\rho)} \ge \frac{\mathrm{const}}{\rho^n \log^2\rho} \ge \mathrm{const}_n.$$

Hence (4-2) implies

$$\int_\Omega \sqrt{-1}^{n^2}\{u_\varepsilon, u_\varepsilon\}_h \le \mathrm{const}_{n,\Omega,\omega} \frac{\sqrt{-1}^{n^2}\{s_{x_0}, s_{x_0}\}_h}{\omega^n(x_0)}.$$

Now combining this with the choice of f_ε and applying the Cauchy–Schwarz inequality, we end up with

$$\int_\Omega \sqrt{-1}^{n^2}\{s_\varepsilon, s_\varepsilon\}_h \le \mathrm{const}_{n,\Omega,\omega} \frac{\sqrt{-1}^{n^2}\{s_{x_0}, s_{x_0}\}_h}{\omega^n(x_0)}.$$

Consequently $s = s_\varepsilon$ is a holomorphic section of $\Lambda^n T^*\Omega \otimes L$ as desired in the theorem, provided $\varepsilon > 0$ is sufficiently small. This completes the proof of Theorem 4.1.

Applying Theorem 4.1 to the trivial line bundle $L = \Omega \times \mathbb{C}$ equipped with the metric $e^{-\varphi}$, $\varphi \in \mathrm{Psh}(\Omega)$, we get immediately the following:

Theorem 4.3. *Let X be an n-dimensional Stein manifold with a continuous volume form $\mathrm{d}V_X$ and Ω be a relatively compact complete Kähler domain in X. Then, for every $x_0 \in \Omega$, $f_{x_0} \in \Lambda^n T^*_{x_0}\Omega$ and $\varphi \in \mathrm{Psh}(\Omega)$, there exists a holomorphic n-form f on Ω such that $f(x_0) = f_{x_0}$ and*

$$\int_\Omega \sqrt{-1}^{n^2} f \wedge \bar f e^{-\varphi} \le \mathrm{const}_{n,\Omega,\mathrm{d}V_X} \frac{\sqrt{-1}^{n^2} f_{x_0} \wedge \bar f_{x_0}}{\mathrm{d}V_X(x_0)} e^{-\varphi(x_0)}. \tag{4-3}$$

Remark 4.4. Note that the conclusion of Theorem 4.3 holds trivially when the right-hand side of (4-3) is infinite: in this case every holomorphic n-form f on Ω that takes the value f_{x_0} at x_0 automatically satisfies the desired estimate (4-3), and such an f always exists (even on the whole X) since X is Stein.

On the other hand, as indicated by a result of [Chen et al. 2015], the conclusion of Theorem 4.3 (and hence Theorem 4.1) does not hold in general when the assumption that Ω is complete Kähler is dropped.

5. Proof of Theorem 1.2

The key ingredients of the proof are Theorem 4.3 and some basic geometric measure theory. As already mentioned in the Introduction, the idea of using Theorem 4.3 in the proof was inspired by [Chen et al. 2015] (and [Demailly 1992]). To get the idea off the ground we first need to prove the following result, which is also of independent interest.

Theorem 5.1. *Let X be a Stein manifold and $E \subset X$ be a closed complete pluripolar set. Then $X \setminus E$ admits a complete Kähler metric.*

When $E \subset X$ is a complex subvariety, this result is a well-known theorem of Grauert [1956] (see also [Demailly 1982] for closely related results). Note that even in this special case, $X \setminus E$ cannot be Stein unless E is empty or has pure codimension 1, in view of the second Riemann extension theorem.

Proof. The result is essentially an easy consequence of Theorem 2.2, as we now explain below.

Let φ be a C^∞ strictly psh exhaustion function for X, and let $\psi : X \setminus E \to \mathbb{R}$ be a C^∞ function such that

$$\psi = -\log(-\rho) \quad \text{on } U \setminus E,$$

where ρ is a psh function as in Theorem 2.2 with $\sup_X \rho > -1$ and $U := \{\rho < -1\}$. Since $-\log(-t)$ is a convex increasing function of $t < 0$, ψ is psh on $U \setminus E$, and hence $\sqrt{-1}\partial\bar\partial\psi \geq 0$ there. One can therefore construct a convex, rapidly increasing C^∞ function χ on $\mathbb{R}$ such that

$$\omega := \sqrt{-1}\partial\bar\partial(\chi \circ \varphi) + \sqrt{-1}\partial\bar\partial\psi \geq \omega_0 \quad \text{on } X \setminus E$$

for some complete Kähler metric ω_0 on X.

We claim that such an ω is complete on $X \setminus E$. To prove this, we may assume without loss of generality that X itself is connected (and so is $X \setminus E$). Suppose $\{x_j\}_{j\in\mathbb{N}}$ is a bounded sequence in the metric space $(X \setminus E, \omega)$. Then there is a sequence of smooth curves $\{\gamma_j\}_{j\in\mathbb{N}} \subset C^\infty([0,1], X \setminus E)$ with uniformly bounded lengths with respect to ω, such that each γ_j joins x_j to a fixed reference point in $X \setminus \overline{U}$. Since $\omega \geq \omega_0$ on $X \setminus E$ and ω_0 is complete on X, we may assume that the sequence $\{x_j\}_{j\in\mathbb{N}}$ itself converges in X by passing to a subsequence if necessary. What now remains is to show that the limit of $\{x_j\}_{j\in\mathbb{N}}$ lies outside E. Suppose the contrary and set

$$t_j := \inf\{t \in [0,1] : \gamma_j([t,1]) \subset U\}, \quad j \in \mathbb{N}.$$

Clearly $0 < t_j < 1$ and $\gamma_j(t_j) \in \partial U = \{\rho = -1\}$ for all sufficiently large j. Observe also that

$$\omega \geq \sqrt{-1}\partial\bar\partial(-\log(-\rho)) \geq \sqrt{-1}\partial\log(-\rho) \wedge \bar\partial\log(-\rho) \quad \text{on } U \setminus E.$$

It then follows that

$$\sqrt{2}\,\mathrm{length}_\omega(\gamma_j) \geq \int_{t_j}^1 \big|(\mathrm{d}\log(-\rho))(\gamma_j'(t))\big|\,\mathrm{d}t \geq \int_{t_j}^1 (\mathrm{d}\log(-\rho))(\gamma_j'(t))\,\mathrm{d}t$$
$$= \log(-\rho(x_j)) \to \infty \qquad \text{as } j \to \infty,$$

contradicting the boundedness of $\{\mathrm{length}_\omega(\gamma_j)\}_{j\in\mathbb{N}}$. Therefore the limit of $\{x_j\}_{j\in\mathbb{N}}$ lies outside E, and hence ω is complete on $X \setminus E$. □

Remark 5.2. If we denote by d_ω and d_{ω_0} the distance functions on $X \setminus E$ and X induced by ω and ω_0, respectively, and by $\chi_{U\setminus E}$ the characteristic function of $U \setminus E$, then the preceding argument actually yields the estimate

$$d_\omega(x, x_0) \geq \max\left\{d_{\omega_0}(x, x_0),\, \frac{\chi_{U\setminus E}(x)\log(-\rho(x))}{\sqrt{2}}\right\}, \quad x \in X \setminus E,\ \ x_0 \in X \setminus \overline{U}.$$

This, of course, implies that ω is complete.

We are now ready to prove Theorem 1.2.

Proof of Theorem 1.2. The uniqueness part of the theorem is clear, since two psh functions on Ω that coincide almost everywhere are actually equal everywhere. So it suffices to prove the existence part.

Let us first reduce the problem to the case when $\Omega \supset K$ is a relatively compact Stein domain in the ambient manifold, which we denote by X. To this end, let ρ be a psh function on X, continuous outside K and satisfying $\rho^{-1}(-\infty) = K$ (see Theorem 2.2). Choose an open set $U \subset X$ such that $K \subset U \Subset X$, and set

$$\tilde{\rho} := \begin{cases} \max\{\rho, \inf_{\partial U}\rho\} & \text{on } X \setminus \overline{U}, \\ \rho & \text{on } \overline{U}. \end{cases}$$

Then $\tilde{\rho}$ is a psh function on X with $\tilde{\rho}^{-1}(-\infty) = K$. Replacing Ω by any connected component of $\{\tilde{\rho} < \inf_{\partial U}\rho\}$, we may assume in what follows that Ω is a relatively compact Stein domain in X.

Now $\Omega \setminus K$ is a relatively compact complete Kähler domain in X, in view of Theorem 5.1. Write $n = \dim_{\mathbb{C}} X$ and fix a continuous volume form $\mathrm{d}V_X$ on X. Given $\varphi \in \mathrm{Psh}(\Omega \setminus K)$, we can therefore invoke Theorem 4.3 to assign to every point $x \in \Omega \setminus K$ a holomorphic n-form f_x on $\Omega \setminus K$ with the property that

$$\frac{\sqrt{-1}^{\,n^2} f_x \wedge \bar{f}_x}{\mathrm{d}V_X}(x) = e^{\varphi(x)} \quad \text{and} \quad \int_{\Omega\setminus K} \sqrt{-1}^{\,n^2} f_x \wedge \bar{f}_x e^{-\varphi} \leq \text{const}, \tag{5-1}$$

where the constant involved is independent of $x \in \Omega\setminus K$. By the Hartogs extension theorem for holomorphic forms (see, e.g., [Serre 1955, p. 22, corollaire du théorème 3]), each such f_x extends holomorphically to Ω. With a slight abuse of notation, we still denote this extension by f_x.

We next prove the psh extendibility of φ across K. By a classical result of Lelong, it suffices to show that every point of K admits a small neighborhood on which φ is bounded above. Fix an arbitrary point $x_0 \in K$. Since the problem is local, we may assume without loss of generality that Ω is a bounded pseudoconvex domain in $\mathbb{C}^n$ containing x_0. Let $z = (z_1, \ldots, z_n)$ be the coordinate of Ω. Identifying x

with its coordinate $z = z(x)$ and f_x with $f_x/\mathrm{d}z_1 \wedge \cdots \wedge \mathrm{d}z_n$, we conclude from (5-1), after shrinking Ω if necessary, that

$$e^{\varphi(z)} \le \text{const}|f_z(z)|^2 \quad \text{and} \quad \int_{\Omega\setminus K} |f_z|^2 e^{-\varphi} \le \text{const}, \tag{5-2}$$

where the integral involved is taken against the Lebesgue measure on $\mathbb{C}^n$. To proceed with the proof we now recall that, as a polar subset of $\mathbb{C}^n \cong \mathbb{R}^{2n}$, K has Hausdorff dimension at most $2n-2$ (see, e.g., [Armitage and Gardiner 2001, Theorem 5.9.6]). This enables us to make use of a result of Shiffman (see [Harvey 1977, Lemma 2.3] or [Demailly 2012b, Chapter III, Lemma 4.7]), according to which we can find — by suitably selecting affine linear coordinates for $\mathbb{C}^n$ — a polydisc $\Delta' \times \Delta'' \subset \mathbb{C}^{n-1} \times \mathbb{C}$ centered at $z_0 = z(x_0) =: (z_0', z_0'')$ such that

$$(\Delta' \times \partial\Delta'') \cap K = \varnothing.$$

By shrinking Δ' if necessary, we further arrive at

$$\big(\Delta' \times (\Delta'' \setminus (1-\varepsilon)\overline{\Delta''})\big) \cap K = \varnothing$$

for some sufficiently small $\varepsilon > 0$. The remaining argument is the same as in the last paragraph of the proof of [Chen et al. 2015, Theorem 1.2]. Choose $R, r > 0$ such that

$$1 - \varepsilon < r < R < 1$$

and a ball $B \Subset \Delta'$ centered at z_0'. Then the Cauchy estimate for holomorphic functions and the inequalities in (5-2) imply

$$\begin{aligned} e^{\varphi(z)} \le \text{const}|f_z(z)|^2 &\le \text{const}_{B,R,r,\varepsilon} \int_{B\times(R\Delta''\setminus r\Delta'')} |f_z|^2 \\ &\le \text{const}_{B,R,r,\varepsilon} \sup_{B\times(R\Delta''\setminus r\Delta'')} e^{\varphi} \int_{B\times(R\Delta''\setminus r\Delta'')} |f_z|^2 e^{-\varphi} \le C \sup_{B\times(R\Delta''\setminus r\Delta'')} e^{\varphi} \end{aligned}$$

for all $z \in ((1-\varepsilon)(B \times \Delta'')) \setminus K$, where $C > 0$ is a constant independent of z. Consequently, φ is bounded above on $(1-\varepsilon)(B \times \Delta'') \ni z_0$ outside K. This completes the proof. □

Acknowledgements

The author is very grateful to Professor Bo-Yong Chen for bringing the Ohsawa problem to his attention and for many helpful discussions. The author would also like to thank Professor Xiangyu Zhou — his mentor when he was a postdoctoral fellow at the Institute of Mathematics, AMSS, Chinese Academy of Sciences from 2017 to 2019 — for his constant support and encouragement over the years. In addition, thanks go to Professor Xiankui Meng for interesting exchanges about compact complete pluripolar sets, as well as Professors Taeyong Ahn, Zhuo Liu, Daowei Ma, Takeo Ohsawa and Xiangyu Zhou for their interest in this work. Last but not least, the author is greatly indebted to the referees for carefully reading the manuscript and suggesting many improvements.

The author was partially supported by the NSFC (Grant No. 12371083) and the Fundamental Research Funds for the Central Universities (Grant Nos. WK0010000099 and WK3470000030).

