ANALYSIS & PDE

Volume 12 No. 7 2019

msp

Analysis & PDE
msp.org/apde

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

The subscription price for 2019 is US $310/year for the electronic version, and $520/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 12, No. 7, 2019
dx.doi.org/10.2140/apde.2019.12.1643

GENERALIZED q-GAUSSIAN VON NEUMANN ALGEBRAS WITH COEFFICIENTS I: RELATIVE STRONG SOLIDITY

MARIUS JUNGE AND BOGDAN UDREA

We define $\Gamma_q(B, S \otimes H)$, the generalized q-gaussian von Neumann algebras associated to a sequence of symmetric independent copies (π_j, B, A, D) and to a subset $1 \in S = S^* \subset A$ and, under certain assumptions, prove their strong solidity relative to B. We provide many examples of strongly solid generalized q-gaussian von Neumann algebras. We also obtain nonisomorphism and nonembedability results about some of these von Neumann algebras.

1. Background and statement of results

1A. *Normalizers in von Neumann algebras.* The study of normalizers in von Neumann algebras is nowadays an intensely active area of research within the field of von Neumann algebras. For a von Neumann algebra M, we denote by $\mathcal{U}(M)$ the group of unitaries in M. Recall that for an inclusion $A \subset M$ of von Neumann algebras, *the normalizing group of* A in M is defined as $\mathcal{N}_M(A) = \{u \in \mathcal{U}(M) : uAu^* = A\}$ and *the normalizer of* A is the von Neumann algebra generated by the normalizing group, i.e., $\mathcal{N}_M(A)'' \subset M$. When A is a maximal abelian von Neumann subalgebra of a type-II_1 factor M, A is called a *Cartan subalgebra* if its normalizer is the whole of M. While some results were obtained in the early 80's, see, e.g., [Connes and Jones 1982; Jones and Popa 1982] and most notably [Connes, Feldman and Weiss 1981], the first truly significant achievement in this area is Voiculescu's ground-breaking result [1996] about the absence of Cartan subalgebras in the free-group factors $L(\mathbb{F}_n)$. After this, in their seminal work, Ozawa and Popa [2010a] established the following result:

MSC2010: 46L10.
Keywords: von Neumann algebras, q-gaussian von Neumann algebras with coefficients, relative strong solidity.

Theorem 1.1. *Let $\mathbb{F}_n \curvearrowright B$ be a* profinite *trace-preserving action of a free group on an amenable von Neumann algebra B. Then for every amenable von Neumann subalgebra $A \subset M = B \rtimes \mathbb{F}_n$, either $A \prec_M B$, or the normalizer of A is amenable.*

The notation $A \prec_M B$ reads "a corner of A embeds into B inside M", in the sense of [Popa 2006b, Theorem 2.1], and it roughly means that A can be conjugated into B by a partial isometry in M. When B is the scalars, this shows that the normalizer of any amenable diffuse von Neumann subalgebra of $L(\mathbb{F}_n)$ is itself amenable, not only reproving and strengthening Voiculescu's result, but also giving a surprisingly far-reaching classification of normalizer algebras in the free group factors. More than merely proving the above theorem, [Ozawa and Popa 2010a] introduced an array of innovative ideas and techniques which remain all-pervasive and highly influential in the field to this day. The results in that paper were then extended in [Ozawa and Popa 2010b] to profinite actions of weakly amenable groups having proper 1-cocycles into (a multiple of) their left regular representations. Subsequent generalizations to the case of profinite actions of groups having quasicocycles or direct products of such have been obtained in [Chifan and Sinclair 2013; Chifan, Sinclair and Udrea 2013]. Recently, Popa and Vaes obtained the definitive form of these results, by completely removing any assumption on the action of the group. Specifically, they proved the following results (see [Popa and Vaes 2014a, Theorem 1.6; 2014b, Theorem 1.4]):

Theorem 1.2. *Let Γ be a weakly amenable group having either a proper 1-cocycle or a proper 1-quasicocycle into a (representation which is weakly contained into) a multiple of its left regular representation. Let $\Gamma \curvearrowright B$ be any trace-preserving action of Γ on the finite von Neumann algebra B, and let $A \subset M = B \rtimes \Gamma$ be a von Neumann subalgebra which is amenable relative to B inside M. Then either $A \prec_M B$, or the normalizer of A is amenable relative to B inside M.*

Theorem 1.3. *Let $\Gamma \curvearrowright B$ be a p.m.p. free ergodic action, where B is abelian diffuse and Γ is weakly amenable and admits an unbounded (rather than proper) 1-cocycle into a mixing representation which is weakly contained into a multiple of the left regular representation of Γ. Then $M = B \rtimes \Gamma$ has a unique Cartan subalgebra, up to unitary conjugacy.*

Popa and Vaes coined the phrase "relative strong solidity" to describe the situation in which the dichotomy in Theorem 1.2 holds. Namely, a von Neumann algebra M is *strongly solid relative to B*, for $B \subset M$ a subalgebra, if for every von Neumann subalgebra $A \subset M$ which is amenable relative to B inside M, see [Ozawa and Popa 2010a], it is either the case that $A \prec_M B$ or that the normalizer of A is amenable relative to B inside M. In the case of B abelian diffuse and of p.m.p. free ergodic actions $\Gamma \curvearrowright B$, the strong solidity of the von Neumann algebra $M = B \rtimes \Gamma$ relative to B implies its uniqueness of Cartan subalgebra, up to unitary conjugacy. Strong solidity relative to the scalars is simply termed strong solidity. Strong solidity is in turn an enhancement of Ozawa's concept of *solidity* [2004]. Ozawa called a von Neumann algebra M solid if for every diffuse von Neumann subalgebra $\mathcal{A} \subset M$ one has that $\mathcal{A}' \cap M$ is amenable. It's easy to see that a nonamenable solid factor M is automatically *prime*, i.e., cannot be written as $M = M_1 \mathbin{\bar{\otimes}} M_2$, with M_i an infinite-dimensional factor for $i = 1, 2$.

Further results pertaining to the classification of normalizers and relative strong solidity were obtained in [Sinclair 2011; Ioana 2015; Isono 2015a; 2015b; Avsec 2012; Boutonnet, Houdayer and Vaes 2018; Houdayer and Vaes 2013; Caspers 2018].

1B. ***Noncommutative probability.*** Voiculescu [1985] introduced his highly influential free probability theory in the early 80's, in order to tackle some problems related to the free group factors. Since then, free probability theory has grown into an immense industry with far-reaching ramifications. Very roughly speaking, in the realm of free probability, classical probability spaces are replaced by C^*- or W^*-algebras endowed with distinguished states (normal in the W^* case), classical random variables by operators in those algebras, classical independence by Voiculescu's free independence, and the classical distribution function by Voiculescu noncommutative distribution of a noncommutative random variable, or joint distribution in the case of a system of random variables. In particular, the normal (gaussian) distribution is replaced by Wigner's semicircular law.

1B1. *Classical gaussian random variables.* We briefly recall the construction in Section 1.1 of [Peterson and Sinclair 2012]. Let H a real Hilbert space and, for $\xi \in H$, let l_ξ be the creation operator on the symmetric Fock space of $H_{\mathbb{C}} = H \oplus iH$. Then $s_1(\xi) = \frac{1}{2}(l_\xi + l_\xi^*)$ is an unbounded self-adjoint operator in the symmetric Fock space. The operators $s_1(\xi)$ and $s_1(\eta)$ commute for all ξ and η and are independent with respect to the vacuum state whenever $\langle \xi, \eta\rangle = 0$. Define $\Gamma_1(H)$ to be the abelian von Neumann algebra generated by the spectral projections of all the $s_1(\xi)$, $\xi \in H$ (or equivalently by all the unitaries $\omega(\xi_1, \dots, \xi_k) = \exp(i\pi s(\xi_1)\cdots s(\xi_k))$), equipped with the trace given by the restriction of the vacuum state. For $\|\xi\| = 1$, we have the moment formula

$$\tau(s(\xi)^m) = \delta_{m\in 2\mathbb{N}} \frac{m!}{2^{\frac{m}{2}}\left(\frac{m}{2}\right)!} = |P_2(m)| = \sum_{\sigma\in P_2(m)} 1^{\mathrm{cr}(\sigma)},$$

where $P_2(m)$ is the collection of pair partitions on the set $\{1, \dots, m\}$, and for $\sigma \in P_2(m)$, $\mathrm{cr}(\sigma)$ denotes the number of crossings of σ. These are exactly the moments of a classical gaussian random variable. By commutativity, independence and multilinearity, the moment formula can be extended to

$$\tau(s_1(\xi_1)\cdots s_1(\xi_m)) = \sum_{\sigma\in P_2(m)} 1^{\mathrm{cr}(\sigma)} \prod_{\{l,r\}\in\sigma} \langle \xi_l, \xi_r\rangle.$$

One also recalls:

Theorem 1.4 (classical central limit theorem). *Let $\{X_n\}_{n\ge 1}$ be a sequence of independent, identically distributed random variables on a probability space (Ω, Σ, P), all having mean equal to zero and variance equal to* 1. *Then the averages $S_n = n^{-\frac{1}{2}} \sum_{j=1}^n X_j$ converge in distribution to a normal (gaussian) random variable with mean zero and variance* 1.

If X_n are chosen such that $\sup_{n\ge 1} \|S_n\|_\infty < \infty$, then one can restate the central limit theorem by saying that the element $S = (S_n)_n \in (L^\infty(\Omega, P)^\omega, \tau_\omega)$ has a gaussian (normal) distribution, where ω is a free ultrafilter on $\mathbb{N}$ and $(L^\infty(\Omega)^\omega, \tau_\omega)$ is the ultraproduct von Neumann algebra. In other words, one could "simulate" gaussian elements using an ultraproduct model.