References

[Armitage and Gardiner 2001] D. H. Armitage and S. J. Gardiner, *Classical potential theory*, Springer, 2001. MR

[Bedford and Taylor 1988] E. Bedford and B. A. Taylor, "Smooth plurisubharmonic functions without subextension", *Math. Z.* **198**:3 (1988), 331–337. MR

[Berndtsson 2010] B. Berndtsson, "An introduction to things $\bar{\partial}$", pp. 7–76 in *Analytic and algebraic geometry*, edited by J. McNeal and M. Mustaţă, IAS/Park City Math. Ser. **17**, Amer. Math. Soc., Providence, RI, 2010. MR

[Berndtsson 2018] B. Berndtsson, "Complex Brunn–Minkowski theory and positivity of vector bundles", pp. 859–884 in *Proceedings of the International Congress of Mathematicians* (*ICM* 2018) (Rio de Janeiro, 2018), edited by B. Sirakov et al., World Sci. Publ., Hackensack, NJ, 2018. MR

[Berndtsson and Lempert 2016] B. Berndtsson and L. Lempert, "A proof of the Ohsawa–Takegoshi theorem with sharp estimates", *J. Math. Soc. Japan* **68**:4 (2016), 1461–1472. MR

[Błocki 2013] Z. Błocki, "Suita conjecture and the Ohsawa–Takegoshi extension theorem", *Invent. Math.* **193**:1 (2013), 149–158. MR

[Błocki 2014] Z. Błocki, "Cauchy–Riemann meet Monge–Ampère", *Bull. Math. Sci.* **4**:3 (2014), 433–480. MR

[Błocki and Kołodziej 2007] Z. Błocki and S. Kołodziej, "On regularization of plurisubharmonic functions on manifolds", *Proc. Amer. Math. Soc.* **135**:7 (2007), 2089–2093. MR

[Boggess et al. 2022] A. Boggess, R. Dwilewicz, and E. Porten, "On the Hartogs extension theorem for unbounded domains in $\mathbb{C}^n$", *Ann. Inst. Fourier* (*Grenoble*) **72**:3 (2022), 1185–1206. MR

[Cao et al. 2024] J. Cao, M. Paun, and B. Berndtsson, "On the Ohsawa–Takegoshi extension theorem", *J. Geom. Anal.* **34**:1 (2024), art. id. 25. MR

[Chan and Choi 2022] T. O. M. Chan and Y.-J. Choi, "Extension with log-canonical measures and an improvement to the plt extension of Demailly–Hacon–Păun", *Math. Ann.* **383**:3-4 (2022), 943–997. MR

[Chen 2011] B.-Y. Chen, "A simple proof of the Ohsawa–Takegoshi extension theorem", preprint, 2011. arXiv 1105.2430

[Chen 2018] B.-Y. Chen, "Hardy–Sobolev type inequalities and their applications", preprint, 2018. arXiv 1712.02044v2

[Chen et al. 2015] B.-Y. Chen, J. Wu, and X. Wang, "Ohsawa–Takegoshi type theorem and extension of plurisubharmonic functions", *Math. Ann.* **362**:1-2 (2015), 305–319. MR

[Colţoiu 1990] M. Colţoiu, "Complete locally pluripolar sets", *J. Reine Angew. Math.* **412** (1990), 108–112. MR

[Demailly 1982] J.-P. Demailly, "Estimations L^2 pour l'opérateur $\bar{\partial}$ d'un fibré vectoriel holomorphe semi-positif au-dessus d'une variété kählérienne complète", *Ann. Sci. École Norm. Sup.* (4) **15**:3 (1982), 457–511. MR

[Demailly 1992] J.-P. Demailly, "Regularization of closed positive currents and intersection theory", *J. Algebraic Geom.* **1**:3 (1992), 361–409. MR

[Demailly 1994] J.-P. Demailly, "Regularization of closed positive currents of type (1, 1) by the flow of a Chern connection", pp. 105–126 in *Contributions to complex analysis and analytic geometry*, edited by H. Skoda and J.-M. Trépreau, Aspects Math. **E26**, Friedr. Vieweg, Braunschweig, 1994. MR

[Demailly 2000] J.-P. Demailly, "On the Ohsawa–Takegoshi–Manivel L^2 extension theorem", pp. 47–82 in *Complex analysis and geometry* (Paris, 1997), edited by P. Dolbeault et al., Progr. Math. **188**, Birkhäuser, 2000. MR

[Demailly 2012a] J.-P. Demailly, *Analytic methods in algebraic geometry*, Surveys of Modern Mathematics **1**, International Press, Somerville, MA, 2012. MR

[Demailly 2012b] J.-P. Demailly, "Complex analytic and differential geometry", open content book, Université de Grenoble I, 2012, available at https://www-fourier.univ-grenoble-alpes.fr/∼demailly/manuscripts/agbook.pdf.

[Deng et al. 2023] F. Deng, J. Ning, Z. Wang, and X. Zhou, "Positivity of holomorphic vector bundles in terms of L^p-estimates of $\bar{\partial}$", *Math. Ann.* **385**:1-2 (2023), 575–607. MR

[Deng et al. 2024] F. Deng, Z. Wang, L. Zhang, and X. Zhou, "New characterization of plurisubharmonic functions and positivity of direct image sheaves", *Amer. J. Math.* **146**:3 (2024), 751–768. MR

[Dieu and Manh 2014] N. Q. Dieu and P. V. Manh, "Complete pluripolar graphs in $\mathbb{C}^N$", *Ann. Polon. Math.* **112**:1 (2014), 85–100. MR

[Edlund 2004] T. Edlund, "Complete pluripolar curves and graphs", *Ann. Polon. Math.* **84**:1 (2004), 75–86. MR

[El Mir 1984] H. El Mir, "Sur le prolongement des courants positifs fermés", *Acta Math.* **153**:1-2 (1984), 1–45. MR

[Grauert 1956] H. Grauert, "Charakterisierung der Holomorphiegebiete durch die vollständige Kählersche Metrik", *Math. Ann.* **131** (1956), 38–75. MR

[Guan et al. 2025] Q. Guan, Z. Mi, and Z. Yuan, "Guan–Zhou's unified version of optimal L^2 extension theorem on weakly pseudoconvex Kähler manifolds", *J. Math. Soc. Japan* **77**:1 (2025), 167–188. MR

[Guan et al. 2026] Q. Guan, Z. Mi, and Z. Yuan, "Optimal L^2 extension for holomorphic vector bundles with singular Hermitian metrics", *Peking Math. J.* **9**:1 (2026), 105–193. MR

[Guan and Zhou 2015a] Q. Guan and X. Zhou, "A solution of an L^2 extension problem with an optimal estimate and applications", *Ann. of Math.* (2) **181**:3 (2015), 1139–1208. MR

[Guan and Zhou 2015b] Q. Guan and X. Zhou, "A proof of Demailly's strong openness conjecture", *Ann. of Math.* (2) **182**:2 (2015), 605–616. MR

[Harvey 1974] R. Harvey, "Removable singularities for positive currents", *Amer. J. Math.* **96** (1974), 67–78. MR

[Harvey 1977] R. Harvey, "Holomorphic chains and their boundaries", pp. 309–382 in *Several complex variables* (Williamstown, MA., 1975), edited by J. Wells, R. O., Proc. Sympos. Pure Math. **XXX:**1, Amer. Math. Soc., Providence, RI, 1977. MR

[Harvey and Polking 1975] R. Harvey and J. Polking, "Extending analytic objects", *Comm. Pure Appl. Math.* **28**:6 (1975), 701–727. MR

[Hosono 2020] G. Hosono, "The optimal jet L^2 extension of Ohsawa–Takegoshi type", *Nagoya Math. J.* **239** (2020), 153–172. MR

[Kim 2023] D. Kim, "L^2 extension of holomorphic functions for log canonical pairs", *J. Math. Pures Appl.* (9) **177** (2023), 198–213. MR

[Levenberg et al. 1992] N. Levenberg, G. Martin, and E. A. Poletsky, "Analytic disks and pluripolar sets", *Indiana Univ. Math. J.* **41**:2 (1992), 515–532. MR

[Li et al. 2023] Z. Li, W. Xu, and X. Zhou, "On Demailly's L^2 extension theorem from non-reduced subvarieties", *Math. Z.* **305**:2 (2023), art. id. 23. MR

[Liu and Xu 2024] Z. Liu and W. Xu, "Characterizations of Griffiths positivity, pluriharmonicity and flatness", *J. Funct. Anal.* **287**:7 (2024), art. id. 110532. MR

[McNeal and Varolin 2015] J. D. McNeal and D. Varolin, "L^2 estimates for the $\bar{\partial}$ operator", *Bull. Math. Sci.* **5**:2 (2015), 179–249. MR

[Ohsawa 1995] T. Ohsawa, "On the extension of L^2 holomorphic functions, III: Negligible weights", *Math. Z.* **219**:2 (1995), 215–225. MR

[Ohsawa 2018] T. Ohsawa, *L^2 approaches in several complex variables*: *towards the Oka–Cartan theory with precise bounds*, 2nd ed., Springer, 2018. MR

[Ohsawa 2020] T. Ohsawa, "A survey on the L^2 extension theorems", *J. Geom. Anal.* **30**:2 (2020), 1366–1395. MR

[Ohsawa and Takegoshi 1987] T. Ohsawa and K. Takegoshi, "On the extension of L^2 holomorphic functions", *Math. Z.* **195**:2 (1987), 197–204. MR

[Ransford 1995] T. Ransford, *Potential theory in the complex plane*, London Mathematical Society Student Texts **28**, Cambridge University Press, Cambridge, 1995. MR

[Rao and Zhang 2021] S. Rao and R. Zhang, "L^2 extension theorem for jets with variable denominators", *C. R. Math. Acad. Sci. Paris* **359** (2021), 181–193. MR

[Serre 1955] J.-P. Serre, "Un théorème de dualité", *Comment. Math. Helv.* **29** (1955), 9–26. MR

[Shiffman 1968] B. Shiffman, "On the removal of singularities of analytic sets", *Michigan Math. J.* **15** (1968), 111–120. MR

[Shiffman 1972] B. Shiffman, "Extension of positive line bundles and meromorphic maps", *Invent. Math.* **15**:4 (1972), 332–347. MR

[Sibony 1985] N. Sibony, "Quelques problèmes de prolongement de courants en analyse complexe", *Duke Math. J.* **52**:1 (1985), 157–197. MR

[Siu 1974] Y. T. Siu, "Analyticity of sets associated to Lelong numbers and the extension of closed positive currents", *Invent. Math.* **27** (1974), 53–156. MR

[Siu 2002] Y.-T. Siu, "Extension of twisted pluricanonical sections with plurisubharmonic weight and invariance of semipositively twisted plurigenera for manifolds not necessarily of general type", pp. 223–277 in *Complex geometry* (Göttingen, 2000), edited by I. Bauer et al., Springer, 2002. MR

[Skoda 1982] H. Skoda, "Prolongement des courants, positifs, fermés de masse finie", *Invent. Math.* **66**:3 (1982), 361–376. MR

[Vîjîitu 2022] V. Vîjîitu, "On Hartogs' extension", *Ann. Mat. Pura Appl.* (4) **201**:1 (2022), 487–498. MR

[Wang 2024a] X. Wang, "Bott–Chern cohomology and the Hartogs extension theorem for pluriharmonic functions", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **25**:1 (2024), 605–610. MR

[Wang 2024b] X. Wang, "Complete pluripolar sets and removable singularities of plurisubharmonic functions", *Indiana Univ. Math. J.* **73**:6 (2024), 1929–1936. MR

[Xu and Zhou 2024] W. Xu and X. Zhou, "Optimal L^2 extensions of openness type", *Math. Ann.* **390**:1 (2024), 1249–1307. MR

[Zhou and Zhu 2018] X. Zhou and L. Zhu, "An optimal L^2 extension theorem on weakly pseudoconvex Kähler manifolds", *J. Differential Geom.* **110**:1 (2018), 135–186. MR

[Zhou and Zhu 2020] X. Zhou and L. Zhu, "Siu's lemma, optimal L^2 extension and applications to twisted pluricanonical sheaves", *Math. Ann.* **377**:1-2 (2020), 675–722. MR

Received 18 Sep 2024. Accepted 2 Nov 2025.

XIEPING WANG: xpwang008@ustc.edu.cn
CAS Wu Wen-Tsun Key Laboratory of Mathematics and School of Mathematical Sciences, University of Science and Technology of China, Hefei, China

ANALYSIS AND PDE
Vol. 19 (2026), No. 7, pp. 1417–1445
DOI: 10.2140/apde.2026.19.1417

msp

FLOW OF THE ZEROS OF POLYNOMIALS UNDER ITERATED DIFFERENTIATION

ANDREI MARTÍNEZ-FINKELSHTEIN AND EVGUENII A. RAKHMANOV

For a monic polynomial Q_n of degree n, let $Q_{n,k}$ be its k-th derivative normalized to be monic. Under the only assumption that the sequence $\{Q_n\}$ has a weak-* limiting zero distribution (an empirical distribution of zeros) represented by a probability measure μ_0 with compact support in the complex plane, we show that, as $n, k \to \infty$ such that $k/n \to t \in (0, 1)$, the Cauchy transform of the normalized zero-counting measure of the polynomials $Q_{n,k}$ converges in a neighborhood of infinity to an analytic function, uniquely determined by μ_0 and t, that can be written as the Cauchy transform of a measure μ_t, not necessarily uniquely determined unless μ_0 is supported on the real line.

The family of these Cauchy transforms and, when well defined, the corresponding measures μ_t, $t \in (0, 1)$, whose dependence on the parameter t can be interpreted as a flow of the zeros under iterated differentiation, has several interesting connections with the inviscid Burgers equation, the fractional free convolution of μ_0, or a nonlocal diffusion equation governing the density of μ_t on $\mathbb{R}$.

We provide an elementary and unified approach that not only recovers but also explains various phenomena observed in prior works — from Burgers-type PDEs to free probability limits.

1. Introduction

In recent years, the problem of repeated differentiation of polynomials has been intensively studied from different points of view and in different contexts. In this introduction, we start by formulating our main result. After that, we make a few remarks about the classical roots of the problem and briefly comment on some recent results.

Hereafter, the asymptotic zero distribution of a sequence of polynomials (also known as its empirical distribution of zeros) is understood in terms of the weak-* limits of their zero-counting measures. We define the zero-counting measure of a polynomial P as

$$\chi(P) := \sum_{P(x)=0} \delta_x, \tag{1-1}$$

where the sum is taken over all zeros of P, accounting for multiplicities, and δ_x is the Dirac measure (unit mass) supported at x. This way, $\chi(P)$ is independent of normalization of P, and

$$\chi(P)(\mathbb{C}) = \deg(P).$$

MSC2020: primary 30C15; secondary 30C10, 37F10, 46L54, 76B99.
Keywords: polynomials, zeros, empirical distribution of zeros, weak convergence, inviscid Burgers equation, free probability.

© 2026 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.

We are interested in the following problem: given a sequence of (monic) polynomials Q_n, $\deg Q_n = n$, $n \in \mathbb{N}$, and the triangular table of their iterated derivatives

$$Q_{n,k}(x) := \frac{(n-k)!}{n!}\frac{d^k}{dx^k}Q_n(x) = x^{n-k} + \text{lower degree terms}, \quad k = 0, 1, \dots, n-1, \tag{1-2}$$

what are the possible limit distributions of the normalized zero-counting measures

$$\sigma_{n,k} := \frac{1}{n}\chi(Q_{n,k}) \tag{1-3}$$

as $n \to \infty$ in such a way that $k/n \to t \in [0, 1)$, if we only know that the $\sigma_{n,0}$ have a weak-* limit, compactly supported on $\mathbb{C}$?

A weaker form of this question deals with the convergence of the logarithmic derivatives

$$v_{n,k}(z) := \frac{Q'_{n,k}(z)}{nQ_{n,k}(z)}. \tag{1-4}$$

The main result of this paper is the following:

Theorem 1.1. *Let Q_n be a sequence of polynomials whose zeros belong to a compact convex set S. Assume also that the limit*

$$\lim_n \frac{Q'_n(z)}{nQ_n(z)} = u(z)$$

exists uniformly on compact subsets of $\Omega = \mathbb{C} \setminus S$.

Then for any $t \in [0, 1)$ and any sequence k_n of natural numbers with $k_n/n \to t$ we have that

$$\lim_n v_{n,k}(z) = \lim_n \frac{Q'_{n,k}(z)}{nQ_{n,k}(z)} = u(z, t) \tag{1-5}$$

uniformly on compact subsets of Ω, where $u(z, t)$ is uniquely determined by the equation

$$\frac{\partial u}{\partial t}(z, t) = \frac{1}{u(z, t)}\frac{\partial u}{\partial z}(z, t) \tag{1-6}$$

with the initial condition $u(z, 0) = u(z)$.

Equation (1-6) is a special case of a nonlinear wave equation, known as the inviscid Burgers or Hopf equation.[1] It is traditionally considered for the real values of the variables z and t, but it has also been studied on the complex plane, both in its viscous and inviscid form; see, for example, [del Monaco and Schleißinger 2016; Hall et al. 2025; 2026; Hotta and Katori 2018; Hotta and Schleißinger 2021; VandenHeuvel et al. 2023].