1B2. *Voiculescu's free semicircular random variables.* Voiculescu [1985] constructed a functor Φ from the category of real Hilbert spaces with contractions to the category of finite von Neumann algebras with completely positive maps. For $h \in H$, the element $\Phi(h) = s_0(h)$ is concretely realized as the real part of the creation operator on the full Fock space of $H_{\mathbb{C}}$. Moreover, he proved that for an orthonormal set $\{h_1, \dots, h_m\} \subset H$, the elements $s_0(h_1), \dots, s_0(h_m)$ are freely independent, have semicircular distributions given by $d\mu(t) = \frac{1}{\pi}\chi_{(-1,1)}(t)\sqrt{1-t^2}\,dt$, and generate a copy of the free group factor $L(\mathbb{F}_m)$. In particular, for a finite-dimensional Hilbert space H, $\Phi(H)$ is $*$-isomorphic to the free group factor $L(\mathbb{F}_{\dim(H)})$. It is well known that the moments of a semicircular variable are given by

$$\tau(s_0(h)^m) = \delta_{m\in 2\mathbb{N}} \frac{m!}{\left(\frac{m}{2}+1\right)\left(\left(\frac{m}{2}\right)!\right)^2} = \sum_{\sigma \in P_2(m)} 0^{\mathrm{cr}(\sigma)} = \sum_{\sigma \in \mathrm{NCP}_2(m)} 0^{\mathrm{cr}(\sigma)},$$

where we denote by $\mathrm{NCP}_2(m)$ the collection of noncrossing pair partitions on the set $\{1, \dots, m\}$ and with the convention $0^0 = 1$. By direct computation, the above formula can be extended to

$$\tau(s_0(h_1)\cdots s_0(h_m)) = \sum_{\sigma \in P_2(m)} 0^{\mathrm{cr}(\sigma)} \prod_{\{l,r\}\in\sigma} \langle h_l, h_r\rangle.$$

Let's recall Voiculescu's central limit theorem [1985]:

Theorem 1.5 (Voiculescu's central limit theorem). *Let $\{a_n\}_{n\geq 1}$ be a sequence of freely independent self-adjoint random variables in a C^*-probability space (A, φ). Assume that $\varphi(a_n) = 0$ for all n, $\sup_{n\geq 1}\|a_n\| < \infty$ and $\lim_{n\to\infty} n^{-1}\sum_{j=1}^n \varphi(a_n^2) = \frac{1}{4}$. Then the elements $S_n = n^{-\frac{1}{2}}\sum_{j=1}^n a_j$ converge in distribution to a semicircular element*; *i.e., their limit distribution is the semicircular law.*

In particular, this says that, beside the Fock space construction, one could create semicircular random variables by taking elements of the form

$$S = (S_n)_n = \left(n^{-\frac{1}{2}} \sum_{j=1}^n a_j\right)_n \in (M^\omega, \tau_\omega),$$

with a_n as above in a finite W^*-probability space $(A, \varphi) = (M, \tau)$ and satisfying the additional condition

$$\sup_{n\geq 1}\left\| n^{-\frac{1}{2}} \sum_{j=1}^n a_j \right\|_\infty < \infty.$$

Let us also recall an informal statement of Voiculescu's matrix limit theorem [1991, Theorem 2.2]:

Theorem 1.6 (Voiculescu's matrix limit theorem). *Any family of random matrices with size going to infinity having independent normalized gaussian entries converges in distribution to a free semicircular family.*

Again, we could interpret this as saying that one can create free semicircular families using elements of some suitable ultraproduct of matrix algebras over abelian von Neumann algebras.

1B3. *q-gaussian von Neumann algebras.* The q-gaussian von Neumann algebras $\Gamma_q(H)$, for H a real Hilbert space, were introduced by Bożejko and Speicher [1991; 1992; 1994; 1996] and further studied in [Bożejko, Kümmerer and Speicher 1997; Ricard 2005; Śniady 2001; 2004; Królak 2000; 2006; Nou 2004; 2006; Shlyakhtenko 2004; 2009; Dykema and Nica 1993; Avsec 2012; Dabrowski 2014]. For $-1 < q < 1$, Bożejko and Speicher constructed a functor Γ_q from the category of real Hilbert spaces with contractions to the category of finite von Neumann algebras with completely positive maps. $\Gamma_q(H)$ is called the q-gaussian von Neumann algebra associated to H. The generators $\Gamma_q(h) = s_q(h)$, for $h \in H$, admit a concrete representation as the real part of the creation operator by h on the q-Fock space of H; for details, see, e.g., Section 2 of [Bożejko and Speicher 1991]. When $q = 0$, the functor Γ_q coincides with Voiculescu's functor Φ, so $\Gamma_0(H) = L(\mathbb{F}_{\dim(H)})$. A direct computation using the concrete realization of the $s_q(h)$'s gives the moment formula

$$\tau(s_q(h_1)\cdots s_q(h_m)) = \sum_{\sigma\in P_2(m)} q^{\mathrm{cr}(\sigma)} \prod_{\{l,r\}\in\sigma} \langle h_l, h_r\rangle,$$

which is why, in view of the above, the $s_q(h)$'s can be called q-semicircular elements. The central limit theorem holds in the q-gaussian context as well, see, e.g., [Speicher 1992, Theorem 1; 1993, Theorems 1 and 2; Bożejko 1991; Junge and Zeng 2015, Appendix A], but its statement is very technical and we omit it. Also, the q-gaussian von Neumann algebras admit random matrix models; see [Śniady 2001, Theorem 3]. We mention that, originally, the q-gaussian von Neumann algebras were studied as concrete implementations of the canonical q-commutation relations, or as examples of nonclassical Brownian motions, see, e.g., [Bożejko and Speicher 1991; 1992; 1994], but we choose to downplay these aspects in the present work. The central limit theorem suggests that the q-gaussians can be introduced via an ultraproduct model. In fact, a concrete ultraproduct embedding which holds a great heuristic value for us is given by

$$\Gamma_q(H) \to (\Gamma_q(\ell^2 \otimes H))^{\omega}, \quad s_q(h) \mapsto \left(n^{-\frac{1}{2}} \sum_{j=1}^{n} s_q(e_j \otimes h)\right)_n,$$

where $\{e_j\}_{j\in\mathbb{N}}$ is the standard orthonormal basis of $\ell^2 = \ell^2(\mathbb{N})$.

1C. ***Generalized q-gaussian von Neumann algebras with coefficients.*** In this article we introduce a new class of von Neumann algebras and prove some structural results about them. Specifically, we introduce the *generalized q-gaussian von Neumann algebras with coefficients* associated to a sequence of *symmetric copies* (π_j, B, A, D). A 4-tuple (π_j, B, A, D) is called a sequence of symmetric copies (of A) if B, A, D are finite tracial von Neumann algebras such that $B \subset A \cap D$ and $\pi_j : A \to D,\ j \in \mathbb{N}$, are unital trace-preserving normal $*$-homomorphisms satisfying

(1) $\pi_j|_B = \mathrm{id}_B$ for all j;

(2) $E_B(\pi_{j_1}(a_1)\cdots\pi_{j_m}(a_m)) = E_B(\pi_{\sigma(j_1)}(a_1)\cdots\pi_{\sigma(j_m)}(a_m))$ for all finite permutations σ on $\mathbb{N}$, all indices $j_1, \ldots, j_m$ in $\mathbb{N}$ and all $a_1, \ldots, a_m$ in A, where $E_B : D \to B$ is the canonical trace-preserving conditional expectation.

We mention that our copies satisfy some additional independence conditions (see Definition 3.2). Let $-1 < q < 1$ be fixed. For H an infinite-dimensional (real) Hilbert space and S a self-adjoint subset of A containing 1, the generalized q-gaussian von Neumann algebra

$$\Gamma_q(B, S \otimes H) \subset (\Gamma_q(\ell^2 \otimes H) \bar{\otimes} D)^\omega$$

with coefficients in B and associated to the symmetric copies (π_j, B, A, D) is defined as the von Neumann subalgebra generated by the elements

$$s_q(a, h) = \left(n^{-\frac{1}{2}} \sum_{j=1}^{n} s_q(e_j \otimes h) \otimes \pi_j(a)\right)_n, \quad a \in BSB = \{b_1 a b_2 : b_1, b_2 \in B, a \in S\},\ h \in H.$$

Here ω is a free ultrafilter on the natural numbers and $\Gamma_q(\ell^2 \otimes H)$ is the q-gaussian von Neumann algebra. When H is finite-dimensional, one needs to further apply a "closure operation" (see Definition 3.4 and Proposition 3.14 for more details). The crucial observation here is that, since a Fock space model is not available, we are forced to introduce our generalized gaussians via an ultraproduct model. The generators $s_q(a, h)$ satisfy the moment formula

$$\tau(s_q(a_1, h_1) \cdots s_q(a_m, h_m)) = \delta_{m \in 2\mathbb{N}} \sum_{\sigma \in P_2(m)} q^{\mathrm{cr}(\sigma)} \prod_{\{l,r\} \in \sigma} \langle h_l, h_r \rangle \tau_D(\pi_{\phi_\sigma(1)}(a_1) \cdots \pi_{\phi_\sigma(m)}(a_m)),$$

as well as the B-valued moment formula

$$E_B(s_q(a_1, h_1) \cdots s_q(a_m, h_m)) = \delta_{m \in 2\mathbb{N}} \sum_{\sigma \in P_2(m)} q^{\mathrm{cr}(\sigma)} \prod_{\{l,r\} \in \sigma} \langle h_l, h_r \rangle E_B(\pi_{\phi_\sigma(1)}(a_1) \cdots \pi_{\phi_\sigma(m)}(a_m)),$$

where for every pair partition $\sigma = \{\{k_1', k_1''\}, \dots, \{k_p', k_p''\}\} \in P_2(m)$, the function $\phi_\sigma : \{1, \dots, m\} \to \{1, \dots, p = \frac{m}{2}\}$ is chosen so that $\phi_\sigma(k_1') = \phi_\sigma(k_1'') = 1, \dots, \phi_\sigma(k_p') = \phi_\sigma(k_p'') = p$. In view of all of the above, the elements $s_q(a, h)$ could thus judiciously be called "B-valued q-semicircular random variables having symmetric B-moments".