The Hopf equation is known for a remarkable phenomenon: its solutions can develop singularities (or shocks) even when posed on the real line. However, this behavior does not occur when the initial condition $u(z, 0)$ is given by the Cauchy transform of a unit *positive* measure supported on a bounded convex set S. In this case, the solution $u(z, t)$ to equation (1-6) remains analytic in the domain $\Omega = \mathbb{C} \setminus S$.

[1] It goes by various other names such as dispersionless Burgers equation, dispersionless KdV (Korteweg-de Vries) equation, inviscid KdV; see, e.g., [Arendt and Urban 2023, Section 2.1.3], [Evans 1998, Section 3.4] or [Markowich 2007, Chapter 9].

Moreover, for any fixed $t \in (0, 1)$, the solution $u(z, t)$ can itself be represented as the Cauchy transform of a positive measure supported on S — although this measure is not necessarily unique. This apparently new observation is implicitly contained in the statement of Theorem 1.1. Indeed, the initial measure can be approximated by the zero-counting measures of a sequence of polynomials Q_n whose zeros are all contained in S. The analyticity of $u(z, t)$ in Ω then follows from the convergence in (1-5). The representation of u as a Cauchy transform follows from the weak compactness of the associated sequence of zero-counting measures.

It is important to note that the analyticity statement fails if the assumption of positivity of μ_0 is dropped; see Remark 2.2 in Section 2.

The fact that, in the situation of complex zeros, not much more than (1-5) can be said is illustrated by the paradigmatic example of two sequences of polynomials, $Q_n(z) = z^n$ and $Q_n(z) = z^n - 1$. In both cases, the sequence of normalized zero-counting measures σ_{n,k_n}, defined in (1-3), converges as $k_n/n \to t$ to the same measure $(1-t)\delta_0$, although the μ_0 are clearly different: δ_0 in the former case, and the unit Lebesgue measure on the unit circle in the latter. In the notation of Theorem 1.1, in the domain $0 < |z| < 1$,

$$u(z) = 0 \ \text{ for } z^n \quad \text{and} \quad u(z) = 1/z \ \text{ for } z^n - 1,$$

which shows that $u(z)$ does not determine $u(z, t)$ there.

The situation is different if the Q_n have only real zeros or, more generally, lie on a segment of a straight line. It will be shown that $u(z, t)$ is the Cauchy transform of a unique positive measure compactly supported on the same segment. In this context, we denote this measure by μ_t. It is completely determined by the values of $u(z, t)$ in a neighborhood of infinity, and measures $\sigma_{n,k}$ weakly converge to μ_t. In more general cases, when uniqueness is not guaranteed, we use μ_t to denote one of the measures whose Cauchy transform coincides with $u(z, t)$ in a neighborhood of infinity; see for example the statement of (iii) in Theorem 2.1.

Let us comment on the classical roots of the problem of zeros of derivatives of polynomials.

A standard example of a flow generated by repeated differentiation is provided by the sequence of polynomials

$$Q_{2n,k}(z) = \frac{(2n-k)!}{(2n)!}\frac{d^k}{dx^k}(x^2-1)^n, \quad n = 0, 1, 2, \ldots, \quad k = [2tn], \tag{1-7}$$

for some $t \in [0, 1)$ (we denote by $[\,\cdot\,]$ the integer part of the argument). The notion of "flow" is justified by considering the parameter t as "time". The corresponding normalized sequence of zero-counting measures $\frac{1}{2n}\chi(Q_{2n,[2tn]})$ converges in the weak-* sense[2] to a measure μ_t with total mass $1-t$, which has an explicit representation

$$\mu_t = \max\left\{\tfrac{1}{2} - t, 0\right\}(\delta_{-1} + \delta_1) + \mu_t^{\mathrm{ac}}, \tag{1-8}$$

[2]The weak-* convergence in this context is a classical result and may be regarded as part of the folklore in the analytic theory of polynomials. For explicit formulas in this setting, see, for example, [Bøgvad et al. 2024; Hoskins and Kabluchko 2023] or the derivation provided in Section 5.

where μ_t^{ac} is an absolutely continuous measure supported on the interval $\big[-2\sqrt{t(1-t)}, 2\sqrt{t(1-t)}\big]$, with

$$d\mu_t^{\mathrm{ac}}(x) = \frac{1}{2\pi}\frac{\sqrt{4t(1-t)-x^2}}{1-x^2}\,dx. \tag{1-9}$$

As t progresses, μ_t represents the flow of zeros of iterated derivatives, starting from its initial distribution $\mu_0 = \frac{1}{2}(\delta(-1)+\delta(1))$. The measure μ_t is well known: it is related to zeros of Gegenbauer polynomials and hence to the angular part of the hyperspherical harmonics; see, e.g., [Saff and Totik 2024, Examples IV.1.18 and IV.6.2]. In fact, for $k = n$, (1-7) is just the Rodrigues formula for the Legendre polynomials, orthogonal in $[-1, 1]$ with respect to the weight $w(x) = 1$. This also explains why μ_t, for $t = \frac{1}{2}$, is one-half of the equilibrium measure (or the Robin distribution) of the interval $[-1, 1]$. Notice also the curious symmetry of μ_t^{ac} with respect to the time reversal ($t \mapsto 1-t$). Since every classical family of orthogonal polynomials satisfies a Rodrigues formula, each of these families originates its own flow of measures with interesting properties.

The formula (1-7) can be extended in several directions. For instance, the type-II Hermite–Padé (or multiple orthogonal) polynomials Q_{2n} of degree $2n$, defined up to normalization by the orthogonality conditions

$$\int_{-1}^{0} Q_{2n}(x)x^k\,dx = \int_{0}^{1} Q_{2n}(x)x^k\,dx = 0, \quad k = 0, 1, \ldots, n-1, \tag{1-10}$$

have the representation

$$Q_{2n}(x) = \frac{(2n)!}{(3n)!}\frac{d^n}{dx^n}(x^3-x)^n, \tag{1-11}$$

obtained originally by Kalyagin [1979], who used it to study the asymptotic properties of Q_{2n}. As before, we can consider

$$Q_{2n,k}(z) = \frac{(3n-k)!}{(3n)!}\frac{d^k}{dx^k}(x^3-x)^n, \quad n = 0, 1, 2, \ldots, \quad k = [3tn],$$

and the corresponding family of measures μ_t of their limiting zero distributions. In this case, the explicit formulas for μ_t are not as simple as before. Similar Rodrigues-type formulas exist for other families of semiclassical and multiple orthogonal polynomials [Aptekarev et al. 1997; Sorokin 2022; 2023]. Recently, the asymptotics of polynomials given by (1-7), where x^2-1 on the right-hand side is replaced by an arbitrary polynomial P, possibly with complex zeros, has been considered in [Bøgvad et al. 2024]; see Section 5 for more details and comments on some results of this work.

The interest in this subject has recently been revitalized by a series of papers by Steinerberger [O'Rourke and Steinerberger 2021; Steinerberger 2019; 2023], and several other works have followed in rapid succession, establishing interesting connections with free probability and the inviscid Burgers (or Hopf) partial differential equation.

First, Steinerberger [2019] wrote a nonlocal diffusion equation that should describe the evolution of the absolutely continuous part μ_t^{ac} of μ_t. Observing the similarities with an equation in [Shlyakhtenko and Tao 2022], he conjectured in [Steinerberger 2023] that μ_t can be written in terms of the fractional free additive convolution of the initial measure μ_0. Hoskins and Kabluchko [2023] rigorously proved this; later,

it was also established in [Arizmendi et al. 2023], this time using elegant arguments involving only finite free convolution of polynomials, a notion developed in a series of works of Markus and collaborators; see [Marcus et al. 2022]. In the case of complex random zeros, the free probability interpretation of repeated differentiation has been obtained in [Campbell et al. 2024].

The paper [Marcus et al. 2022] also contains another interesting motivation for studying our main problem: the average characteristic polynomial of all $k \times k$ minors of an $n \times n$ matrix A can be expressed in terms of the $(n-k)$-th derivative of the characteristic polynomial of A; see [Marcus et al. 2022, Lemma 1.17] for a precise formulation.

Other works addressing several aspects of repeated differentiation (many of them in the context of random polynomials) are [Angst et al. 2024; Byun et al. 2022; Campbell et al. 2024; Galligo 2022; Hoskins and Steinerberger 2022; Kabluchko 2026; Kiselev and Tan 2022]. A related differential operator (the heat-flow operator), which is degree-preserving and contains derivatives of all orders, was studied in [Hall et al. 2025]. The authors present a transport map T_t, defined in terms of the Cauchy (or Stieltjes) transform of μ_0, such that $T_t(\mu_0) = \mu_t$. Moreover, they show that the Cauchy transform of μ_t satisfies the inviscid Burgers equation (written in a slightly different form than the one appearing in the following); see [Hall et al. 2025, Equation (5.4)]. As they point out, the complex Burgers equation with initial value given by a Cauchy transform has recently been used in the description of the growth of multiple SLE; see [del Monaco and Schleißinger 2016; Hotta and Katori 2018; Hotta and Schleißinger 2021]. It also appears in [Aptekarev and Rykov 2006] in the description of the evolution of the end-points of the support of the equilibrium measure in a specific external field depending on t; see also [Martínez-Finkelshtein et al. 2015]. Also, equation (1.17) in [Shlyakhtenko and Tao 2022] can be reduced to (1-6). Another equivalent formulation of (1-6) appears again in [Hall et al. 2026]; the authors even address the behavior of individual zeros and conjecture (proving it in the rotationally invariant case) that the Cauchy transform of the measure, evaluated along the path of a single root, must remain constant, and the roots should move along the characteristic curves of a certain nonlinear PDE.

In this paper, we address the problem of the macroscopic description of the zeros of polynomials under repeated differentiation, with the most general assumptions stated in Theorem 1.1. However, our goal is not only to prove (and partially reprove) the results but also to show why they are true. The main feature of our approach is that it is totally elementary and uses only basic facts from complex analysis.

Our main findings can be summarized as follows. The complex Hopf equation (1-6) plays the main role. The connection of the problem on the flow of derivatives under differentiation with this equation lies in the elementary fact that the Cauchy transform of the zero-counting measure of a polynomial can be expressed in terms of the logarithmic derivative of this polynomial. As a first step, we show that if its initial value is given by the Cauchy transform of a compactly supported probability measure μ_0, then there is a universal domain, depending only on the support of μ_0, where not only the solutions of the equation remain analytic, but they can be written as the Cauchy transform of a measure for all values of the parameter $t \in (0, 1)$.

We show that the solution of the Hopf equation by the method of characteristics is precisely the expression of the fact that μ_t can be written as the fractional free convolution of μ_0.

Finally, we show that the nonlocal diffusion from [Steinerberger 2019] is merely a projection of the Hopf equation onto the real line via the Sohotski–Plemelj identities (or the Stieltjes–Perron inversion formula).

2. Main results

Let us start by introducing the following notation: for $0 < R < +\infty$, define $\mathbb{D}_R = \{z \in \mathbb{C} : |z| < R\}$, $\overline{\mathbb{D}}_R = \{z \in \mathbb{C} : |z| \le R\}$, and $\Omega_R = \{z \in \mathbb{C} : R < |z| < +\infty\} = \mathbb{C} \setminus \overline{\mathbb{D}}_R$. Notice that the open domain $\Omega_R \subset \mathbb{C}$ does not include infinity, while its closure in $\overline{\mathbb{C}}$ — that is, $\overline{\Omega}_R = \{z \in \mathbb{C} : R \le |z| \le +\infty\}$ — does. We also denote by ∂_z the differentiation operator with respect to the real or complex parameter z (each time, it will be clear from the context).

The *inviscid Burgers* or *Hopf equation*, considered in this paper, is written as

$$\partial_t u = \frac{\partial_z u}{u}, \quad u = u(z; t). \tag{2-1}$$

A more familiar form $\partial_t u + u\,\partial_z u = 0$ can be obtained, for instance, by replacing u by $-1/u$. We require that the parameter t takes real values in $[0, 1]$, while z is in the complex plane; in our case, z is in a neighborhood of infinity, that is, a domain of the form Ω_R.

It is well known that the solutions of (2-1) can exhibit shocks in a finite time t; we will be interested in a very particular type of initial value problem for (2-1), when $u(z, 0)$ is the Cauchy transform of a probability measure.

Given a finite Borel (in general, signed) measure μ on $\mathbb{C}$, we define its *Cauchy* (also known as Stieltjes or Markov) *transform* as

$$\mathcal{C}^\mu(z) := \int \frac{d\mu(y)}{z - y}, \tag{2-2}$$

which is a holomorphic function in $\mathbb{C} \setminus \operatorname{supp}(\mu)$. Moreover, if μ is compactly supported, then

$$\mathcal{C}^\mu(z) = \sum_{k=0}^{\infty} \frac{m_k}{z^{k+1}} \tag{2-3}$$

converges in a neighborhood of infinity; the coefficients

$$m_j = \int z^j \, d\mu(z), \quad j = 0, 1, 2, \ldots, \tag{2-4}$$

are the *moments* of the measure μ.

Let μ_0 be a compactly supported Borel probability measure on $\mathbb{C}$, and let $0 < r < +\infty$ be such that $\operatorname{supp} \mu_0 \subset \mathbb{D}_r$. One of the central results of this paper is the fact that the initial value problem for equation (2-1) with initial datum

$$u(z, 0) = \mathcal{C}^{\mu_0}(z) \tag{2-5}$$

has an analytic solution $u(z, t)$ for all $t \in [0, T]$, $0 < T < 1$, and z in a neighborhood of infinity that depends only on r and T but not on the measure μ_0 itself:

Theorem 2.1. *Let* $0 < T < 1$ *and* $0 < r < +\infty$ *be fixed. Then*:

(i) *There exists an* $R = R(T) > 2r$ *such that, for any probability measure* μ_0 *with* $\operatorname{supp}(\mu_0) \subset \mathbb{D}_r$ *and for* $(z, t) \in \Omega_R \times [0, T]$, *there exists a unique solution* $u(z, t)$ *of* (2-1) *satisfying the initial condition*

$$u(z, 0) = \mathcal{C}^{\mu_0}(z).$$

(ii) *For* $(z, t) \in \Omega_R \times [0, T]$, $u(z, t)$ *is given by*

$$u(z, t) = u(s, 0), \tag{2-6}$$

where the unique $s = s(z)$ *is implicitly defined by the equation* $F(z, s) = 0$, *with*

$$F(z, s) := z - s + \frac{t}{u(s, 0)}. \tag{2-7}$$

(iii) *For each* $t \in [0, 1)$, *there is a measure* μ_t, *with* $\operatorname{supp}(\mu_t) \subset \mathbb{D}_r$, *such that*

$$u(z, t) = \mathcal{C}^{\mu_t}(z), \quad z \in \Omega_r. \tag{2-8}$$

The moments

$$m_j(t) := \int z^j \, d\mu_t(z), \quad j = 0, 1, 2, \ldots, \tag{2-9}$$

of μ_t *are determined uniquely by* μ_0,[3] *with* $m_0(t) = 1 - t$. *Moreover, for* $k \in \mathbb{N}$, $m_k(t)$ *is a polynomial in* t *of degree at most* k *such that*

$$\lim_{t \to 1} \frac{m_k(t)}{1 - t} = m_1(0)^k, \quad k \geq 1. \tag{2-10}$$

In essence, this theorem guarantees that, for any compactly supported initial positive measure μ_0, the flow governed by the Hopf equation preserves analyticity and admits a representation as a Cauchy transform throughout its evolution.

Remark 2.2. (a) Using the change of variables $w = u(s, 0)$ we can rewrite (2-6)–(2-7) as

$$u^{-1}(w, t) = u^{-1}(w, 0) - \frac{t}{w}, \tag{2-11}$$

where u^{-1} is the functional inverse of u, or equivalently as

$$w = u\Big(z + \frac{t}{w}, 0\Big), \tag{2-12}$$

which is an equation on $w = u(z, t)$.

(b) The assertion in (i) of Theorem 2.1 about R depending only on r and T is in general false for a unit ($m_0 = 1$) but signed measure. A simple example is the case

$$\mu = (a + 1)\delta_0 - a\delta_1, \quad a > 0,$$

for which

$$u(z, 0) = \frac{a+1}{z} - \frac{a}{z-1} \tag{2-13}$$

[3]Recall that μ_t itself is not necessarily uniquely defined by its moments.

is holomorphic in Ω_1, but

$$u(z,t)=\frac{(1-2t)z-(a+1-t)+\sqrt{z^2-2(t-1+a(1-2t))z+(a+1-t)^2}}{2(z-1)z}$$

has branch points at $\zeta_+(t)=\overline{\zeta_-(t)}$, with $|\zeta_\pm|=a+1-t$. This shows that in this case there is no bounded R that works for all values of the parameter $a>0$.

(c) We can formally define the derivative with respect to t of the measures μ_t:

$$\tau_t:=-\lim_{\varepsilon\to 0+}\frac{1}{\varepsilon}(\mu_{t+\varepsilon}-\mu_t).$$

By (iii) of Theorem 2.1, $\tau_t(\mathbb{C})=1$ for all $t\in(0,1)$. In the terminology of [Kerov 1998], Theorem 2.1 states that τ_t is the *inverse Markov transform* of the measure μ_t.

(d) Formula (2-10) can be interpreted as follows: μ_t disappears, as $t\to 1-$, as a mass point at $m_1(0)$, i.e., at the center of mass of the initial distribution μ_0. This is the case, at least, when $\operatorname{supp}\mu_0\subset\mathbb{R}$.

Recall that the Hilbert transform of a measure μ on $\mathbb{R}$ is defined by

$$\mathcal{H}_\mu(x):=\frac{1}{\pi}\,\text{p.v.}\int\frac{d\mu(y)}{x-y}. \tag{2-14}$$

Theorem 2.1 has the following consequence when the initial data is supported on the real line: the flow remains uniquely determined and real, and satisfies a nonlinear nonlocal transport equation.

Corollary 2.3. *Let μ_0 be a compactly supported probability measure on $\mathbb{R}$, and let μ_t, $t\in[0,1)$, be as described in* (iii) *of Theorem 2.1. Then, for each $t\in[0,1)$, the measure μ_t can also be taken as supported on $\mathbb{R}$, and under this condition it is unique. If μ_t is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ then the density*

$$f(x,t):=\frac{d\mu_t(x)}{dx}$$

satisfies

$$\partial_t f(x,t)=-\frac{1}{\pi}\partial_x\arctan\left(\frac{\mathcal{H}_{\mu_t}(x)}{f(x,t)}\right), \tag{2-15}$$

where H_{μ_t} is its Hilbert transform.

The "nonlocal transport equation" (2-15) coincides with the one found in [Steinerberger 2019].

The initial value problem for the Hopf equation, whose existence and analyticity have been established in Theorem 2.1, is directly related to the central problem of this paper, on the flow of zeros of a sequence of polynomials under iterated differentiation, as explained in the Introduction.

Namely, recall that for a sequence of monic polynomials Q_n, with $\deg Q_n=n$, $n\in\mathbb{N}$, and the associated polynomials (1-2), we defined the normalized zero-counting measures (1-3) such that

$$\sigma_{n,k}(\mathbb{C})=\frac{n-k}{n}.$$

The only assumption on $\{Q_n\}$ is that there exists $0<r<+\infty$ such that, for each sufficiently large $n\in\mathbb{N}$, all zeros of Q_n belong to $\mathbb{D}_r$, and by the Gauss–Lucas theorem the same property is shared by all

polynomials $Q_{n,k}$, defined in (1-2). Equivalently, we can say that, for all sufficiently large $n \in \mathbb{N}$, the measures $\sigma_{n,k},\ k \in \{0, 1, \dots, n-1\}$, are supported in $\mathbb{D}_r$.

A key consequence of Theorem 2.1 is the existence of a universal value $R = R(T) > 2r$ such that Hopf's equation (2-1) has a holomorphic and uni-valued solution in Ω_R for any $0 < t < T$ and for *any* initial condition given by the Cauchy transform of *any* probability measure supported in $\overline{\mathbb{D}}_r$. In particular, if for each $n \in \mathbb{N}$ we denote by $u_n(z, t)$ the solution of (2-1) with the initial condition

$$u_n(z, 0) := \mathcal{C}^{\sigma_{n,0}}(z) = \frac{Q_n'(z)}{n Q_n(z)}, \tag{2-16}$$

then all $u_n(z, t)$ are guaranteed to exist and to be analytic in $\Omega_R \times [0, T]$.

As in (1-4), define for $k \in \{0, 1, \dots, n-1\},\ n \in \mathbb{N}$,

$$v_{n,k}(z) := \mathcal{C}^{\sigma_{n,k}}(z) = \frac{Q_{n,k}'(z)}{n Q_{n,k}(z)} = \frac{n-k}{n} \frac{Q_{n,k+1}(z)}{Q_{n,k}(z)}. \tag{2-17}$$

For all such k and n, $v_{n,k}$ is holomorphic in $\overline{\Omega}_r$, and

$$v_{n,k}(z) = \frac{n-k}{n} \frac{1}{z} + \mathcal{O}(z^{-2}), \quad z \to \infty. \tag{2-18}$$

Taking the logarithmic derivative on both sides of (2-17) we obtain that

$$v_{n,k+1}(z) - v_{n,k}(z) = \frac{1}{n} \frac{\partial_z v_{n,k}(z)}{v_{n,k}(z)}. \tag{2-19}$$

This identity can be regarded as a discrete version of Hopf's equation (2-1), and thus it is natural to expect that, under our assumptions, functions $v_{n,k}(z)$ and $u_n(z, k/n)$ are close for sufficiently large n.

If additionally

$$\sigma_{n,0} \xrightarrow{*} \mu_0, \quad n \to \infty, \tag{2-20}$$

where $\xrightarrow{*}$ denotes the weak-* convergence and μ_0 is a probability measure on $\mathbb{C}$ supported on $\overline{\mathbb{D}}_r$, then the initial conditions (2-16) converge to the initial condition (2-5), at least in the fixed neighborhood of infinity, and we can expect that the solutions $u_n(z, t)$ of equation (2-1) with initial conditions (2-16) will converge there to $u(z, t)$, the solution of the same equation (2-1) but with initial condition (2-5). These heuristic arguments can be made rigorous, although the trivial-looking assertion about the proximity of the solutions to the "discrete" and "continuous" Hopf equations requires some careful analysis due to the nonlinear character of these equations. For convenience of the reader, we formulate Theorem 1.1 here in a more precise form:

Theorem 2.4. *Let Q_n be a sequence of polynomials with uniformly bounded zeros such that* (2-20) *holds. Then, for any $t \in [0, 1)$, there exists a positive Borel measure μ_t with* $\operatorname{supp} \mu_t \subset \operatorname{conv}(\operatorname{supp} \mu_0)$ *and $\mu_t(\mathbb{C}) = 1 - t$ such that in a neighborhood of infinity*

$$\mathcal{C}^{\mu_t}(z) = u(z, t), \tag{2-21}$$

where u is the solution of (2-1) *with the initial condition* (2-5).

Furthermore, for any sequence k_n of natural numbers with $k_n/n \to t$ we have that

$$\lim_n \mathcal{C}^{\sigma_{n,k_n}}(z) = \mathcal{C}^{\mu_t}(z)$$

locally uniformly in $\mathbb{C} \setminus (\mathrm{conv}(\mathrm{supp}\,\mu_0))$.

Finally, if all zeros of Q_n are real, then μ_t is uniquely determined from u, and

$$\sigma_{n,k_n} \xrightarrow{*} \mu_t, \quad n \to \infty.$$

Remark 2.5. In the statement above, $\mathrm{conv}(K)$ denotes the convex hull of the set K.

With the assumptions of the theorem, let $T = \frac{1}{2}(1+t)$, and let $R = R(T)$ be as in Theorem 2.1. Then to establish Theorem 2.4, it is sufficient to prove that

$$\lim_n \mathcal{C}^{\sigma_{n,k_n}}(z) = u(z,t)$$

locally uniformly in $\Omega_R = \{z \in \mathbb{C} : R < |z| < +\infty\} = \mathbb{C} \setminus \overline{\mathbb{D}}_R$. Indeed, by the weak-* compactness of the sequence σ_{n,k_n}, there exists a measure μ_t, with the properties indicated above, such that $\sigma_{n,k_n} \xrightarrow{*} \mu_t$ along a subsequence of $\mathbb{N}$, and thus (2-21) holds in Ω_R. The rest of the assertion follows by elementary analytic continuation arguments.

As a consequence, the Cauchy transform of the measure μ_t is completely determined in the complement of $\mathrm{conv}(\mathrm{supp}\,\mu_0)$. However, in its maximum generality, this does not determine measure μ_t uniquely, which makes the statement of Theorem 2.4 apparently less straightforward.

Finally, we recast Theorem 2.4 in the terminology of free probability; see e.g., [Arizmendi et al. 2023; Marcus et al. 2022; Mingo and Speicher 2017; Nica and Speicher 2006] for a background. Along with the Cauchy transform $\mathcal{C}^{\mu}$ of a compactly supported probability measure μ, as in (2-2), we can define its *R-transform* implicitly by

$$\mathcal{R}_\mu(\mathcal{C}^\mu(z)) = z - \frac{1}{\mathcal{C}^\mu(z)}; \tag{2-22}$$

in particular, $\mathcal{R}_\mu$ is an analytic function in a neighborhood of the origin. For $s > 0$, the *fractional free additive convolution* $\mu^{\boxplus s}$ of μ is defined via the identity (see [Shlyakhtenko and Tao 2022])

$$\mathcal{R}_{\mu^{\boxplus s}}(w) = s\mathcal{R}_\mu(w). \tag{2-23}$$

Corollary 2.6. *Let Q_n be a sequence of polynomials as in Theorem 2.4 such that* (2-20) *holds, where μ_0 is a compactly supported probability measure on $\mathbb{R}$. For $t \in [0,1)$, define*

$$\widehat{Q}_{n,k}(x) := Q_{n,k}((1-t)x),$$

where we use the notation (1-2).

Then, for $k/n \to t$,

$$\frac{1}{n-k}\chi(\widehat{Q}_{n,k}) \xrightarrow{*} \mu_0^{\boxplus 1/(1-t)}. \tag{2-24}$$

As mentioned in the Introduction, (2-24) — established heuristically in [Steinerberger 2023] — was rigorously proven in [Hoskins and Kabluchko 2023]. A short proof was presented in [Arizmendi et al. 2023],

based entirely on arguments from free probability. We will show that this result is an immediate consequence of solving Hopf's equation by the method of characteristics.

3. Proof of Theorem 2.1 and Corollary 2.3

We start by addressing the evolution in a neighborhood of infinity of the Cauchy transform (2-2) of a finite Borel *signed* measure μ, compactly supported on $\mathbb{C}$, according to the Hopf equation (2-1). Recall that $\mathcal{C}^\mu$ is a holomorphic function in $\mathbb{C}\setminus \operatorname{supp}(\mu)$, and that the series in (2-3) converges locally uniformly in any $\Omega_r = \{z\in\mathbb{C} : r < |z| < +\infty\}$, disjoint from $\operatorname{supp}(\mu)$. We assume additionally that

$$m_0 = \mu(\mathbb{C}) = 1 \tag{3-1}$$

and denote by $|\mu|$ the total variation of μ; since μ is signed, in general, $|\mu|(\mathbb{C}) \geq m_0$.

As a first step, we need some uniform estimates on $\mathcal{C}^\mu$:

Proposition 3.1. *Let μ be a finite Borel and compactly supported signed measure on $\mathbb{C}$ such that* (3-1) *holds, and let $0 < r < +\infty$ be such that* $\operatorname{supp}(\mu)\subset \mathbb{D}_r$. *Then, for any $R > (1+|\mu|(\mathbb{C}))r$, $\mathcal{C}^\mu$ is analytic in $\overline{\Omega}_R$ and does not vanish for $z\in\Omega_R$.*

If in addition μ is a positive (*probability*) *measure, then for $|z|\geq R > 2r$*

$$\frac{2}{9} < \frac{1-r/R}{(1+r/R)^2} \leq |z\mathcal{C}^\mu(z)| \leq \frac{R}{R-r}, \tag{3-2}$$

$$\frac{1-2r/R}{(1+r/R)^4} \leq \left|z^2\frac{d}{dz}\mathcal{C}^\mu(z)\right| \leq \frac{1}{(1-r/R)^2} < 4. \tag{3-3}$$

Moreover, if $R > (\sqrt{2}+1)r$, then C^μ is univalent (*injective*) *in Ω_R, and*

$$\overline{\mathbb{D}}_{1/(4R)} \subset C^\mu(\overline{\Omega}_R) \subset \overline{\mathbb{D}}_{1/(R-r)}. \tag{3-4}$$

Remark 3.2. As formula (2-13) in Remark 2.2(b) shows, for a signed measure μ, R in the first assertion of Proposition 3.1, in general, cannot be made independent of $|\mu|$.

Proof. For an arbitrary $R > r$, the function $z\mathcal{C}^\mu(z)$ is analytic in $\overline{\Omega}_R$ and can be written as

$$z\mathcal{C}^\mu(z) = 1 + \int \frac{y}{z-y}\, d\mu(t).$$

Since for $|z| = R$

$$\left|\int \frac{y}{z-y}\, d\mu(t)\right| \leq \frac{r}{R-r}|\mu(\mathbb{C})|,$$

if $R > r$ is such that

$$\frac{r}{R-r}|\mu|(\mathbb{C}) < 1, \quad \text{or equivalently } \ R > r(1+|\mu|(\mathbb{C})),$$

then by Rouché's theorem, $\mathcal{C}^\mu(z)\neq 0$ in $\overline{\Omega}_R$. In particular, if μ is a probability measure, this is true for any $R > 2r$.

Assuming that μ is a probability measure with $\operatorname{supp}(\mu) \subset \mathbb{D}_r$, we have $|\mu|(\mathbb{C}) = 1$, and the bound obtained above gives us that, for $|z| = R > 2r$,

$$|z\mathcal{C}^\mu(z)| = \left|1 + \int \frac{y}{z-y}\, d\mu(t)\right| \le 1 + \frac{r}{R-r} = \frac{R}{R-r}.$$

On the other hand,

$$|z\mathcal{C}^\mu(z)| \ge \operatorname{Re}(z\mathcal{C}^\mu(z)) = \int \operatorname{Re}\left(\frac{1}{1-y/z}\right) d\mu(y),$$

which gives us that, for $|z| = R > 2r$ and $|y| \le r < R$,

$$\operatorname{Re}\left(\frac{1}{1-y/z}\right) = \frac{1-\operatorname{Re}(y/z)}{|1-y/z|^2} \ge \frac{1-r/R}{(1+r/R)^2} > \frac{2}{9},$$

and the bounds in (3-2) follow.