When compared to pure q-gaussians, the generalized q-gaussian von Neumann algebras with coefficients can be viewed as an analogue of the cross-product von Neumann algebras $B \rtimes \Gamma$ as opposed to pure group von Neumann algebras $L(\Gamma)$. This analogy can be given some substance along the lines of [Shlyakhtenko 1999]. However, in the present work we do not pursue this insight and use this analogy merely as a guideline for the implementation of Popa's deformation-rigidity strategy. The main result we prove about our generalized q-gaussian algebras is:

Theorem A. *Let (π_j, B, A, D) be a sequence of symmetric independent copies, H be a finite-dimensional Hilbert space and $\mathcal{A} \subset M = \Gamma_q(B, S \otimes H)$ be a diffuse von Neumann subalgebra which is amenable relative to B inside M. For every $s \geq 0$, define $D_s(S)$ to be the following right B-submodule of $L^2(D)$:*

$$\overline{\mathrm{span}}^{\|\cdot\|_2}\{E_{A_{\{1,\dots,s\}}}(\pi_{\phi_\sigma(1)}(x_1) \cdots \pi_{\phi_\sigma(m)}(x_m)) : m \geq 1,\ \sigma \in P_{1,2}(m),\ x_i \in BSB\},$$

where $m = s + 2p$, $P_{1,2}(m)$ is the set of pair-singleton partitions of $\{1, \dots, m\}$, σ runs over all pair-singleton partitions in $P_{1,2}(m)$ having s singletons and p pairs and for such a partition $\sigma =$

$\{\{k_1\},\ldots,\{k_s\},\{k_1',k_1''\},\ldots,\{k_p',k_p''\}\}$, *the function* $\phi_\sigma:\{1,\ldots,m\}\to\{1,\ldots,s+p\}$ *satisfies* $\phi_\sigma(k_1)=1,\ldots,\phi_\sigma(k_s)=s$ *and* $\phi_\sigma(k_1')=\phi_\sigma(k_1'')=s+1,\ldots,\phi_\sigma(k_p')=\phi_\sigma(k_p'')=s+p$. *Assume that there exist constants* $d,C>0$ *such that* $\dim_B(D_s(S))\le Cd^s$ *for all* $s\ge 1$. *Then at least one of the following statements is true*:

(1) $\mathcal{A}\prec_M B$.

(2) *The von Neumann algebra* $P=\mathcal{N}_M(\mathcal{A})''$ *generated by the normalizer of* $\mathcal{A}$ *in* M *is amenable relative to* B *inside* M.

The technical condition on the dimension of the B-modules $D_k(S)$ implies in particular that the subspace of Wick words of length k is finitely generated over B for all $k\ge 1$ (see Theorem 3.16 and Proposition 3.20). This last condition in turn is the exact analogue of the group cocycle being proper in the case of cross-product von Neumann algebras.

As a consequence of our Theorem A, we find a number of examples of generalized q-gaussians which are strongly solid (when $B=\mathbb{C}$ or B is finite-dimensional) or strongly solid relative to B for diffuse B. While the class of generalized q-gaussian von Neumann algebras with coefficients is huge (roughly speaking such a von Neumann algebra can be constructed starting from any action of the infinite symmetric group on another finite von Neumann algebra), the range of examples to which our Theorem A applies is greatly restricted by the technical assumptions we make. The examples in the corollary below are introduced in more detail in Section 4.

Corollary B. *The following von Neumann algebras are strongly solid relative to* B:

(1) (*see Section 4A*) $B\mathbin{\bar{\otimes}}\Gamma_q(H)$ *for* H *a finite-dimensional Hilbert space.*

(2) (*see Section 4C2*) $\Gamma_q(B,S\otimes H)$ *associated to the symmetric independent copies* (π_j,B,A,D) *constructed in the following way*: *take a trace preserving action* α *of* $\mathbb{Z}$ *on a finite von Neumann algebra* N. *Let* $\mathcal{H}=\langle g_j : j\ge 0\rangle$ *be the Heisenberg group, take* $\eta:\mathcal{H}\to\mathbb{Z}$ *an onto group homomorphism and define* $\beta:\mathcal{H}\curvearrowright N$ *by*

$$\beta_g(x)=\alpha_{\eta(g)}(x),\quad g\in\mathcal{H},\ x\in N.$$

Let $\mathcal{H}_1=\langle g_0,g_1\rangle$ *and take* $B=N\rtimes\mathbb{Z}=N\mathbin{\bar{\otimes}}L(\mathbb{Z})$, $A=N\rtimes\mathcal{H}_1$ *and* $S=\{1,g_1,g_1^{-1}\}$. *Define* $\pi_j:A\to D$ *by*

$$\pi_j(xu_{g_1})=\alpha_{\eta(g_j)}(x)u_{g_j},\quad \pi_j(xu_{g_0})=xu_{g_0},\quad x\in N,\ j,k\in\mathbb{N}.$$

(3) (*see Section 4D1*) $\Gamma_q(\mathbb{C},S\otimes K)$ *associated to the symmetric copies* $(\pi_j,B=\mathbb{C},A=\Gamma_{q_0}(H),D=\Gamma_q(\ell^2\otimes H))$, *where* $\pi_j(s_{q_0}(h))=s_q(e_j\otimes h)$ *and* K *is a finite-dimensional Hilbert space.*

(4) (*see Section 4D2*) $\Gamma_q(B_d,S\otimes H)$ *associated to the symmetric copies* (π_j,B_d,A_d,D_d), *where* $B_d=L(\Sigma_{[-d,0]})$, $A_d=L(\Sigma_{[-d,1]})$, $D_d=L(\Sigma_{[-d,\infty)})=\{u_\sigma:\sigma\in\Sigma_{[-d,\infty)}\}''$ *and* $S=\{1,u_{(01)}\}$ *for a fixed* $d\in\mathbb{N}\setminus\{0\}$; *here* $\Sigma_{\mathbb{Z}}$ *is the group of finite permutations on* $\mathbb{Z}$ *and for a subset* $F\subset\mathbb{Z}$, $\Sigma_F\subset\Sigma_{\mathbb{Z}}$ *is the group of finite permutations on* F *naturally embedded into* $\Sigma_{\mathbb{Z}}$. *The copies are defined by* $\pi_j(a)=u_{(1j)}au_{(1j)}$, $a\in A_d$.

(4) *(see Section 4B)* $\Gamma_q(\mathbb{C}, S \otimes H)$ *associated to the symmetric copies* $\big(\pi_j, B = \mathbb{C}, A = L(\mathbb{Z}) = \{u\}'', D = \ast_{\mathbb{N}} L(\mathbb{Z})\big)$, *where* u *is a Haar unitary, the symmetric copies* $\pi_j : A \to D$ *are defined by the relations* $\pi_j(u) = \cdots * 1 * u * 1 * \cdots$, *and* $S = \{1, u, u^*\}$.

It follows that the examples in (3), (4) *and* (5) *are strongly solid and hence solid nonamenable von Neumann algebras. In particular, they are prime von Neumann algebras. Note that when* $q = 0$ *and* H *is trivial, the example in* (5) *is* $*$*-isomorphic to* $L(\mathbb{F}_\infty)$, *thus reproving the strong solidity of the free group factors.*

Using Theorem A we also deduce the following:

Corollary C. *Let* $M_i = \Gamma_{q_i}(B_i, S_i \otimes H_i)$ *be associated with two sequences of symmetric independent copies* (π_j^i, B_i, A_i, D_i) *and two subsets* $S_i \subset A_i$, *and* $-1 < q_i < 1$, $i = 1, 2$. *Assume that* $2 \le \dim(H_i) < \infty$, $\dim_{B_i}(D_k(S_i)) \le C d^k$ *for fixed constants* $d, C > 0$ *and* B_i *are amenable for* $i = 1, 2$. *If* $M_1 \subset M_2$, *then* $B_1 \prec_{M_2} B_2$. *Moreover, if* $M_1 = M_2 = M$, *it follows that* $B_1 \prec_M B_2$ *and* $B_2 \prec_M B_1$.

This result can be regarded as an analogue of the "uniqueness of Cartan subalgebra" results in the group measure space construction setting. Note however, that even when B is abelian, it is not a MASA in $M = \Gamma_q(B, S \otimes H)$. Indeed, B always commutes with a copy of $\Gamma_q(H)$ inside $M = \Gamma_q(B, S \otimes H)$; hence it can never be maximally abelian. Thus, even when B_1 and B_2 are both abelian diffuse, we cannot avail ourselves of Popa's results [2006a, Appendix, Theorem A.1] about unitary conjugacy of Cartan subalgebras to conclude that B_1 is unitarily conjugate to B_2, so this double intertwining result is optimal in our case. Finally, we deduce some nonisomorphism and nonembedability results for generalized q-gaussians.