In a similar fashion,

$$\left|z^2 \frac{d}{dz}\mathcal{C}^\mu(z)\right| = \left|\int \frac{1}{(1-y/z)^2}\, d\mu(t)\right| \le \frac{1}{(1-r/R)^2} < 4$$

and

$$\left|z^2 \frac{d}{dz}\mathcal{C}^\mu(z)\right| \ge \operatorname{Re}\left(-z^2 \frac{d}{dz}\mathcal{C}^\mu(z)\right) = \int \operatorname{Re}\frac{1}{(1-y/z)^2}\, d\mu(y).$$

For $|z| \ge R > 2r$ and $y \in \mathbb{D}_r$,

$$\begin{aligned}
\operatorname{Re}(1-y/z)^2 &= 1 - 2\operatorname{Re}\left(\frac{y}{z}\right) + \operatorname{Re}\left(\frac{y}{z}\right)^2 = 1 - 2\operatorname{Re}\left(\frac{y}{z}\right) + \left(\operatorname{Re}\frac{y}{z}\right)^2 - \left(\operatorname{Im}\frac{y}{z}\right)^2 \\
&= \left(1 - \operatorname{Re}\frac{y}{z}\right)^2 - \left(\operatorname{Im}\frac{y}{z}\right)^2 \ge \left(1 - \operatorname{Re}\frac{y}{z}\right)^2 - \left|\frac{y}{z}\right|^2 \ge \left(1 - \frac{r}{R}\right)^2 - \left(\frac{r}{R}\right)^2 \\
&= 1 - \frac{2r}{R} > 0,
\end{aligned}$$

so that

$$\operatorname{Re}\frac{1}{(1-y/z)^2} = \frac{\operatorname{Re}(1-y/z)^2}{|1-y/z|^4} \ge \frac{1-2r/R}{(1+r/R)^4},$$

and hence

$$\left|z^2 \frac{d}{dz}\mathcal{C}^\mu(z)\right| \ge \frac{1-2r/R}{(1+r/R)^4}, \quad |z| \ge R > 2r.$$

This concludes the proof of (3-3).

Furthermore, for $R > r$, consider the function

$$f_R(z) := R\mathcal{C}^\mu\left(\frac{R}{z}\right) = z \int \frac{1}{1-zy/R}\, d\mu(y) = z + \mathcal{O}(z^2), \quad z \to 0, \tag{3-5}$$

defined in the unit disk $\mathbb{D}_1 = \{z \in \mathbb{C} : |z| < 1\}$. We have that

$$\operatorname{Re}\left(\frac{f_R(z)}{z}\right) = \int \operatorname{Re}\left(\frac{1}{1-zy/R}\right) d\mu(y) = \int \frac{1-\operatorname{Re}(zy)/R}{|1-zy/R|^2}\, d\mu(y).$$

Since

$$1-\frac{\operatorname{Re}(zy)}{R}\geq 1-\frac{|zy|}{R}\geq 1-\frac{r}{R}>0,$$

we conclude that $\operatorname{Re}(f_R(z)/z)>0$ in $|z|<1$, and by [MacGregor 1962, Theorem 3], f_R is univalent in $|z|<\sqrt{2}-1$ (or, equivalently, $\mathcal{C}^\mu$ is univalent in $\Omega_{(\sqrt{2}+1)R}$).

Moreover, with $R>(\sqrt{2}+1)r$, by the Koebe $1/4$ theorem (see, e.g., [Duren 1983, §2.2]), the image $f_R(\mathbb{D}_1)$ of the unit disk by the function f_R contains the disk $\overline{\mathbb{D}}_{1/4}$, or, in other words, the image of $\overline{\Omega}_R$ by $\mathcal{C}^\mu$ contains the disk $\overline{\mathbb{D}}_{1/(4R)}$. □

It is well known that the Hopf equation (2-1) can be solved by the method of characteristics. In preparation for the analysis of this solution, we prove the following auxiliary result:

Proposition 3.3. *Let μ be a finite Borel signed measure μ on $\mathbb{C}$ such that* (3-1) *holds. Assume $0<r<+\infty$ is such that* $\operatorname{supp}\mu\subset\mathbb{D}_r$, *and let $0<T<1$ be fixed. Then there exists an $R>2r$, depending only on the values of T, $|m_1|$, and $|\mu|(\mathbb{C})$, such that*

$$f_1(s)=\mathcal{C}^\mu\Big(\frac{1}{s}\Big)=\int\frac{s}{1-sy}\,d\mu(y)$$

is analytic in $\overline{\mathbb{D}}_{1/R}$, does not vanish for $0<|s|\leq 1/R$, and, for each $z\in\mathbb{D}_{1/(2R)}$ and each $t\in[0,T]$, the equation

$$z-s-\frac{zst}{f_1(s)}=0 \tag{3-6}$$

has a unique solution $s\in\mathbb{D}_{1/R}$.

Proof. By Proposition 3.1, if $R>r(1+|\mu|(\mathbb{C}))$ then $f_1(z)$ is analytic and nonvanishing for $0<|z|<1/R$. The assertion about the unisolvence of (3-6) is obviously true for $z=0$, so we assume that $z\neq 0$.

Write

$$f_1(s)=s+s^2g(s), \tag{3-7}$$

with

$$g(s):=\int\frac{y}{1-sy}\,d\mu(y), \tag{3-8}$$

so that

$$g(0)=\int y\,d\mu(y)=m_1,$$

the first moment of μ; see (2-4). Notice that

$$|g(s)-g(0)|=\left|\int\frac{y}{1-sy}\,d\mu(y)-\int y\,d\mu(y)\right|=|s|\left|\int\frac{y^2}{1-sy}\,d\mu(y)\right|,$$

and for $|s|=1/R$

$$|g(s)-g(0)|=\frac{1}{R}\left|\int\frac{y^2}{1-sy}\,d\mu(y)\right|\leq\frac{r^2}{R-r}|\mu|(\mathbb{C}).$$

If we take $R>r(1+|\mu|(\mathbb{C}))$ such that

$$\frac{r^2}{R-r}|\mu|(\mathbb{C})<1,\quad\text{or equivalently}\quad R>r(1+r|\mu|(\mathbb{C})),$$

then

$$|g(s)| \leq 1 + |m_1|, \quad s \in \overline{\mathbb{D}}_{1/R}. \tag{3-9}$$

Now, given $T \in (0, 1)$, let $R > \max\{r(1 + |\mu|(\mathbb{C})), r(1 + r|\mu|(\mathbb{C}))\}$ be such that

$$R > (1 + |m_1|)\frac{1+T}{1-T}, \quad \text{or equivalently } 0 < \frac{T}{1 - (|m_1| + 1)/R} < \frac{1+T}{2} < 1.$$

Then, for $s \in \mathbb{D}_{1/R}$,

$$0 \leq \frac{|g(s)|}{R} \leq \frac{1 + |m_1|}{R} < \frac{1-T}{1+T} < 1,$$

so that

$$\left|\frac{t}{1 + sg(s)}\right| \leq \frac{T}{|1 - |g(s)|/R|} \leq \frac{T}{1 - |g(s)|/R} < \frac{1+T}{2} < 1, \tag{3-10}$$

and as a consequence, for $z \neq 0$ and $|s| = 1/R$,

$$\left|\frac{zst}{f_1(s)}\right| = \left|\frac{zt}{1 + sg(s)}\right| < |z|. \tag{3-11}$$

Thus, since for $0 < |z| < 1/(2R)$ and $|s| = 1/R$ we have

$$|z - s| > \frac{1}{2R} > |z|,$$

we can apply Rouché's theorem to claim that, for every $z \in \mathbb{D}_{1/(2R)} \setminus \{0\}$, equation (3-6) has a unique solution.

All assertions have been proved for any

$$R > \max\left\{r(1 + |\mu|(\mathbb{C})),\ r(1 + r|\mu|(\mathbb{C})),\ (1 + |m_1|)\frac{1+T}{1-T}\right\} \geq 2r,$$

which concludes the proof. □

Remark 3.4. Under the assumption of Proposition 3.3, for a signed measure μ, both $|\mu|$ and $|m_1|$ can be arbitrarily large even though $m_0 = \mu(\mathbb{C}) = 1$.

However, if μ is a probability measure, then $|\mu|(\mathbb{C}) = m_0 = 1$ and

$$\operatorname{supp}(\mu) \subset \mathbb{D}_r \quad \Longrightarrow \quad |m_1| = \left|\int t\, d\mu(t)\right| \leq r,$$

showing that we can take

$$R = R(T) > \max\left\{2r,\ r(1 + r),\ (1 + r)\frac{1+T}{1-T}\right\}, \tag{3-12}$$

which depends only on the values of T and r.

Now we are ready to prove the first two statements of Theorem 2.1.

Proof of statements (i)–(ii) *of Theorem 2.1.* Let $0 < T < 1$ and $0 < r < +\infty$ be fixed and take $R = R(T) > 2r$, satisfying (3-12). Consider equation (2-1) with initial condition (2-5), where μ_0 is a probability measure supported by $\mathbb{D}_r$. We solve (2-1) using the standard method of characteristics; see, e.g., [Salih 2018; Whitham 1974].

If $z = z(t)$ is differentiable, then

$$\frac{d}{dt}u(z(t), t) = \frac{\partial u}{\partial t} + z'(t)\frac{\partial u}{\partial z}.$$

Comparing it with (2-1) we see that if $z(t)$ satisfies

$$z'(t) = -\frac{1}{u(z(t), t)}, \tag{3-13}$$

then for a solution u of (2-1) we should have

$$\frac{d}{dt}u(z(t), t) = 0. \tag{3-14}$$

Differentiating (3-13) with respect to t again and using (3-14) we conclude that $z''(t) \equiv 0$, so that $z(t)$ must be linear in t: $z(t) = Lt + s$.

To find the slope L, we use the initial values,

$$L = z'(0) = -\frac{1}{u(z(0), 0)} = -\frac{1}{u(s, 0)},$$

so that $z(t) = s - t/u(s, 0)$. This is the characteristics line, along which $u(z, t)$ is constant, with the constant given by $u(z(0), 0) = u(s, 0)$. Thus, for (z, t), the solution of (2-1) is $u(z, t) = u(s, 0)$, where $s = s(z)$ is the solution of

$$z - s + \frac{t}{u(s, 0)} = 0,$$

that is, of $F(z, s) = 0$. Making the change of variables $z \mapsto 1/z$ and $s \mapsto 1/s$, this is equivalent to writing

$$u\Big(\frac{1}{z}, t\Big) = u\Big(\frac{1}{s}, 0\Big), \tag{3-15}$$

where $s = s(z)$ is defined implicitly by

$$z - s - \frac{zst}{u(1/s, 0)} = 0. \tag{3-16}$$

Thus, it remains to apply Proposition 3.3 with $f_1(z) = u(1/z, 0)$ to conclude the proof. □

Now we turn to the proof of statement (iii) of Theorem 2.1. The fact that there exists a measure μ_t such that (2-8) holds will follow directly from Theorem 2.4. Hence we only need to establish the second part of the statement, on the properties of the moments $m_k(t)$. Notice that the fact that they are uniquely determined by the initial data $\{m_k(0)\}_{k\geq 0}$ is also a consequence of the uniqueness of the solution $u(z, t)$, already proved before.

Proof of statement (iii) *of Theorem 2.1.* Assume that μ_0 is a probability measure compactly supported in the disk $\mathbb{D}_r$, $0 < T < 1$, and take $R = R(T) > 2r$ satisfying (3-12). The unique solution $u(z, t)$ of

equation (2-1) for $t \in [0, T]$, satisfying the initial condition (2-5), has the Laurent expansion, convergent in Ω_R, given by

$$u(z,t) = \sum_{k=0}^{\infty} \frac{m_k(t)}{z^{k+1}}, \quad |z| > R, \tag{3-17}$$

with $m_0(0) = 1$. Replacing this expansion in (2-1), rewritten as

$$\partial_z u = u\, \partial_t u, \tag{3-18}$$

we get an infinite system of equations

$$\sum_{j=0}^{k} m_{k-j}(t) m_j'(t) = -(k+1) m_k(t), \quad k = 0, 1, \ldots, \tag{3-19}$$

where we have written

$$m_k'(t) = \frac{d}{dt} m_k(t).$$

For $k = 0$, we immediately get $m_0'(t) = -1$, so that $m_0(t) = 1 - t$. For $k = 1$, we obtain $m_1(t) = m_1(0)(1-t)$, so that (2-10) is valid for $k = 1$.

We conclude the proof by induction, assuming that the assertion is established for $m_1(t), \ldots, m_{k-1}(t)$, with $k \geq 2$. We need to show that $m_k(t)$ is a polynomial in t of degree at most k, whose coefficients depend only on the initial moments $m_j(0)$, $0 \leq j \leq k$, and such that

$$m_k(1) = 0, \quad -m_k'(1) = \lim_{t \to 1} \frac{m_k(t)}{1-t} = m_1(0)^k.$$

Let us rewrite (3-19) as

$$(1-t) m_k'(t) + k m_k(t) = -\sum_{j=1}^{k-1} m_{k-j}(t) m_j'(t) = -f_k'(t), \tag{3-20}$$

where

$$f_k(t) := \frac{1}{2} \sum_{j=1}^{k-1} m_{k-j}(t) m_j(t).$$

Multiplying both sides of (3-20) by the integrating factor $(1-t)^{-k-1}$ we get

$$\left(\frac{m_k(t)}{(1-t)^k} \right)' = -\frac{f_k'(t)}{(1-t)^{k+1}},$$

so that, taking into account the initial conditions,

$$m_k(t) = (1-t)^k m_k(0) - (1-t)^k \int_0^t \frac{f_k'(u)}{(1-u)^{k+1}}\, du. \tag{3-21}$$

Notice that

$$f_2(t) = \tfrac{1}{2} m_1(t)^2 = \tfrac{1}{2} m_1(0)^2 (1-t)^2, \quad f_2''(1) = m_1(0)^2,$$

and by the induction hypothesis, f_k is a polynomial of degree at most k in t, with

$$\begin{aligned} &f_k(1) = f_k'(1) = 0, \\ &f_k''(1) = \lim_{t\to 1} \frac{2f_k(t)}{(1-t)^2} = -\lim_{t\to 1} \frac{f_k'(t)}{1-t} = \sum_{j=1}^{k-1} m_{k-j}'(1) m_j'(1) = (k-1)m_1(0)^k. \end{aligned} \tag{3-22}$$

Integrating in (3-21) by parts k times and using that $f^{(k+1)} \equiv 0$, we get

$$\int_0^t \frac{f_k'(u)}{(1-u)^{k+1}}\,du = \left(\sum_{j=1}^{k} \frac{(-1)^{j-1}(k-j)!}{k!} \frac{f_k^{(j)}(u)}{(1-u)^{k+1-j}}\right)\Bigg|_0^t,$$

so that

$$-(1-t)^k\left[\int_0^t \frac{f_k'(u)}{(1-u)^{k+1}}\,du\right] = (1-t)^k d + \sum_{j=1}^{k} \frac{(-1)^j (k-j)!}{k!} f_k^{(j)}(t)(1-t)^{j-1},$$

where d is a constant:

$$d = \sum_{j=1}^{k} \frac{(-1)^{j-1}(k-j)!}{k!} f_k^{(j)}(0).$$

Combining all in (3-21), we conclude that

$$\begin{aligned} m_k(t) &= m_k(0)(1-t)^k + \sum_{j=1}^{k} (-1)^j \frac{(k-j)!}{k!}\big[f_k^{(j)}(t) - f_k^{(j)}(0)(1-t)^{k-j+1}\big](1-t)^{j-1} \\ &= -\frac{1}{k} f_k'(t) + \frac{1}{k(k-1)} f_k''(t)(1-t) + \mathcal{O}((1-t)^2), \quad t \to 1. \end{aligned}$$