Corollary D. *Under the assumptions of Corollary C, if we moreover assume that*

(1) B_1 *is finite-dimensional and* B_2 *is amenable diffuse, or*

(2) B_1 *is abelian and* B_2 *is the hyperfinite* II_1 *factor,*

then $M_2 = \Gamma_{q_2}(B_2, S_2 \otimes H_2)$ *cannot be realized as a von Neumann subalgebra of* $M_1 = \Gamma_{q_1}(B_1, S_1 \otimes H_1)$. *In particular* M_1 *and* M_2 *are not* $*$*-isomorphic.*

1D. ***Comments on the proofs and structure of the article.*** Finally, a couple of words about the main ideas behind the proof of Theorem A. We mention that actually Theorem A will be derived from the technical Theorem 7.2 much along the lines of Theorem 3.1 in [Popa and Vaes 2014a], whose statement and proof can be found in Section 7. We follow the approach of Popa and Vaes [2014a; 2014b], which in turn is a development of the original ground-breaking insight in [Ozawa and Popa 2010a; 2010b]. Let $\mathcal{A} \subset M = \Gamma_q(B, S \otimes H)$ be a diffuse von Neumann subalgebra which is amenable relative to B. The two main ingredients of the proof are, just as in [Popa and Vaes 2014a]:

(1) The fact that the embedding $\mathcal{A} \subset M$ is *weakly compact relative to* B. This is the existence of a sequence of normal states viewed as unit vectors $\xi_n \in L^2(\mathcal{N})$, where $\mathcal{N} \supset M$ is a suitable (in general nontracial) von Neumann algebra, which are asymptotically invariant to the action of the "tensor double" of the normalizer of $\mathcal{A}$ in M; the existence of these states is a consequence of the weak amenability (with Cowling–Haagerup constant 1) of the pure q-gaussian von Neumann algebras $\Gamma_q(H)$.

(2) The existence of a one-parameter group of $*$-automorphisms (α_t) of a suitable dilation $\widetilde{\mathcal{N}}$ of $\mathcal{N}$ having good properties.

The proof proceeds by applying the deformation α_t to the vectors ξ_n. Then either the deformation significantly displaces the vectors, or it does not. The first case yields the amenability of $P = \mathcal{N}_M(\mathcal{A})''$ relative to B, while the second implies that $\mathcal{A} \prec_M B$, via the fact that the maps T_t (where $t \to T_t$ is the canonical semigroup of u.c.p. maps on M) are compact over B, in the terminology of Popa and Ozawa.

While it's true that conceptually we follow closely the approach of [Popa and Vaes 2014a], it has to be strongly emphasized that the technical difficulties of our approach are vastly larger. First of all, since our objects are much more elusive and complicated than cross-product von Neumann algebras, being defined as subalgebras of an ultraproduct to begin with, the proof of Theorem 5.1 (the existence of the invariant states), which is the key ingredient in the proof of the technical theorem, is riddled with daunting difficulties. Among these, constructing the von Neumann algebras involved in the deformation-rigidity argument (e.g., $\mathcal{N}$ and $\widetilde{\mathcal{N}}$ above) and the spaces on which they act was a particularly challenging task. Also, the complete boundedness of certain maps used in the proof turns out to be surprisingly nontrivial and requires the use of delicate operator-space techniques; in the pure Hilbert space setting, somewhat similar techniques have been used in [Avsec 2012; Nou 2004; 2006]. Second, and just as important, we cannot use the reduction to the "trivial-action case" (i.e., the tensor product case), as Popa and Vaes do. The reduction step plays a crucial role in their proof, because it is only in the tensor product setting that they are able to prove the relative weak compactness property and subsequently carry out the deformation-rigidity arguments. The reduction is essentially based on the use of the comultiplication map in the cross-product case. Since we have no good substitute for the comultiplication map, we cannot reduce to the tensor-product case, and hence everything becomes much more complicated and technically involved, including the standard forms of the von Neumann algebras involved, which are in general nontracial.

The article contains six sections beside the introduction, and is organized as follows: Section 2 contains some needed technical preliminaries. In Section 3 we introduce the generalized q-gaussian von Neumann algebras and prove their basic properties; among other things, we exhibit the canonical generators of $\Gamma_q(B, S \otimes H)$ (the Wick words), prove that they actually belong to the algebra and prove a very useful reduction result about them. Section 4 lists a rather wide range of examples of generalized q-gaussian von Neumann algebras constructed from a variety of symmetric independent copies. We devote Section 5 to the proof of the relative weak compactness of the embedding $\mathcal{A} \subset M$; the second half of this section contains some technical results about the complete boundedness of certain multipliers used in the proof. In Section 6 we prove that under the assumption of polynomial growth of the dimensions of the modules $D_k(S)$ over B, the natural deformation bimodules used in the technical theorem are weakly contained in $L^2(M) \otimes_B L^2(M)$, a fact which will be further used in combination with the technical theorem to derive Theorem A. The proof is based on a novel and "nondeterministic" approach. Indeed, the calculation of the deformation bimodules in the q-gaussian setting is a real challenge even in the case of pure $\Gamma_q(H)$ von Neumann algebras, see [Avsec 2012], and it becomes even more so when we allow q-gaussians with coefficients. Section 7 contains the proof of the main technical theorem and its applications. Beside many examples of strongly solid generalized q-gaussian von Neumann algebras, we also obtain some nonisomorphism and nonembedability results.

2. Preliminaries

2A. *Popa's intertwining techniques.* We will briefly review the concept of intertwining two subalgebras inside a finite von Neumann algebra, along with the main technical tools developed by Popa [2006a; 2006b]. Let (M,τ) be a finite von Neumann algebra, let $f\in\mathcal{P}(M)$ and $Q\subset fMf,\ B\subset M$ be two von Neumann subalgebras. We say that *a corner of Q can be intertwined into B inside M* and denote it by $Q\prec_M B$ (or simply $Q\prec B$) if there exist two nonzero projections $q\in Q,\ p\in B$, a nonzero partial isometry $v\in qMp$, and a $*$-homomorphism $\psi: qQq\to pBp$ such that $v\psi(x)=xv$ for all $x\in qQq$. The partial isometry v is called an intertwiner between Q and B. Popa [2006b] proved the following intertwining criterion:

Theorem 2.1 [Popa 2006b, Corollary 2.3]. *Let M be a von Neumann algebra and let $Q\subset fMf,\ B\subset M$ be diffuse subalgebras for some projection $f\in M$. Then the following are equivalent*:

(1) $Q\prec_M B$.

(2) *There exists a finite set $\mathcal{F}\subset fMf$ and $\delta>0$ such that for every unitary $v\in\mathcal{U}(Q)$ we have*

$$\sum_{x,y\in\mathcal{F}}\|E_B(xvy^*)\|_2^2\geq\delta.$$

Let (M,τ) be a finite von Neumann algebra and $\Phi: M\to M$ a normal, completely positive map. We say that Φ is subtracial if $\tau\circ\Phi\leq\tau$. If Φ is subtracial, then, due to the Schwarz inequality, we automatically have

$$\|\Phi(x)\|_2^2=\tau(\Phi(x)^*\Phi(x))\leq\tau(\Phi(x^*x))\leq\tau(x^*x)=\|x\|_2^2;$$

i.e., Φ is automatically $\|\cdot\|_2$-contractive, and hence extends to a bounded operator on $L^2(M)$ defined by

$$T_\Phi: L^2(M)\to L^2(M),\quad T_\Phi(\hat{x})=\widehat{\Phi(x)},\quad x\in M.$$

Let $B\subset(M,\tau)$ be an inclusion of finite von Neumann algebras. The basic construction (of M with B) is defined by, see, e.g., [Popa 2006a],

$$\langle M,e_B\rangle=(M\cup\{e_B\})''=(JBJ)'\subset B(L^2(M)),$$

where $L^2(M)$ is the standard form of M and $J: L^2(M)\to L^2(M)$ is the associated conjugation. The definition of the *compact ideal space* of the basic construction (more generally of any semifinite von Neumann algebra) can be found in [Popa 2006a, 1.3.3].

Definition 2.2. Let (M,τ) be a finite von Neumann algebra, $B\subset M$ a von Neumann subalgebra and $\Phi: M\to M$ a normal, completely positive, B-bimodular, subunital, subtracial map. We say that Φ is compact over B if the canonical operator $T_\Phi: L^2(M)\to L^2(M)$ belongs to the compact ideal space of the basic construction $\langle M,e_B\rangle$.

The following result is Proposition 2.7 in [Ozawa and Popa 2010a]; see also [Popa 2006a, 1.3.3].

Proposition 2.3. *Let (M, τ) be a finite von Neumann algebra and let $B, P \subset M$ be two von Neumann subalgebras. Let $\Phi : M \to M$ be a normal, completely positive, subunital, subtracial map which is compact over B and assume that*

$$\inf_{u \in \mathcal{U}(P)} \|\Phi(u)\|_2 > 0.$$

Then $P \prec_M B$.