By (3-22), $f_k'(1) = 0$, so that $m_k(1) = 0$, and

$$-m_k'(1) = \lim_{t\to 1} \frac{m_k(t)}{1-t} = \lim_{t\to 1}\left(-\frac{1}{k}\frac{f_k'(t)}{1-t} + \frac{1}{k(k-1)} f_k''(t)\right) = \frac{f_k''(1)}{k-1} = m_1(0)^k,$$

which proves the assertion. □

Proof of Corollary 2.3. Assume that μ_t is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ with density

$$f(x,t) := \frac{d\mu_t(x)}{dx}.$$

The Sokhotski–Plemelj formulas

$$(\mathcal{C}^\mu)_\pm(x) = \lim_{y\to 0+} \mathcal{C}^\mu(x \pm iy) = -\pi(\mathcal{H}_\mu(x) \pm i\mu'(x)), \quad x \in \mathbb{R}, \tag{3-23}$$

where $\mathcal{H}_\mu$ is the Hilbert transform defined in (2-14), imply that

$$u_\pm(x,t) = -\pi(1-t)(\mathcal{H}_{\mu_t}(x) \pm if).$$

Using this in (2-1) and subtracting both boundary values, we get the identity

$$2\pi i\,\partial_t[(1-t)f(x,t)] = 2i\,\partial_x\left(\frac{f(x,t)}{\mathcal{H}_{\mu_t}(x)}\right)\frac{1}{1+\left(\frac{f(x,t)}{\mathcal{H}_{\mu_t}(x)}\right)^2},$$

or

$$\partial_t[(1-t)f(x,t)] = \frac{1}{\pi}\partial_x \arctan\left(\frac{f(x,t)}{\mathcal{H}_{\mu_t}(x)}\right) = -\frac{1}{\pi}\partial_x \arctan\left(\frac{\mathcal{H}_{\mu_t}(x)}{f(x,t)}\right),$$

which proves (2-15). □

4. Proof of Theorem 2.4 and Corollary 2.6

The triangular table of monic polynomials $Q_{n,k}$, $\deg Q_{n,k} = n-k$, $n \in \mathbb{N}$,

$$Q_{n,k}(z) = \frac{(n-k)!}{n!}\frac{d^k}{dz^k}Q_n(z), \quad k \in \{0, 1, \ldots, n-1\},$$

was introduced in (1-2) along with their zero-counting measures $\sigma_{n,k}$ in (1-3).

The heuristic arguments supporting Theorem 2.4 were presented in Section 2, where we also introduced the functions $u_n(z,t)$, the solution of Hopf's equation (2-1) with the initial condition (2-16). As explained, our goal is to show that

$$u_{n,k}(z) := u(z, k/n)$$

and $v_{n,k}$, defined in (2-17), are asymptotically close. The solution of the Hopf equation is characterized by the identity (2-11), which motivates us to study the behavior of their functional inverses.

Since $v_{n,k}$ is the Cauchy transform of a measure of size $1-k/n$, by Proposition 3.1, the inverse function $v_{n,k}^{-1}(\zeta)$ is guaranteed to exist for all $(\zeta, k/n) \in \mathbb{D}_\rho \times [0,T]$, $0 < T < 1$, where $\rho = (1-T)/(4R)$ and $R > (\sqrt{2}+1)r$; moreover, $v_{n,k}^{-1}(\mathbb{D}_\rho) \subset \Omega_R$.

On the other hand, recall that

$$u_n(z,t) = u_n(s,0) = \mathcal{C}^{\sigma_{n,0}}(s) = \frac{Q_n'(s)}{nQ_n(s)},$$

and by Proposition 3.3, the inverse function $u_{n,k}^{-1}(\zeta)$ is guaranteed to exist for all $(\zeta, k/n) \in \mathbb{D}_{1/(4R)} \times [0,T]$, $0 < T < 1$, where $R > (\sqrt{2}+1)r$; also, $u_{n,k}^{-1}(\mathbb{D}_\rho) \subset \Omega_R$.

In summary, given $R > (\sqrt{2}+1)r$ and $0 < T < 1$, there exists a $0 < \rho < (1-T)/(4R)$ such that we can consider the inverse functions $u_{n,k}^{-1}(\zeta)$ and $v_{n,k}^{-1}(\zeta)$, conformal mappings of $\mathbb{D}_\rho$ into Ω_R for every $k/n \in [0,T]$. We fix these values of R and ρ for the rest of the proof.

Proposition 4.1. *For the inverse functions $v_{n,k}^{-1}(\zeta)$ and $v_{n,k+1}^{-1}(\zeta)$, with $\zeta \in \mathbb{D}_\rho$, we have*

$$v_{n,k+1}^{-1}(\zeta) - v_{n,k}^{-1}(\zeta) = -\frac{1}{n\zeta} + \Delta_{n,k}(\zeta). \tag{4-1}$$

There is $\rho' < \rho$ and $M > 0$, depending only on r, such that the $\Delta_{n,k}$ are analytic in $\mathbb{D}_{\rho'}$ and

$$|\Delta_{n,k}(\zeta)| \le \frac{M}{n^2}, \quad \zeta \in \mathbb{D}_{\rho'}. \tag{4-2}$$

Proof. Notice that, by (2-18),

$$v_{n,k+1}^{-1}(\zeta) - v_{n,k}^{-1}(\zeta) + \frac{1}{n\zeta}$$

is analytic at $\zeta = 0$, so it is enough to establish the bound (4-2) on a circle $|\zeta| = \rho'$.

To simplify the notation, define $\varphi_{n,k}(\zeta) := v_{n,k}^{-1}(\zeta)$ and let $'$ stand for the derivative with respect to the corresponding variable that should be clear in each context: $v'_{n,k}(z) = \partial_z v_{n,k}(z)$, $\varphi'_{n,k}(\zeta) = \partial_\zeta v_{n,k}^{-1}(\zeta)$, etc. We also define $z_k := \varphi_{n,k}(\zeta)$ and $z_{k+1} := \varphi_{n,k+1}(\zeta)$.

Substituting $z_k = \varphi_{n,k}(\zeta)$ in (2-19) and using that $v'_{n,k}(z_k) = 1/\varphi'_{n,k}(\zeta)$, we obtain

$$v_{n,k+1}(\varphi_{n,k}(\zeta)) = \zeta + \frac{1}{n\zeta}\frac{1}{\varphi'_{n,k}(\zeta)}.$$

Applying now $\varphi_{n,k+1} = v_{n,k+1}^{-1}$ to both sides yields

$$\varphi_{n,k}(\zeta) = \varphi_{n,k+1}\left(\zeta + \frac{1}{n\zeta}\frac{1}{\varphi'_{n,k}(\zeta)}\right),$$

and thus

$$\varphi_{n,k+1}(\zeta) - \varphi_{n,k}(\zeta) = \varphi_{n,k+1}(\zeta) - \varphi_{n,k+1}\left(\zeta + \frac{1}{n\zeta}\frac{1}{\varphi'_{n,k}(\zeta)}\right). \tag{4-3}$$

Using that, for small $s \in \mathbb{C}$,

$$\varphi_{n,k+1}(\zeta + s) = \varphi_{n,k+1}(\zeta) + \varphi'_{n,k+1}(\zeta)s + \tfrac{1}{2}\varphi''_{n,k+1}(\zeta + \theta)s^2, \tag{4-4}$$

we get that

$$\begin{aligned} z_{k+1} - z_k = \varphi_{n,k+1}(\zeta) - \varphi_{n,k}(\zeta) &= -\frac{1}{n\zeta}\frac{\varphi'_{n,k+1}(\zeta)}{\varphi'_{n,k}(\zeta)} - \frac{\delta_{n,k}(\zeta)}{n^2\zeta^2} \\ &= -\frac{1}{n\zeta} - \frac{1}{n\zeta}\frac{\varphi'_{n,k+1}(\zeta) - \varphi'_{n,k}(\zeta)}{\varphi'_{n,k}(\zeta)} - \frac{\delta_{n,k}(\zeta)}{n^2\zeta^2}. \end{aligned} \tag{4-5}$$

With the notation introduced above, let us rewrite

$$\frac{\varphi'_{n,k+1}(\zeta) - \varphi'_{n,k}(\zeta)}{\varphi'_{n,k}(\zeta)} = \frac{v'_{n,k}(z_k) - v'_{n,k+1}(z_{k+1})}{v'_{n,k+1}(z_{k+1})} = \frac{1}{n}\varepsilon_{n,k}^{(1)}(\zeta) - \varepsilon_{n,k}^{(2)}(\zeta)(u - z), \tag{4-6}$$

where, again by (2-19),

$$\begin{aligned} \varepsilon_{n,k}^{(1)}(\zeta) = n\frac{v'_{n,k}(z_{k+1}) - v'_{n,k+1}(z_{k+1})}{v'_{n,k+1}(z_{k+1})} &= -\frac{1}{v'_{n,k+1}(z_{k+1})}\left(\frac{v'_{n,k}(u)}{v_{n,k}(u)}\right)\Big|_{u=z_{k+1}} \\ &= -\frac{1}{z_{k+1}^2 v'_{n,k+1}(z_{k+1})} z_{k+1}^2\left(\frac{v'_{n,k}(u)}{v_{n,k}(u)}\right)'\Big|_{u=z_{k+1}} \end{aligned} \tag{4-7}$$

and

$$\varepsilon_{n,k}^{(2)}(\zeta) = \frac{v'_{n,k}(z_k) - v'_{n,k}(z_{k+1})}{z_k - z_{k+1}}\frac{1}{v'_{n,k+1}(z_{k+1})}.$$

Combining (4-5) and (4-6) we get

$$(z_{k+1}-z_k)\left(1-\frac{\varepsilon_{n,k}^{(2)}(\zeta)}{n\zeta}\right)=-\frac{1}{n\zeta}-\frac{\varepsilon_{n,k}^{(1)}(\zeta)}{n^2\zeta}-\frac{\delta_{n,k}(\zeta)}{n^2\zeta^2}. \tag{4-8}$$

To prove (4-2), it suffices to show that $\varepsilon_{n,k}^{(1)}(\zeta)$, $\varepsilon_{n,k}^{(2)}(\zeta)$, and $\delta_{n,k}(\zeta)$ are uniformly bounded for sufficiently small $|\zeta|$ (or equivalently sufficiently large z_k and z_{k+1}).

By (4-6) and Proposition 3.1, $\varepsilon_{n,k}^{(1)}(\zeta)$ is analytic for $|z_{k+1}|\geq R$, and the functions

$$u\frac{v_{n,k}'(u)}{v_{n,k}(u)} \quad \text{and} \quad u^2v_{n,k+1}'(u)$$

are analytic and nonvanishing in $|u|\geq R$. By (3-2)–(3-3), for $|u|\geq R$,

$$\left|u\frac{v_{n,k}'(u)}{v_{n,k}(u)}\right|=\left|\frac{u^2v_{n,k}'(u)}{uv_{n,k}(u)}\right|\leq 18,$$

and in consequence,

$$\left|\frac{v_{n,k}'(u)}{v_{n,k}(u)}\right|\leq\frac{18}{R}, \quad |u|=R.$$

By the Cauchy integral formula,

$$\left|u^2\left(\frac{v_{n,k}'(u)}{v_{n,k}(u)}\right)'\right|\leq 18(R+1)^2, \quad |u|=R+1. \tag{4-9}$$

On the other hand, by (3-3),

$$\left|\frac{1}{u^2v_{n,k+1}'(u)}\right|\leq\frac{1}{1-(k+1)/n}\frac{(1+r/R)^4}{1-2r/R}\leq\frac{4(3+2\sqrt{2})}{1-(k+1)/n}, \quad |u|\geq R\geq(\sqrt{2}+1)r. \tag{4-10}$$

Using these estimates in (4-7), we conclude that $|\varepsilon_{n,k}^{(1)}(\zeta)|$ is uniformly bounded for $|z_{k+1}|\geq R+1$ and $k/n\leq T<1$.

As for $\varepsilon_{n,k}^{(2)}(\zeta)$, we notice that for $z_k, z_{k+1}\in\Omega_R$

$$\left|\frac{v_{n,k}'(z_k)-v_{n,k}'(z_{k+1})}{z_k-z_{k+1}}\right|\leq\max_{s\in\Omega_R}|v_{n,k}''(s)|.$$

Appealing again to (3-2) and Cauchy's formula, we get the uniform boundedness of the function $\varepsilon_{n,k}^{(2)}(\zeta)$ for $z_k, z_{k+1}\in\Omega_{R+1}$.

Finally, by (4-4), $\delta_{n,k}(\zeta)$ in (4-5) satisfies

$$|\delta_{n,k}(\zeta)|\leq\frac{1}{2}\max_{\zeta\in\mathbb{D}_\rho}\frac{\varphi_{n,k+1}''(\zeta)}{(\varphi_{n,k}'(\zeta))^2}, \quad \zeta\in\mathbb{D}_\rho. \tag{4-11}$$

We have

$$\begin{aligned}\frac{\varphi_{n,k+1}''(\zeta)}{(\varphi_{n,k}'(\zeta))^2}=-\frac{v_{n,k+1}''(z_{k+1})}{(v_{n,k+1}'(z_{k+1}))^3}(v_{n,k}'(z_k))^2&=-\frac{v_{n,k+1}''(z_{k+1})}{v_{n,k+1}'(z_{k+1})}\left(\frac{v_{n,k}'(z_k)}{v_{n,k+1}'(z_{k+1})}\right)^2\\&=-\frac{v_{n,k+1}''(z_{k+1})}{v_{n,k+1}'(z_{k+1})}\left(1+\frac{\varepsilon_{n,k}(\zeta)}{n}\right)^2.\end{aligned}$$

Reasoning as above, by (3-2),

$$|v''_{n,k+1}(u)| \le \frac{4(1-(k+1)/n)}{R}, \quad |u| = R+1,$$

while, by (4-10),

$$\left|\frac{1}{v'_{n,k+1}(u)}\right| \le \frac{4(3+2\sqrt{2})R^2}{1-(k+1)/n}, \quad |u| \ge R \ge (\sqrt{2}+1)r,$$

showing that $\delta_{n,k}$ in (4-5) is uniformly bounded also. □

We are one step away from proving Theorem 2.4. Since $u(z,t)$ is the solution of (2-1) with the initial condition (2-5), it satisfies (2-11), which implies that

$$(u^{-1}_{n,k+1} - u^{-1}_{n,k})(\zeta) = -\frac{1}{n\zeta}.$$

Combining it with (4-1) we get that

$$(v^{-1}_{n,k+1}(\zeta) - u^{-1}_{n,k+1}(\zeta)) - (v^{-1}_{n,k}(\zeta) - u^{-1}_{n,k}(\zeta)) = \Delta_{n,k}(\zeta) = \mathcal{O}\Big(\frac{1}{n^2}\Big), \quad |\zeta| = \rho',$$

uniformly over $n, k \in \mathbb{N}$ such that $k/n < T$. Since $u^{-1}_{n,0} = v^{-1}_{n,0}$, we conclude that

$$u^{-1}_{n,k}(\zeta) - v^{-1}_{n,k}(\zeta) = \mathcal{O}\Big(\frac{1}{n}\Big) \tag{4-12}$$

locally uniformly in $\mathbb{D}_\rho$.