2B. ***Bimodules over von Neumann algebras and weak containment.*** Let M, Q be two von Neumann algebras. An M-Q Hilbert bimodule $\mathcal{K}$ is simply a Hilbert space together with a pair of normal $*$-representations $\lambda : M \to B(\mathcal{K})$, $\rho : Q^{\mathrm{op}} \to B(\mathcal{K})$ with commuting ranges. To these one can associate a $*$-representation $\pi : M \otimes_{\mathrm{bin}} Q^{\mathrm{op}} \to B(\mathcal{K})$ by

$$\pi\Bigl(\sum_k x_k \otimes y_k^{\mathrm{op}}\Bigr)\xi = \sum_k \lambda(x_k)\rho(y_k^{\mathrm{op}})\xi, \quad x_k \in M,\ y_k \in Q,\ \xi \in \mathcal{K}.$$

Definition 2.4. Let M, Q be two von Neumann algebras and $\mathcal{H}, \mathcal{K}$ be two M-Q bimodules. We say that $\mathcal{K}$ *is weakly contained in* $\mathcal{H}$ and denote it by $\mathcal{K} \prec \mathcal{H}$ if $\|\pi_{\mathcal{K}}(x)\| \le \|\pi_{\mathcal{H}}(x)\|$ for all $x \in M \otimes_{\mathrm{alg}} Q$, where $\pi_{\mathcal{H}}, \pi_{\mathcal{K}}$ are the $*$-representations canonically associated to the left and right actions on $\mathcal{H}, \mathcal{K}$ respectively.

Give an M-Q bimodule $\mathcal{K}$ and an Q-N bimodule $\mathcal{H}$ we will denote by $\mathcal{K} \otimes_Q \mathcal{H}$ their Connes tensor product, which is an M-N bimodule. For the definition and basic properties of the Connes tensor product, see Sections 2.3, 2.4 in [Popa and Vaes 2014a]. The Connes tensor product is well-behaved with respect to weak containment; see [loc. cit.].

Definition 2.5 [Popa and Vaes 2014a, Definition 2.3 and Proposition 2.4]. Let (M, τ_M) and (Q, τ_Q) be finite tracial von Neumann algebras and $P \subset M$ a von Neumann subalgebra. We say that an M-Q bimodule $\mathcal{K}$ is *left P-amenable* if one of the following equivalent conditions holds:

(1) There exists a P-central state Ω on $B(\mathcal{K}) \cap (Q^{\mathrm{op}})'$ such that $\Omega|_M = \tau_M$.

(2) $L^2(M) \prec \mathcal{K} \otimes_Q \overline{\mathcal{K}}$ as M-P bimodules.

Definition 2.6. Let (M, τ) be a tracial von Neumann algebra, and let $B, P \subset M$ be two von Neumann algebras. We say that P *is amenable relative to B inside M* if one of the following equivalent conditions holds:

(1) The M-B bimodule $L^2(M)$ is left P-amenable.

(2) $L^2(M) \prec L^2(M) \otimes_B L^2(M)$ as M-P bimodules.

Remark 2.7. Let (M, τ) be a finite von Neumann algebra and $B, P \subset M$ be two von Neumann subalgebras. Let $\mathcal{K}$ be a left P-amenable M-M bimodule such that $\mathcal{K} \prec L^2(M) \otimes_B \mathcal{H}$ for some B-M bimodule $\mathcal{H}$. Then P is amenable relative to B inside M. Indeed, we have that, as M-P bimodules,

$$L^2(M) \prec \mathcal{K} \otimes_M \overline{\mathcal{K}} \prec (L^2(M) \otimes_B \mathcal{H}) \otimes_M (\overline{\mathcal{H}} \otimes_B L^2(M)) \prec L^2(M) \otimes_B L^2(M).$$

2C. ***Standard forms of nontracial von Neumann algebras.*** In some instances we will have to consider nontracial von Neumann algebras M and their standard forms. Let us recall that a (hyper-)standard form for a von Neumann algebra is given by (M, H, J, P), where $J : H \to H$ is an antilinear unitary, $P \subset H$ is a self-dual cone such that

(i) the map $M \to M'$, $x \mapsto Jx^*J$, is a $*$-anti-isomorphism acting trivially on $\mathcal{Z}(M)$;

(ii) $J\xi = \xi$ for $\xi \in P$;

(iii) $xJxJ(P) \subset P$ for $x \in M$.

The standard form of M is unique up to $*$-isomorphism; see, e.g., [Haagerup 1975]. A particularly useful way of describing the standard form of M is the abstract Haagerup $L^2(M)$ space, which we briefly describe below. The reader can find more details in [Haagerup 1979; Terp 1982; Haagerup, Junge and Xu 2010]. Let (M, φ) be a von Neumann algebra endowed with a normal semifinite faithful (n.s.f.) weight. Consider $\mathcal{M} = M \rtimes_{\sigma^\varphi} \mathbb{R}$, the cross-product von Neumann algebra of M with $\mathbb{R}$ by the modular automorphism group σ_t^φ. Then $\mathcal{M}$ is semifinite and there exists an n.s.f. trace τ on $\mathcal{M}$ such that

$$(D\hat{\varphi} : D\tau)_t = \lambda(t), \quad t \in \mathbb{R},$$

where $\hat{\varphi}$ is the dual weight, $(D\hat{\varphi} : D\tau)_t$ is the Connes cocycle and $\lambda(t)$ is the group of translations on $\mathbb{R}$. Moreover, τ is the unique n.s.f. trace on $\mathcal{M}$ which satisfies

$$\tau \circ \hat{\sigma}_t^\varphi = e^{-t}\tau, \quad t \in \mathbb{R}.$$

Given another n.s.f. weight ψ on M, denote by h_ψ the Radon–Nikodym derivative of $\hat{\psi}$ with respect to τ, i.e., the unique positive self-adjoint operator affiliated to $\mathcal{M}$ such that

$$\hat{\psi}(x) = \tau(h_\psi^{\frac{1}{2}} x h_\psi^{\frac{1}{2}}), \quad x \in \mathcal{M}_+.$$

Then the following condition holds:

$$\hat{\sigma}_t^\varphi(h_\psi) = e^{-t}h_\psi, \quad t \in \mathbb{R}.$$

Moreover, the map $\psi \mapsto h_\psi$ is a bijection from the set of n.s.f. weights on M to the set of positive self-adjoint operators affiliated to $\mathcal{M}$ which satisfy the above condition. Let $L_0(\mathcal{M}, \tau)$ be the $*$-algebra consisting of all the operators on $L^2(\mathbb{R}, H)$ which are measurable with respect to $(\mathcal{M}, \tau)$. For $p > 0$, the Haagerup $L^p(M, \varphi)$ is defined by

$$L^p(M, \varphi) = \{x \in L_0(\mathcal{M}, \tau) : \hat{\sigma}_t^\varphi(x) = e^{-\frac{t}{p}}x \text{ for all } t \in \mathbb{R}\}.$$

One can define a bicontinuous linear isomorphism from M_* to $L^1(M, \varphi)$ as the linear extension of the map

$$M_*^+ \to L^1(M, \varphi), \quad \psi \mapsto h_\psi.$$

The norm $\|\cdot\|_1$ on $L^1(M, \varphi)$ is defined by requiring that the above isomorphism be isometric. One can define a norm-1 linear functional tr on $L^1(M, \varphi)$ by $\mathrm{tr}(h_\psi) = \psi(1)$, and thus $\|h\|_1 = \mathrm{tr}(|h|)$, $h \in L^1(M, \varphi)$.

This "trace" is indeed tracial; i.e.,

$$\mathrm{tr}(xy) = \mathrm{tr}(yx) \quad \text{for } x, y \in L^2(M).$$

Let $x = u|x|$ be the polar decomposition of an element $x \in L_0(\mathcal{M}, \tau)$. Then we have

$$x \in L^p(M, \varphi) \Longleftrightarrow u \in M \quad \text{and} \quad |x| \in L^p(M, \varphi) \Longleftrightarrow u \in M \quad \text{and} \quad |x|^p \in L^1(M, \varphi).$$

This allows one to introduce the $\|\cdot\|_p$-norm on $L^p(M, \varphi)$, by $\|x\|_p = \||x|^p\|_1^{\frac{1}{p}}$ for $x \in L^p(M, \varphi)$. Let's also remark that the weight φ can be recovered from the trace. Define

$$N_\varphi = \{x \in M : \varphi(x^*x) < \infty\}, \quad M_\varphi = N_\varphi^* N_\varphi = \mathrm{span}\{y^*x : x, y \in N_\varphi\}.$$

The dual weight $\hat{\varphi}$ has a Radon–Nikodym derivative with respect to τ, which will be denoted by d_φ. Then for every $x \in M_\varphi$ the operator $d_\varphi^{\frac{1}{2}} x d_\varphi^{\frac{1}{2}}$ is closable, its closure belongs to $L^1(M, \varphi)$ and we have the relation

$$\varphi(x) = \mathrm{tr}(d_\varphi^{\frac{1}{2}} x d_\varphi^{\frac{1}{2}}), \quad x \in M_\varphi.$$

If φ is a bounded functional, then $d_\varphi \in L^1(M, \varphi)$ and the above identity becomes

$$\varphi(x) = \mathrm{tr}(d_\varphi^{\frac{1}{2}} x d_\varphi^{\frac{1}{2}}) = \mathrm{tr}(x d_\varphi), \quad x \in M.$$

The Haagerup space $L^p(M, \varphi)$ does not depend on the choice of the n.s.f. weight φ up to isomorphism; hence it can simply be denoted by $L^p(M)$. It's easy to see that M is naturally represented in standard form on the Haagerup space $L^2(M)$ via the obvious left and right actions. When M is finite and τ is a faithful trace on M, the Haagerup space $L^2(M) = L^2(M, \tau)$ coincides with the usual one.