In order to complete the proof, it remains to show that (4-12) implies proximity of the direct functions; that is,

$$u_{n,k}(z) - v_{n,k}(z) = \mathcal{O}\Big(\frac{1}{n}\Big) \tag{4-13}$$

locally uniformly in Ω_R for some $R > r$. For that, it is more convenient to work with the functions

$$\hat{u}_{n,k}(z) := u_{n,k}\Big(\frac{1}{z}\Big), \quad \hat{v}_{n,k}(z,t) := v_{n,k}\Big(\frac{1}{z}, t\Big),$$

which are holomorphic in $\overline{\mathbb{D}}_{1/R}$ and nonvanishing in $\overline{\mathbb{D}}_{1/R} \setminus \{0\}$. Also,

$$\hat{u}'_{n,k}(0) = \hat{v}'_{n,k}(0) = 1 - \frac{k}{n} \ge 1 - T > 0,$$

and both $\hat{u}_{n,k}$ and $\hat{v}_{n,k}$ are uniformly bounded in $\overline{\mathbb{D}}_{1/R}$: there exists $M > 0$, depending on T and R only, such that

$$|\hat{u}_{n,k}(z)| \le M, \quad |\hat{v}_{n,k}(z)| \le M, \qquad z \in \overline{\mathbb{D}}_{1/R},$$

for every pair $(k,n) \in \mathbb{N} \cup \{0\} \times \mathbb{N}$ with $k/n \in [0,T]$. The assertion (4-13) is a consequence of the general facts gathered in Propositions A.1–A.3 that, for the sake of readability, we defer to the Appendix. This concludes the proof of Theorem 2.4.

Proof of Corollary 2.6. By (2-22)–(2-23),

$$\mathcal{R}_{\mu_0^{\boxplus 1/(1-t)}}(w) = \frac{1}{1-t}\mathcal{R}_{\mu_0}(w) = \frac{1}{1-t}[\mathcal{C}^{\mu_0}]^{-1}(w) - \frac{1}{(1-t)w},$$

where $[\mathcal{C}^{\mu}]^{-1}$ is the functional inverse of $\mathcal{C}^{\mu}$. Taking into account (2-11), we have that

$$\begin{aligned}\mathcal{R}_{\mu_0^{\boxplus 1/(1-t)}}(w) = \frac{1}{1-t}u^{-1}(w,0) - \frac{1}{(1-t)w} &= \frac{1}{1-t}\left(u^{-1}(w,t) + \frac{t}{w}\right) - \frac{1}{(1-t)w} \\ &= \frac{1}{1-t}u^{-1}(w,t) - \frac{1}{w},\end{aligned}$$

so that

$$\left[\mathcal{C}^{\mu_0^{\boxplus 1/(1-t)}}\right]^{-1}(w) = \frac{1}{1-t}u^{-1}(w,t),$$

or equivalently

$$\mathcal{C}^{\mu_0^{\boxplus 1/(1-t)}}(z) = \mathcal{C}^{\mu_t}((1-t)z) = u((1-t)z,t). \tag{4-14}$$

On the other hand, for

$$\hat{\sigma}_{n,k} := \frac{1}{n-k}\chi(\widehat{Q}_{n,k})$$

we have that

$$\mathcal{C}^{\hat{\sigma}_{n,k}}(z) = \frac{1}{n-k}\frac{\partial_z \widehat{Q}_{n,k}(z)}{\widehat{Q}_{n,k}(z)} = (1-t)\frac{1}{n-k}\frac{Q'_{n,k}(z)}{Q_{n,k}(z)}((1-t)z) = (1-t)\frac{n}{n-k}\mathcal{C}^{\sigma_{n,k}}((1-t)z),$$

with $\sigma_{n,k}$ defined in (1-3). Hence, for $k/n \to t$, by Theorem 2.4,

$$\lim_n \mathcal{C}^{\hat{\sigma}_{n,k}}(z) = \mathcal{C}^{\mu_t}((1-t)z),$$

so that, by (4-14),

$$\lim_n \mathcal{C}^{\hat{\sigma}_{n,k}}(z) = \mathcal{C}^{\mu_0^{\boxplus 1/(1-t)}}(z).$$

For measures on the real line, this is equivalent to (2-24). □

5. Some examples

We finish our exposition by presenting a few simple examples, starting with some paradigmatic zero distributions on the real line.

It is well known that under very mild assumptions on the orthogonality measure on the interval $[-1, 1]$ (such as having a strictly positive Radon–Nikodym derivative with respect to the Lebesgue measure almost everywhere on $[-1, 1]$; see, e.g., [Stahl and Totik 1992]), the limit zero distribution of the corresponding orthogonal polynomials is the equilibrium measure (or arcsine distribution) of the interval $[-1, 1]$; that is,

$$d\mu_0(x) = \frac{1}{\pi\sqrt{1-x^2}}\,dx, \quad x \in (-1, 1),$$

for which

$$u(z,0) = \mathcal{C}^{\mu_0}(z) = \frac{1}{\sqrt{z^2-1}} = \frac{1}{z} + \mathcal{O}\left(\frac{1}{z^2}\right), \quad z \in \mathbb{C} \setminus [-1, 1],$$

and consequently,

$$u^{-1}(w,0)=\frac{\sqrt{w^2+1}}{w}.$$

By (2-11), solving the equation

$$z=\frac{\sqrt{w^2+1}-t}{w}$$

we get that

$$w=u(z,t)=\frac{-tz+\sqrt{z^2-(1-t^2)}}{z^2-1}.$$

Since the residues of the right-hand side at ± 1 are 0, we conclude that μ_t is absolutely continuous and

$$d\mu_t(x)=\frac{1}{\pi}\frac{\sqrt{1-t^2-x^2}}{1-x^2}\,dx,\quad x\in\left[-\sqrt{1-t^2},\sqrt{1-t^2}\right], \tag{5-1}$$

where we have used the Stieltjes–Perron inversion formula.

Another notable case is the semicircle distribution on the interval $[-1,1]$,

$$d\mu_0(x)=\frac{2}{\pi}\sqrt{1-x^2}\,dx,\quad x\in(-1,1),$$

corresponding, for instance, to the zero limit distribution of the appropriately rescaled Hermite polynomials. Since

$$u(z,0)=2(z-\sqrt{z^2-1})=\frac{1}{z}+\mathcal{O}\left(\frac{1}{z^2}\right),\quad z\in\mathbb{C}\setminus[-1,1],$$

and

$$u^{-1}(w,0)=\frac{w^2+4}{4w},$$

solving the equation

$$z=\frac{w^2+4-4t}{4w},$$

we get that

$$w=u(z,t)=2(z-\sqrt{z^2-(1-t)}).$$

Again, μ_t is the absolutely continuous probability measure given by

$$d\mu_t(x)=\frac{2}{\pi}\sqrt{1-t-x^2}\,dx,\quad x\in\left[-\sqrt{1-t},\sqrt{1-t}\right], \tag{5-2}$$

the rescaled semicircle distribution for all values of $t\in[0,1)$.

The case of the Marchenko–Pastur distribution on $[0,4]$, given by

$$d\mu_0(x)=\frac{1}{2\pi}\frac{\sqrt{x(4-x)}}{x}\,dx,$$

that describes the asymptotic behavior of singular values of large random matrices (or equivalently, the asymptotic zero distribution of rescaled Laguerre polynomials) is analyzed in a similar way.

On the other hand, a natural case to consider is of the normalized uniform distribution on $[-1, 1]$. It turns out it is remarkably challenging; indeed, in this case

$$u(z, 0) = \frac{1}{2}\log\Big(\frac{z+1}{z-1}\Big), \quad z \in \mathbb{C} \setminus [-1, 1],$$

and

$$u^{-1}(w, 0) = \frac{v+1}{v-1}, \quad v = e^{2w}.$$

However, the equation

$$z = \frac{v+1}{v-1} - \frac{t}{\log v}, \quad v = e^{2w}, \tag{5-3}$$

is highly transcendental and does not have an explicit solution. We can use this representation to find the support of the measure μ_t, which again appears to be a formidable task.

Let us end with the case where μ_0 is discrete and is the normalized zero-counting measure of a given polynomial P of degree m, with real or complex zeros. This is the case, for instance, addressed at the beginning, when the Q_n are generated by a Rodrigues-type formula. As mentioned, this situation has been studied before in [Bøgvad et al. 2024] using different techniques.

Now

$$u(z, 0) = \frac{1}{m}\frac{d}{dz}\log P(z),$$

and (2-12) reduces to[4]

$$mP(z + tv) - vP'(z + tv) = 0, \quad v = \frac{1}{w}.$$

Using that

$$P(z + s) = \sum_{j=0}^{m} \frac{P^{(j)}(z)}{j!} s^j,$$

this yields an algebraic equation for $w = u(z, t)$,

$$\sum_{j=0}^{m} \left(m - \frac{j}{t}\right) \frac{P^{(j)}(z)}{j!} t^j w^{m-j} = 0, \tag{5-4}$$

also derived in [Bøgvad et al. 2024].

For instance, in our first example (1-7), when $P(z) = z^2 - 1$ and $m = 2$,

$$u(z, 0) = \frac{z}{z^2 - 1},$$

and (5-4) yields the quadratic equation for $w = u(z, t)$:

$$(z^2 - 1)w^2 - (1 - 2t)zw - (1 - t)t = 0.$$

Thus,

$$u(z, t) = \frac{(1 - 2t)z - \sqrt{4(t - 1)t + z^2}}{2(z^2 - 1)},$$

and we get the expressions (1-8)–(1-9) for the measure μ_t.

[4]This equation appears in [Bøgvad et al. 2024, Formula (1.6)] in the context that is a particular case of a discrete measure μ_0.

Remark 5.1. Ravichandran [2020] (see also [Totik 2021]) proved that, for $t \geq \frac{1}{2}$, the ratio of the diameters of the zero sets of $Q_{n,k}(x)$, defined in (1-2), and Q_n, with $k = [nt]$, is bounded above by $2\sqrt{t(1-t)}$. Obviously, both (5-1) and (5-2) are in agreement with this bound, and formulas (1-8)–(1-9) show that the assumption that $t \geq \frac{1}{2}$ cannot be omitted.

On the other hand, in Kalyagin's example (1-11), when $P(z) = z(z^2-1)$ and $m = 3$, (5-4) yields the following cubic equation for $w = u(z,t)$:

$$3z(z^2-1)w^3 + (3t-1)(3z^2-1)w^2 + 3t(3t-2)zw - 3t^2(1-t) = 0.$$

For $t \in (0,1)$, the end-points of the support of μ_t can be found from the branch points of w, which are solutions of the biquadratic equation

$$9z^4 + 3(9t^2-6t-2)z^2 - (1-t)(3t-1)^3 = 0;$$

see [Kalyagin 1979].

Finally, if P has complex roots, such as $P(z) = z^3 - 1$, (5-4) yields again a cubic equation for $w = u(z,t)$,

$$(z^3-1)w^3 - (1-3t)z^2w^2 - (2-3t)tzw - (1-t)t^2 = 0,$$

whose branch points are solutions of the equation

$$27t^3 - 54t^2 + 27t - 4z^3 = 0,$$

that is, the three cube roots

$$z = 3\sqrt[3]{\frac{t(1-t)^2}{4}}.$$

Appendix: Auxiliary results

Denote by $\mathcal{H}(R, M, c)$ the class of functions f satisfying the following conditions:

(i) f is analytic in the closed disc $\overline{\mathbb{D}}_R$;

(ii) $|f(z)| \leq M$ for $z \in \overline{\mathbb{D}}_R$;

(iii) $f'(0) = c > 0$.

Proposition A.1. *For $R, M, c > 0$, let*

$$r_1 := \min\left\{\frac{R}{2}, \frac{cR^2}{4M}\right\}. \tag{A-1}$$

Then, for $r \leq r_1$, any function $f \in \mathcal{H}(R, M, c)$ is univalent in $\overline{\mathbb{D}}_r$ and $f(\overline{\mathbb{D}}_r)$ covers the disc $\overline{\mathbb{D}}_\rho$ with $\rho = \frac{1}{2}cr$.

Consequently, f^{-1} is univalent in $\overline{\mathbb{D}}_\rho$ and $f^{-1}(\overline{\mathbb{D}}_\rho) \subset \mathbb{D}_r$.

Proof. Let

$$f(z) = cz + h(z) \in \mathcal{H}(R, M, c) \quad \text{and} \quad h(z) = \sum_{n=2}^{\infty} c_n z^n.$$

Since h is holomorphic in the closed disc $\overline{\mathbb{D}}_R$, the Cauchy inequalities imply $|c_n| \le M/R^n,\ n = 2, 3, \ldots,$ so that for any $r < R$

$$|h(z)| \le M \sum_{n=2}^{\infty} \left(\frac{r}{R}\right)^n = M \frac{(r/R)^2}{1-(r/R)}, \quad z \in \mathbb{D}_r.$$

If $r \le \frac{1}{2}R$ then $1 - r/R \ge \frac{1}{2}$, and

$$|h(z)| \le 2M \left(\frac{r}{R}\right)^2, \quad z \in \mathbb{D}_r.$$

Also,

$$r \le \frac{cR^2}{4M} \quad \Longleftrightarrow \quad cr \ge 4M \left(\frac{r}{R}\right)^2.$$

This means that if we assume $r \le r_1$, it implies, for $z \in C_r := \partial\mathbb{D}_r = \{z \in \mathbb{C} : |z| = r\}$, that $|cz| \ge 2|h(z)| > 0$. This means that the image of the circle C_r by f lies in the annulus $\frac{1}{2}cr = \rho \le |z| \le 3\rho$ and, by Rouché's theorem, its index with respect to the origin is one. The index of $f(C_r)$ with respect to any point $f(a) \notin f(C_r)$ is constant in any connected component of the complement to $f(C_r)$ and is equal to the multiplicity of $f(a)$ in $\mathbb{D}_r$. We conclude that this index is 1 for any $a \in \mathbb{D}_r$, concluding that f is univalent in $\mathbb{D}_r$. □

Proposition A.2. *Let $f \in \mathcal{H}(R, M, c)$ and*

$$r_0 := \min\left\{\frac{R}{2}, \frac{cR^2}{16M}\right\} \le r_1, \tag{A-2}$$

where r_1 was defined in (A-1).

Then for any $r \le r_0$ we have $|f'(z)| \ge \frac{1}{2}c$ for $z \in \mathbb{D}_r$.

Proof. Since $f'(z) = c + h'(z),\ h(z) = \sum_{n=2}^{\infty} c_n z^n$, the Cauchy inequalities imply again that, for $z \in \mathbb{D}_r$,

$$|h'(z)| \le \left|\sum_{n=2}^{\infty} n c_n z^{n-1}\right| \le M \sum_{n=2}^{\infty} n \frac{r^{n-1}}{R^n} = \frac{Mr}{R^2} \sum_{n=1}^{\infty} (n+1)\left(\frac{r}{R}\right)^{n-1}$$
$$\le \frac{2Mr}{R^2} \sum_{n=1}^{\infty} n\left(\frac{r}{R}\right)^{n-1} = \frac{2Mr}{R^2} \frac{1}{(1-r/R)^2} = \frac{2Mr}{(R-r)^2}.$$

If $r \le \frac{1}{2}R$, then $|h'(z)| \le 2Mr/r^2 = 2M/r$ for $z \in \mathbb{D}_r$. On the other hand, if additionally $r \le cR^2/(16M)$ then $c \ge 16Mr/R^2 \ge 4M/r$, which implies that, for $z \in \mathbb{D}_r$,

$$|f'(z)| \ge |c + h'(z)| \ge c - |h'(z)| \ge c - \frac{2M}{r} \ge c - \frac{8Mr}{R^2} \ge \frac{c}{2}.$$ □

Proposition A.3. *Let $f, g \in \mathcal{H}(R, M, c)$ and, for some $\epsilon > 0$, $|f(z) - g(z)| < \epsilon$ for $z \in \mathbb{D}_R$. Then there exist positive ρ and C, depending only on R, M, and c, such that the inverse functions f^{-1} and g^{-1} exist in $\mathbb{D}_\rho$ and*

$$m(\rho) := \max_{\zeta \in \overline{\mathbb{D}}_\rho} |f^{-1}(\zeta) - g^{-1}(\zeta)| \le C\epsilon. \tag{A-3}$$

Proof. Fix r_0 defined in (A-2) and consider $r \in (0, r_0)$, as long as the corresponding ρ equals $\frac{1}{2}cr$. By Proposition A.1, for any such r and ρ, the functions $f^{-1}(\zeta)$, $g^{-1}(\zeta)$ exist, are univalent in $\overline{\mathbb{D}}_\rho$, and map $\overline{\mathbb{D}}_\rho$ into $\mathbb{D}_r$. Thus, for any $\rho \le \frac{1}{2}c$ and $\zeta \in \mathbb{D}_\rho$,

$$\begin{aligned} |f^{-1} - g^{-1}|(\zeta) = \left| \int_0^\zeta \left[(f^{-1})'(w) - (g^{-1})'(w) \right] dw \right| &= \int_0^\zeta \left| \frac{1}{f'(f^{-1}(w))} - \frac{1}{g'(g^{-1}(w))} \right| |dw| \\ &\le \rho \max_{w \in \mathbb{D}_\rho} \left| \frac{g'(g^{-1}(w)) - f'(f^{-1}(w))}{g'(g^{-1}(w)) f'(f^{-1}(w))} \right| \\ &\le \frac{4\rho}{c^2} \max_{w \in \mathbb{D}_\rho} \left| g'(g^{-1}(w)) - f'(f^{-1}(w)) \right|, \quad \text{(A-4)} \end{aligned}$$

where we have used Proposition A.2. For $w \in \mathbb{D}_\rho$,

$$g'(g^{-1}(w)) - f'(f^{-1}(w)) = \big(g'(g^{-1}(w)) - f'(g^{-1}(w)) \big) + \big(f'(g^{-1}(w)) - f'(f^{-1}(w)) \big),$$

and thus

$$\left| g'(g^{-1}(w)) - f'(f^{-1}(w)) \right| \le \left| (g' - f')(g^{-1}(w)) \right| + \left| f'(g^{-1}(w)) - f'(f^{-1}(w)) \right|. \quad \text{(A-5)}$$

Observe that, by the Cauchy integral formula, for $f \in \mathcal{H}(R, M, c)$ and $z \in \mathbb{D}_r$, with any $r < R$, we have $|f'(z)| \le M/(R - r)$ and $|2f''(z)| \le M/(R - r)^2$. Thus, since $z = g^{-1}(w) \in \mathbb{D}_r$ for $w \in \mathbb{D}_\rho$, we have

$$\left| g'(g^{-1}(w)) - f'(g^{-1}(w)) \right| = |(g' - f')(z)| \le \frac{\epsilon}{R - r} \le 2\epsilon$$

and

$$\left| f'(g^{-1}(w)) - f'(f^{-1}(w)) \right| \le \max_{z \in \mathbb{D}_r} |f''(z)| \max_{w \in \mathbb{D}_\rho} |g^{-1}(w) - f^{-1}(w)| \le \frac{M}{2(R - r)^2} m(\rho) \le \frac{2M}{R^2} m(\rho).$$

Combining these bounds in (A-5) we obtain from (A-4)

$$m(\rho) \le \frac{4\rho}{c^2} \left(2\epsilon + \frac{2M}{R^2} m(\rho) \right).$$

Thus, selecting for instance $\rho = c^2 R^2/(16M)$ we obtain (A-3) with $C = M/R^2$. □

Acknowledgements

Martínez-Finkelshtein was partially supported by Simons Foundation Collaboration Grants for Mathematicians (grant MPS-TSM-00710499). He also acknowledges the support of the project PID2021-124472NB-I00, funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe", as well as the support of Junta de Andalucía (research group FQM-229 and Instituto Interuniversitario Carlos I de Física Teórica y Computacional).

The authors also thank Brian Hall and Edward Saff for interesting discussions, as well as the anonymous referee for very useful observations.

References

[Angst et al. 2024] J. Angst, D. Malicet, and G. Poly, "Almost sure behavior of the critical points of random polynomials", *Bull. Lond. Math. Soc.* **56**:2 (2024), 767–782. MR

[Aptekarev and Rykov 2006] A. I. Aptekarev and Y. G. Rykov, "On the variational representation of solutions to some quasilinear equations and systems of hyperbolic type on the basis of potential theory", *Russ. J. Math. Phys.* **13**:1 (2006), 4–12. MR

[Aptekarev et al. 1997] A. I. Aptekarev, F. Marcellán, and I. A. Rocha, "Semiclassical multiple orthogonal polynomials and the properties of Jacobi–Bessel polynomials", *J. Approx. Theory* **90**:1 (1997), 117–146. MR

[Arendt and Urban 2023] W. Arendt and K. Urban, *Partial differential equations—an introduction to analytical and numerical methods*, Graduate Texts in Mathematics **294**, Springer, 2023. MR

[Arizmendi et al. 2023] O. Arizmendi, J. Garza-Vargas, and D. Perales, "Finite free cumulants: multiplicative convolutions, genus expansion and infinitesimal distributions", *Trans. Amer. Math. Soc.* **376**:6 (2023), 4383–4420. MR

[Bøgvad et al. 2024] R. Bøgvad, C. Hägg, and B. Shapiro, "Rodrigues' descendants of a polynomial and Boutroux curves", *Constr. Approx.* **59**:3 (2024), 737–798. MR

[Byun et al. 2022] S.-S. Byun, J. Lee, and T. R. Reddy, "Zeros of random polynomials and their higher derivatives", *Trans. Amer. Math. Soc.* **375**:9 (2022), 6311–6335. MR

[Campbell et al. 2024] A. Campbell, S. O'Rourke, and D. Renfrew, "The fractional free convolution of R-diagonal elements and random polynomials under repeated differentiation", *Int. Math. Res. Not.* **2024**:13 (2024), 10189–10218. MR

[Duren 1983] P. L. Duren, *Univalent functions*, Grundl. Math. Wissen. **259**, Springer, 1983. MR

[Evans 1998] L. C. Evans, *Partial differential equations*, Graduate Studies in Mathematics **19**, Amer. Math. Soc., Providence, RI, 1998. MR

[Galligo 2022] A. Galligo, "Modeling complex root motion of real random polynomials under differentiation", pp. 245–253 in *ISSAC '22—Proceedings of the* 2022 *International Symposium on Symbolic and Algebraic Computation*, edited by A. Hashemi, ACM, New York, 2022. MR

[Hall et al. 2025] B. C. Hall, C.-W. Ho, J. Jalowy, and Z. Kabluchko, "Zeros of random polynomials undergoing the heat flow", *Electron. J. Probab.* **30** (2025), art. id. 159. MR

[Hall et al. 2026] B. C. Hall, C.-W. Ho, J. Jalowy, and Z. Kabluchko, "Roots of polynomials under repeated differentiation and repeated applications of fractional differential operators", *Trans. Amer. Math. Soc. Ser. B* **13** (2026), 190–239. MR

[Hoskins and Kabluchko 2023] J. Hoskins and Z. Kabluchko, "Dynamics of zeroes under repeated differentiation", *Exp. Math.* **32**:4 (2023), 573–599. MR

[Hoskins and Steinerberger 2022] J. G. Hoskins and S. Steinerberger, "A semicircle law for derivatives of random polynomials", *Int. Math. Res. Not.* **2022**:13 (2022), 9784–9809. MR

[Hotta and Katori 2018] I. Hotta and M. Katori, "Hydrodynamic limit of multiple SLE", *J. Stat. Phys.* **171**:1 (2018), 166–188. MR

[Hotta and Schleißinger 2021] I. Hotta and S. Schleißinger, "Limits of radial multiple SLE and a Burgers–Loewner differential equation", *J. Theoret. Probab.* **34**:2 (2021), 755–783. MR

[Kabluchko 2026] Z. Kabluchko, "Repeated differentiation and free unitary Poisson process", *Trans. Amer. Math. Soc.* **379**:3 (2026), 2217–2253. MR

[Kalyagin 1979] V. A. Kalyagin, "On a class of polynomials defined by two orthogonality relations", *Mat. Sb.* (*N.S.*) **110(152)**:4(12) (1979), 609–627. In Russian; translated in *Math. USSR Sb.* **38**:4 (1981), 563–580.

[Kerov 1998] S. Kerov, "Interlacing measures", pp. 35–83 in *Kirillov's seminar on representation theory*, edited by G. I. Olshanski, Amer. Math. Soc. Transl. Ser. 2 **181**, Amer. Math. Soc., Providence, RI, 1998. MR

[Kiselev and Tan 2022] A. Kiselev and C. Tan, "The flow of polynomial roots under differentiation", *Ann. PDE* **8**:2 (2022), art. id. 16. MR

[MacGregor 1962] T. H. MacGregor, "Functions whose derivative has a positive real part", *Trans. Amer. Math. Soc.* **104** (1962), 532–537. MR

[Marcus et al. 2022] A. W. Marcus, D. A. Spielman, and N. Srivastava, "Finite free convolutions of polynomials", *Probab. Theory Related Fields* **182**:3-4 (2022), 807–848. MR

[Markowich 2007] P. A. Markowich, *Applied partial differential equations*: *a visual approach*, Springer, 2007. MR

[Martínez-Finkelshtein et al. 2015] A. Martínez-Finkelshtein, R. Orive, and E. A. Rakhmanov, "Phase transitions and equilibrium measures in random matrix models", *Comm. Math. Phys.* **333**:3 (2015), 1109–1173. MR

[Mingo and Speicher 2017] J. A. Mingo and R. Speicher, *Free probability and random matrices*, Fields Institute Monographs **35**, Springer, 2017. MR

[del Monaco and Schleißinger 2016] A. del Monaco and S. Schleißinger, "Multiple SLE and the complex Burgers equation", *Math. Nachr.* **289**:16 (2016), 2007–2018. MR

[Nica and Speicher 2006] A. Nica and R. Speicher, *Lectures on the combinatorics of free probability*, London Mathematical Society Lecture Note Series **335**, Cambridge University Press, Cambridge, 2006. MR

[O'Rourke and Steinerberger 2021] S. O'Rourke and S. Steinerberger, "A nonlocal transport equation modeling complex roots of polynomials under differentiation", *Proc. Amer. Math. Soc.* **149**:4 (2021), 1581–1592. MR

[Ravichandran 2020] M. Ravichandran, "Principal submatrices, restricted invertibility, and a quantitative Gauss–Lucas theorem", *Int. Math. Res. Not.* **2020**:15 (2020), 4809–4832. MR

[Saff and Totik 2024] E. B. Saff and V. Totik, *Logarithmic potentials with external fields*, 2nd ed., Grundl. Math. Wissen. **316**, Springer, 2024. MR

[Salih 2018] A. Salih, "Inviscid Burgers' equation", Lecture Notes, Indian Institute of Space Science and Technology, Thiruvananthapuram, 2018, available at https://stageweb.iist.ac.in/sites/default/files/2025-06/Burgers_equation_inviscid.pdf.

[Shlyakhtenko and Tao 2022] D. Shlyakhtenko and T. Tao, "Fractional free convolution powers", *Indiana Univ. Math. J.* **71**:6 (2022), 2551–2594. MR

[Sorokin 2022] V. N. Sorokin, "A generalization of the discrete Rodrigues formula for Meixner polynomials", *Mat. Sb.* **213**:11 (2022), 79–101. In Russian; translated in *Sb. Math.* **213**:11 (2022), 1559–1581. MR

[Sorokin 2023] V. N. Sorokin, "On polynomials defined by the discrete Rodrigues formula", *Mat. Zametki* **113**:3 (2023), 423–439. MR

[Stahl and Totik 1992] H. Stahl and V. Totik, *General orthogonal polynomials*, Encyclopedia of Mathematics and its Applications **43**, Cambridge University Press, 1992. MR

[Steinerberger 2019] S. Steinerberger, "A nonlocal transport equation describing roots of polynomials under differentiation", *Proc. Amer. Math. Soc.* **147**:11 (2019), 4733–4744. MR

[Steinerberger 2023] S. Steinerberger, "Free convolution powers via roots of polynomials", *Exp. Math.* **32**:4 (2023), 567–572. MR

[Totik 2021] V. Totik, "Critical points of polynomials", *Acta Math. Hungar.* **164**:2 (2021), 499–517. MR

[VandenHeuvel et al. 2023] D. J. VandenHeuvel, C. J. Lustri, J. R. King, I. W. Turner, and S. W. McCue, "Burgers' equation in the complex plane", *Phys. D* **448** (2023), art. id. 133686. MR

[Whitham 1974] G. B. Whitham, *Linear and nonlinear waves*, Wiley, New York, 1974. MR

Received 1 Dec 2024. Revised 16 Jul 2025. Accepted 19 Sep 2025.

ANDREI MARTÍNEZ-FINKELSHTEIN: a_martinez-finkelshtein@baylor.edu
Department of Mathematics, Baylor University, Waco, TX, United States

and

Department of Mathematics, University of Almería, Almería, Spain

EVGUENII A. RAKHMANOV: rakhmano@usf.edu
Department of Mathematics, University of South Florida, Tampa, FL, United States

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 19 No. 7 2026

Existence, uniqueness and characterisation of local minimisers in second-order calculus of variations in L^∞ 1251
NIKOS KATZOURAKIS and ROGER MOSER

An $L^{3/2}$ SL_2 Kakeya maximal inequality 1295
JOHN GREEN, TERENCE L. J. HARRIS and YUMENG OU

The Brunn–Minkowski inequality for the first eigenvalue of the Ornstein–Uhlenbeck operator and log-concavity of the relevant eigenfunction 1339
ANDREA COLESANTI, ELISA FRANCINI, GALYNA LIVSHYTS and PAOLO SALANI

Almost-fuchsian structures on disc bundles over a surface 1361
SAMUEL BRONSTEIN

On the Ohsawa–Takegoshi L^2 extension theorem and removable singularities of plurisubharmonic functions 1395
XIEPING WANG

Flow of the zeros of polynomials under iterated differentiation 1417
ANDREI MARTÍNEZ-FINKELSHTEIN and EVGUENII A. RAKHMANOV