2D. *W^*-Hilbert modules.* We also have to recall some facts about (right) Hilbert W^*-modules. According to [Paschke 1973; 1974], see also [Junge and Sherman 2005], a right Hilbert C^*-module X over a *von Neumann algebra M* is self-dual if and only if admits a module basis, i.e., a family $\{\xi_\alpha\} \subset X$ such that

$$X = \overline{\mathrm{span}} \sum_\alpha \xi_\alpha M \quad \text{and} \quad \langle \xi_\alpha, \xi_\beta \rangle = \delta_{\alpha\beta} e_\alpha \in \mathcal{P}(M).$$

Here, $\langle \cdot, \cdot \rangle$ denotes the M-valued inner product. In this situation, there exists an index set I, a projection $e \in B(\ell_2(I)) \bar{\otimes} M$, and a right module isomorphism $u : X \to e(\ell_2(I)^c \bar{\otimes} M)$. Indeed, for a basis ξ_α with $\langle \xi_\alpha, \xi_\alpha \rangle = e_\alpha$, the map u is given by $u\big(\sum_\alpha \xi_\alpha m_\alpha\big) = [e_\alpha m_\alpha]$. Here $\ell_2(I)^c \bar{\otimes} M$ denotes the space of strongly convergent columns indexed by I. Then it is easy to see that the C^*-algebra $\mathcal{L}(X)$ of adjointable operators on X is indeed a von Neumann algebra, and isomorphic to $e(B(\ell_2(I)) \bar{\otimes} M)e$. Moreover, the M-compact operators $\mathcal{K}(X)$ spanned by the maps $\Phi_{\xi,\eta}(\zeta) = \xi\langle \eta, \zeta \rangle$ are weakly dense in $\mathcal{L}(X)$, because $K(\ell_2(I)) \otimes_{\min} M$ is weakly dense in $B(\ell_2(I)) \bar{\otimes} M$. With the help of a normal faithful state, we can complete X to the Hilbert space $L_2(X, \phi)$ with inner product $(\xi, \eta) = \phi(\langle \xi, \eta \rangle)$. Let $\iota_\phi : X \to L_2(X, \phi)$ be the inclusion map. Then

$$\pi : \mathcal{L}(X) \to B(L_2(X, \phi)), \quad \pi(T)(\iota_\phi(x)) = \iota_\phi(Tx),$$

defines a normal faithful $*$-homomorphism such that

$$\pi(\mathcal{L}(X)) = B(L_2(X,\phi)) \cap (M^{\mathrm{op}})'.$$

This is indeed very easy to check for $\mathcal{L}(X) = e(B(\ell_2(I)) \mathbin{\overline{\otimes}} M)e$. See [Paschke 1973; 1974; Junge and Sherman 2005] for more details and references.

3. The generalized gaussian von Neumann algebras with coefficients: definition and basic properties

Throughout this section we will freely use the basic properties of the pure Hilbert space q-gaussian von Neumann algebras $\Gamma_q(H)$, as they can be found in Section 4 of [Junge, Longfield and Udrea 2014]; see also [Avsec 2012]. The following result is due to Voiculescu, Dykema and Nica [1992].

Proposition 3.1. *Let (M,φ) and (N,ψ) be two von Neumann algebras endowed with faithful normal tracial states. Let $(x_i)_{i=1}^{\infty}$ and $(y_j)_{j=1}^{\infty}$ be countable systems of generators for M and N, respectively. Assume that for every $m \geq 1$, every $i_1,\dots,i_m \in \mathbb{N}$ and every $\varepsilon_i \in \{1,*\}$ we have*

$$\varphi(x_{i_1}^{\varepsilon_1}\cdots x_{i_m}^{\varepsilon_m}) = \psi(y_{i_1}^{\varepsilon_1}\cdots y_{i_m}^{\varepsilon_m}).$$

Then there exists a $*$-isomorphism $\pi : M \to N$ such that $\psi \circ \pi = \varphi$ and $\pi(x_i) = y_i$ for all $i \geq 1$.

Definition 3.2. Let A and D be two finite tracial von Neumann algebras and B a von Neumann subalgebra of $A \cap D$. Let $\pi_j : A \to D$, $j \in \mathbb{N}$, be a countable family of unital, normal, faithful, trace-preserving $*$-homomorphisms. The 4-tuple (π_j, B, A, D) is called a sequence of *symmetric independent copies* of A if the following properties hold:

(1) $\pi_j|_B = \mathrm{id}_B$ for all j.

(2) $E_B(\pi_{j_1}(a_1)\cdots\pi_{j_m}(a_m)) = E_B(\pi_{\sigma(j_1)}(a_1)\cdots\pi_{\sigma(j_m)}(a_m))$ for all finite permutations σ on $\mathbb{N}$, all indices $j_1,\dots,j_m$ in $\mathbb{N}$ and all $a_1,\dots,a_m$ in A, where $E_B : D \to B$ is the canonical trace-preserving conditional expectation.

(3) For $i \in \mathbb{N}$ set $A_i = \pi_i(A) \subset D$ and for $I \subset \mathbb{N}$, set $A_I = \bigvee_{i\in I}\pi_i(A) = \bigvee_{i\in I} A_i \subset D$ (by convention, set $A_\varnothing = B$); then, for any finite subsets $I \subset J \subset \mathbb{N}$, $j \notin J$, $d \in A_I$ and $a, a' \in A$, we have

$$E_{A_I}(\pi_j(a)d\pi_j(a')) = E_{A_J}(\pi_j(a)d\pi_j(a')),$$

where $E_{A_I} : D \to A_I$ is the canonical conditional expectation.

(4) For any finite subsets $I, J \subset \mathbb{N}$, we have $E_{A_I}E_{A_J} = E_{A_{I\cap J}}$. Note that this automatically implies $E_{A_I}E_{A_J} = E_{A_J}E_{A_I} = E_{A_I \cap A_J}$ and in particular $A_I \cap A_J = A_{I\cap J}$.

(5) $A_{\mathbb{N}} = D$.

If the 4-tuple (π_j, B, A, D) only satisfies axioms (1) and (2), we call it a sequence of symmetric copies.

The role played by the copies $\pi_j(A)$ is analogous to that of tensor copies in a classical product probability space; in fact such an infinite product probability space over a commutative or noncommutative base constitutes the first obvious example of symmetric independent copies. To be more precise, let (A, τ) be a tracial von Neumann algebra, let $D = \bigotimes_{i\in\mathbb{N}}(A_i, \tau_i)$, where $(A_i, \tau_i) = (A, \tau)$ for all $i \in \mathbb{N}$, let π_j be the obvious embedding of A in D as the j-th tensor copy and let $B = \mathbb{C}$. Then all the axioms (1) to (5) are satisfied. In particular, one could take $(A, \tau) = \big(L^\infty(X, \mu), \int_X d\mu\big)$ for a probability measure space (X, μ). Axiom (2), while convenient because it greatly simplifies some of our technical computations, doesn't seem to be indispensable to the development of a general theory of B-valued q-gaussian von Neumann algebras. Indeed, the generalized q-gaussian von Neumann algebras can still be introduced in the presence of a weaker "subsymmetry" assumption, but the technicalities become even more cumbersome, and it is unclear whether some of our results can still be obtained. Axioms (3) and (4) can both be viewed as describing some sort of independence of the copies over B, with (4) being the more obvious one, since for example it gives that for $I \cap J = \varnothing$, we have $E_{A_I} E_{A_J} = E_B$. In the case of an abelian D and $B = \mathbb{C}$, this amounts to classical probabilistic independence. Axiom (5), while added for completeness, can always be made redundant by shrinking the algebra D.

In what follows, the expectations E_{A_I} will be denoted by E_I.

Proposition 3.3. *Let (π_j, B, A, D) be a sequence of symmetric copies. Let $\Sigma = \mathbb{S}(\infty)$ be the group of finite permutations on $\mathbb{N} \setminus \{0\}$. Then for every $\sigma \in \Sigma$ there exists a trace-preserving automorphism α_σ of $D_0 = A_{\mathbb{N}\setminus\{0\}} \subset D$ such that*

$$\alpha_\sigma(\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)) = \pi_{\sigma(j_1)}(x_1)\cdots\pi_{\sigma(j_m)}(x_m) \quad \textit{for all } x_1, \dots, x_m \in A \textit{ and } j_1, \dots, j_m \in \mathbb{N}.$$

Moreover,

$$\Sigma \to \mathrm{Aut}(D_0, \tau), \quad \sigma \mapsto \alpha_\sigma,$$

is an action of Σ on D_0 by trace-preserving automorphisms. Additionally, *if the symmetric copies satisfy axiom* (4), *then the fixed points algebra of this action is B.*

Proof. The map $V_\sigma : L^2(D_0) \to L^2(D_0)$ defined by

$$\sum \pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m) \mapsto \sum \pi_{\sigma(j_1)}(x_1)\cdots\pi_{\sigma(j_m)}(x_m)$$

is easily seen to be a well-defined unitary because of axiom (2). Then $\alpha_\sigma = \mathrm{Ad}(V_\sigma)|_D$ is a trace-preserving automorphism of D which satisfies the required condition. The verification of the second statement is straightforward and we leave it to the reader. □

Symmetric copies can also be introduced in the following alternative way, which is a converse to the previous proposition: assume that $\alpha : \Sigma \to \mathrm{Aut}(D, \tau)$ is a trace-preserving action by $*$-automorphism of the finite von Neumann algebra D, where Σ is now the finite permutation group on $\mathbb{N} \cup \{0\}$ instead of $\mathbb{N}$. Denote by $B = D^\Sigma$ the fixed points algebra of this action. Set

$$\Sigma_0 = \mathrm{Stab}_\Sigma(0) = \{\sigma \in \Sigma : \sigma(0) = 0\},$$
$$A = D^{\Sigma_0} = \{d \in D : \alpha_\sigma(d) = d \text{ for all } \sigma \in \Sigma_0\}.$$

Note that $\Sigma_0 \subset \Sigma$ is a subgroup isomorphic to $\mathbb{S}(\infty)$ and that $B \subset A \subset D$. For every $j \geq 1$, define $\pi_j : A \to D$ by the formula $\pi_j(a) = \alpha_{(0j)}(a)$, $a \in A$, where $(0j) \in \Sigma$ is the transposition interchanging

0 and j. Then (π_j, B, A, D) represents a sequence of symmetric copies. Indeed, for any $j \geq 1$ and $b \in B$ we have $\pi_j(b) = \alpha_{(0j)}(b) = b$ because B is the fixed points algebra of the action α, so (1) is true. Note that $\alpha_\sigma(a) = a$ for every $\sigma \in \Sigma_0$ and $a \in A$ and $E_B \circ \alpha_\sigma = E_B$ for all $\sigma \in \Sigma$, due to (1) and the facts that α is trace-preserving and the trace-preserving conditional expectation $E_B : D \to B$ is unique. Then for every $\sigma \in \Sigma_0 \cong \mathbb{S}(\infty)$ and for all $j_1, \ldots, j_m \geq 1$ and $a_1, \ldots, a_m \in A$ we have

$$\begin{aligned} E_B(\pi_{\sigma(j_1)}(a_1) \cdots \pi_{\sigma(j_m)}(a_m)) &= E_B(\alpha_{(0\sigma(j_1))}(a_1) \cdots \alpha_{(0\sigma(j_m))}(a_m)) \\ &= E_B(\alpha_{\sigma(0j_1)\sigma^{-1}}(a_1) \cdots \alpha_{\sigma(0j_m)\sigma^{-1}}(a_m)) \\ &= E_B\big((\alpha_\sigma \circ \alpha_{(0j_1)} \circ \alpha_{\sigma^{-1}})(a_1) \cdots (\alpha_\sigma \circ \alpha_{(0j_m)} \circ \alpha_{\sigma^{-1}})(a_m)\big) \\ &= E_B(\alpha_\sigma(\alpha_{(0j_1)}(a_1) \cdots \alpha_{(0j_m)}(a_m))) \\ &= E_B(\alpha_{(0j_1)}(a_1) \cdots \alpha_{(0j_m)}(a_m)) \\ &= E_B(\pi_{j_1}(a_1) \cdots \pi_{j_m}(a_m)), \end{aligned}$$

so (2) is also true. As noted before, we can also assume without loss of generality that

$$D = \bigvee_{j \geq 1} \pi_j(A) = \bigvee_{j \geq 1} A_j,$$

by simply replacing D with a von Neumann subalgebra.

Notation. Let $(j_1, \ldots, j_m)$ be an m-tuple with $1 \leq j_k \leq n,\ 1 \leq k \leq m$. We denote by $P(m)$ the set of partitions of $\{1, \ldots, m\}$ and by $\dot{0}, \dot{1}$ the finest and the coarsest partitions in $P(m)$, respectively. The notation $P_{1,2}(m)$ stands for the collection of all the partitions of $\{1, \ldots, m\}$ consisting only of singletons and pairs. For $\sigma \in P(m)$, we say that

(1) $(j_1, \ldots, j_m) \leq \sigma$ if $j_i = j_k$ whenever $i, k \in A \in \sigma$;

(2) $(j_1, \ldots, j_m) \geq \sigma$ if $j_i = j_k$ implies that there exists an $A \in \sigma$ with $i, k \in A$;

(3) $(j_1, \ldots, j_m) = \sigma$ if $j_i = j_k$ exactly when there exists an $A \in \sigma$ such that $i, k \in A$.

Given an m-tuple $(j_1, \ldots, j_m) = \dot{0}$ with $1 \leq j_k \leq n$ for $1 \leq k \leq m$, we set $\alpha_{j_1,\ldots,j_m} = \alpha_{\sigma_{j_1,\ldots,j_m}}$, where $\sigma_{j_1,\ldots,j_m}(i) = j_i,\ 1 \leq i \leq m$.

Definition 3.4. Let (π_j, B, A, D) be a sequence of symmetric independent copies, S a subset of A such that $1 \in S = S^*$, H a Hilbert space and ω a free ultrafilter on $\mathbb{N}$. Denote by $\{e_j\}$ the canonical orthonormal basis of $\ell^2 = \ell^2(\mathbb{N})$. Let $-1 < q < 1$. Define

$$\Gamma_q^0(B, S \otimes H) = (B \cup \{s_q(a,h) : a \in S,\ h \in H\})'' \subset (\Gamma_q(\ell^2 \otimes H) \,\overline{\otimes}\, D)^\omega,$$

where

$$s_q(a,h) = \left(n^{-\frac{1}{2}} \sum_{j=1}^{n} s_q(e_j \otimes h) \otimes \pi_j(a)\right)_n.$$

Finally define

$$\Gamma_q(B, S \otimes H) = (E_{\Gamma_q(\ell_n^2 \otimes H)} \otimes \mathrm{id})_n(\Gamma_q^0(B, S \otimes K)),$$

where K is an infinite-dimensional Hilbert space containing H, $\ell_n^2 = \text{span}\{e_1, \dots, e_n\}$ and for each n

$$E_{\Gamma_q(\ell_n^2 \otimes H)} : \Gamma_q(\ell^2 \otimes K) \to \Gamma_q(\ell_n^2 \otimes H)$$

is the canonical conditional expectation.

As q will be fixed throughout this section, we will simply use the notation $s(x,h)$ instead of $s_q(x,h)$ from now on.

Remark 3.5. Due to functoriality, the definition of $\Gamma_q(B, S \otimes H)$ does not depend on the particular choice of $K \supset H$. When H is infinite-dimensional $\Gamma_q(B, S \otimes H) = \Gamma_q^0(B, S \otimes H)$.

Remark 3.6. $\Gamma_q^0(B, S \otimes H) = (\{s(a,h) : a \in B \cup S,\ h \in H\})'' \subset (\Gamma_q(\ell^2 \otimes H) \,\bar{\otimes}\, D)^\omega$.

Remark 3.7. $\Gamma_q(B, S \otimes H)$ is a von Neumann algebra. Indeed, since the map

$$E = (E_{\Gamma_q(\ell_n^2 \otimes H)} \otimes \mathrm{id})_n : (\Gamma_q(\ell^2 \otimes K) \,\bar{\otimes}\, D)^\omega \to (\Gamma_q(\ell^2 \otimes H) \,\bar{\otimes}\, D)^\omega$$

is a normal linear projection (i.e., idempotent map) of norm 1, it follows that $\Gamma_q(B, S \otimes H)$ is an ultraweakly closed, self-adjoint subspace of $(\Gamma_q(\ell^2 \otimes H) \,\bar{\otimes}\, D)^\omega$ containing the identity. It's straightforward to see that the map E has the following bimodularity property:

$$E(x)E(y)E(z) = E(E(x)yE(z)) \quad \text{for all } x, y, z \in \Gamma_q^0(B, S \otimes K).$$

Thus, for $x, y \in \Gamma_q^0(B, S \otimes K)$ we have

$$E(x)E(y) = E(E(x)y) \in \Gamma_q(B, S \otimes H).$$

The canonical generators $s_q(a,h)$ are not easy to work with in a variety of situations. The classical q-gaussians possess a system of generators, the so-called Wick words, whose linear span is an ultraweakly dense $*$-subalgebra. Generalized q-gaussians also have such a well-behaved system of linear generators, which will be called Wick words by analogy with the classical case. In order to find these Wick words let us first define, for every $n \in \mathbb{N}$, $x \in A$ and $h \in H$,

$$u_n(x,h) = n^{-\frac{1}{2}} \left(\sum_{j=1}^{n} s(e_j \otimes h) \otimes \pi_j(x) \right) \in \Gamma_q(\ell^2 \otimes H) \,\bar{\otimes}\, D.$$

It's easy to see that $s(x,h) = (u_n(x,h))_n \in (\Gamma_q(\ell^2 \otimes H) \,\bar{\otimes}\, D)^\omega$ for $x \in A, h \in H$. For $x_1, \dots, x_m \in BSB = \{b_1 a b_2 : b_1, b_2 \in B,\ a \in S\}$ and $h_1, \dots, h_m \in H$ we will analyze the product

$$\begin{aligned} u_n(x_1,h_1)\cdots u_n(x_m,h_m) &= n^{-\frac{m}{2}} \sum_{1 \le j_1, \dots, j_m \le n} s(e_{j_1} \otimes h_1) \cdots s(e_{j_m} \otimes h_m) \otimes \pi_{j_1}(x_1)\pi_{j_2}(x_2) \cdots \pi_{j_m}(x_m) \\ &= \sum_{\sigma \in P(m)} \left(n^{-\frac{m}{2}} \sum_{\substack{(j_1,\dots,j_m)=\sigma \\ 1 \le j_1,\dots,j_m \le n}} s(e_{j_1} \otimes h_1) \cdots s(e_{j_m} \otimes h_m) \otimes \pi_{j_1}(x_1) \cdots \pi_{j_m}(x_m) \right). \end{aligned}$$

For $\sigma \in P(m)$ let's define

$$x_\sigma^n(x_1,h_1,\dots,x_m,h_m) = n^{-\frac{m}{2}} \sum_{\substack{1\le j_1,\dots,j_m\le n\\ (j_1,\dots,j_m)=\sigma}} s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\otimes \pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m),$$

and

$$x_\sigma(x_1,h_1,\dots,x_m,h_m) = (x_\sigma^n(x_1,h_1,\dots,x_m,h_m))_n \in (\Gamma_q(\ell^2\otimes H)\bar{\otimes} D)^\omega.$$

To keep the notation less cumbersome, we will omit the parameters x_k, h_k whenever they are clearly understood from the context. Next we see that

$$u_n(x_1,h_1)\cdots u_n(x_m,h_m) = \sum_{\sigma\in P(m)} x_\sigma^n,$$

and also

$$s(x_1,h_1)\cdots s(x_m,h_m) = (u_n(x_1,h_1)\cdots u_n(x_m,h_m))_n = \sum_{\sigma\in P(m)} x_\sigma.$$

Lemma 3.8. *Let (π_j, B, A, D) be a sequence of symmetric copies. Then*:

(o) $\sup_n \|x_\sigma^n\|_\infty < \infty$ *for all $m \ge 1$ and $\sigma \in P_{1,2}(m)$.*

(i) *If $\sigma \notin P_{1,2}(m)$ and $0 < p < \infty$ then*

$$\lim_n \|x_\sigma^n\|_p = 0.$$

In particular $s(x_1,h_1)\cdots s(x_m,h_m) = \sum_{\sigma\in P_{1,2}(m)} x_\sigma$.

Proof. The proof is the same as that of Proposition 4.1 in [Junge, Longfield and Udrea 2014]. □

Proposition 3.9. *We have the following convolution formula for the multiplication of Wick words*:

$$x_\sigma(x_1,h_1,\dots,x_m,h_m)\, x_\theta(y_1,k_1,\dots,y_{m'},k_{m'}) = \sum_{\substack{\gamma\in P_{1,2}(m+m')\\ \gamma_p|_{1,\cdots,m}=\sigma_p,\ \gamma_p|_{1,\cdots,m'}=\theta_p}} x_\gamma(x_1,h_1,\dots,y_{m'},k_{m'}).$$

Moreover, *item* (i) *in the lemma above shows that in the summation we can restrict ourselves to pair-singleton partitions whose only additional pairings are between the singletons of σ and θ. In particular*, *the linear span of the Wick words is a $*$-algebra.*

Proof. We have

$$\begin{aligned}
&x_\sigma(x_1,h_1,\dots,x_m,h_m)\, x_\theta(y_1,k_1,\dots,y_{m'},k_{m'})\\
&\quad= \Bigg(n^{-\frac{m+m'}{2}} \sum_{\substack{(j_1,\dots,j_m)=\sigma\\ (l_1,\dots,l_{m'})=\theta}} s(e_{j_1}\otimes h_1)\cdots s(e_{l_{m'}}\otimes k_{m'})\otimes\pi_{j_1}(x_1)\cdots\pi_{l_{m'}}(y_{m'})\Bigg)\\
&\quad= \sum_{\substack{\gamma\in P_{1,2}(m+m')\\ \gamma_p|_{\{1,\dots,m\}}=\sigma_p,\ \gamma_p|_{\{1,\dots,m'\}}=\theta_p}} \Bigg(n^{-\frac{m+m'}{2}} \sum_{(j_1,\dots,l_{m'})=\gamma} s(e_{j_1}\otimes h_1)\cdots s(e_{l_{m'}}\otimes k_{m'})\otimes\pi_{j_1}(x_1)\cdots\pi_{l_{m'}}(y_{m'})\Bigg)\\
&\quad= \sum_{\substack{\gamma\in P_{1,2}(m+m')\\ \gamma_p|_{\{1,\dots,m\}}=\sigma_p,\ \gamma_p|_{\{1,\dots,m'\}}=\theta_p}} x_\gamma(x_1,h_1,\dots,y_{m'},k_{m'}).
\end{aligned}$$

Now if $\gamma \in P_{1,2}(m+m')$ connects a singleton in σ with a leg of a pair in θ or the leg of pair in σ with either a singleton or a leg of a pair in θ, the resulting x_γ is associated to a partition containing a 3-set or a 4-set and hence vanishes according to Lemma 3.8. So in the above sum we may only allow γ's which preserve the pair sets of both σ and θ and can only additionally pair singletons "on different sides of the marker", which ends the proof. $\square$

Our next result provides a reduction method for the Wick words.

Lemma 3.10. *Let $\pi_j : A \to D$ be symmetric independent copies, and $1 \in S = S^* \subset A$. Let $x_1, \dots, x_m \in BSB$, $\sigma \in P_{1,2}(m)$ having s singletons and p pairs and $\phi : \{1, \dots, m\} \to \{1, \dots, s+p\}$ which encodes σ, i.e., $\phi(k_t) = t$ for every singleton $\{k_t\} \in \sigma$, $1 \le t \le s$, and $\phi(k'_t) = \phi(k''_t) = t+s$ for every pair $\{k'_t, k''_t\} \in \sigma$, $1 \le t \le p$. Consider (ε_k) a sequence of Rademacher variables, i.e., Bernoulli independent random variables on a probability space (X, μ) satisfying $\varepsilon_k : X \to \{\pm 1\}$, $\mathbb{E}(\varepsilon_k = 1) = \mathbb{E}(\varepsilon_k = -1) = \frac{1}{2}$. Then*

$$\Big\| \sum_{(l_1,\dots,l_{s+p})=\dot{0}} \varepsilon_{l_1} \cdots \varepsilon_{l_s} \otimes \big(\pi_{l_{\phi(1)}}(x_1) \cdots \pi_{l_{\phi(m)}}(x_m) - E_{l_1,\dots,l_s}(\pi_{l_{\phi(1)}}(x_1) \cdots \pi_{l_{\phi(m)}}(x_m))\big) \Big\|_2 \le C(m, x_j) n^{\frac{m-1}{2}}.$$

In particular we have

$$\begin{aligned} \Big(n^{-\frac{m}{2}} &\sum_{(l_1,\dots,l_s,l_{s+1},\dots,l_{s+p})=\dot{0}} \varepsilon_{l_1} \cdots \varepsilon_{l_s} \otimes (\pi_{l_{\phi(1)}}(x_1) \cdots \pi_{l_{\phi(m)}}(x_m))\Big) \\ &= \Big(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} \varepsilon_{l_1} \cdots \varepsilon_{l_s} \otimes E_{l_1,\dots,l_s}(\pi_{l_{\phi(1)}}(x_1) \cdots \pi_{l_{\phi(m)}}(x_m))\Big) \\ &= \Big(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} \varepsilon_{l_1} \cdots \varepsilon_{l_s} \otimes \alpha_{l_1,\dots,l_s}(F_\sigma(x_1, \dots, x_m))\Big), \end{aligned}$$

where $F_\sigma(x_1, \dots, x_m) = E_{1,\dots,s}(\pi_{\phi(1)}(x_1) \cdots \pi_{\phi(m)}(x_m))$ and the second equality takes place in $(L^\infty(X) \bar{\otimes} D)^\omega$.

Proof. Throughout the proof we endow $L^\infty(X) \bar{\otimes} D$ with the natural trace $\mu \otimes \tau$, where τ is the faithful trace on D. The $\|\cdot\|_2$ in the first statement is the one corresponding to $\mu \otimes \tau$. The approach we take is somewhat similar to the one in [Junge and Zeng 2015].

Step 1. Let $x_1, \dots, x_m \in BSB$ and n be fixed. Consider

$$\Omega_n = \{(C_1, \dots, C_{s+p}) : C_1 \sqcup \cdots \sqcup C_{s+p} = \{1, \dots, n\},\ C_i \neq \varnothing \text{ for all } i\}.$$

Make Ω_n into a probability space with the normalized counting measure. For every $(s+p)$-tuple $(l_1, \dots, l_{s+p}) = \dot{0}$, consider the indicator function $\delta_{l_1,\dots,l_{s+p}} : \Omega_n \to \{0, 1\}$ which is 1 if $l_i \in C_i$ for all $1 \le i \le s+p$ and 0 otherwise. According to the proof of Lemma 3.6 in [Junge and Zeng 2015], $\mathbb{E}(\delta_{l_1,\dots,l_{s+p}}) = (s+p)^{-s-p} = C$. Put

$$F(l_1, \dots, l_{s+p}) = \varepsilon_{l_1} \cdots \varepsilon_{l_s} \otimes (\pi_{l_{\phi(1)}}(x_1) \cdots \pi_{l_{\phi(m)}}(x_m)) - E_{l_1,\dots,l_s}(\pi_{l_{\phi(1)}}(x_1) \cdots \pi_{l_{\phi(m)}}(x_m)).$$

Then we have

$$\begin{aligned}
&\Big\| \sum_{(l_1,\dots,l_{s+p})=\dot{0}} F(l_1,\dots,l_{s+p}) \Big\|_2 \\
&\qquad = C^{-1} \Big\| C \sum_{(l_1,\dots,l_{s+p})=\dot{0}} F(l_1,\dots,l_{s+p}) \Big\|_2 \\
&\qquad = C^{-1} \Big\| \sum_{(l_1,\dots,l_{s+p})=\dot{0}} CF(l_1,\dots,l_{s+p}) \Big\|_2 \\
&\qquad = C^{-1} \Big\| \sum_{(l_1,\dots,l_{s+p})=\dot{0}} \mathbb{E}(\delta_{l_1,\dots,l_{s+p}}) F(l_1,\dots,l_{s+p}) \Big\|_2 \\
&\qquad = C^{-1} \Big\| \frac{1}{|\Omega_n|} \sum_{(l_1,\dots,l_{s+p})=\dot{0}} \ \sum_{(C_1,\dots,C_{s+p})\in\Omega_n} \delta_{l_1,\dots,l_{s+p}}((C_1,\dots,C_{s+p})) F(l_1,\dots,l_{s+p}) \Big\|_2 \\
&\qquad = C^{-1} \Big\| \frac{1}{|\Omega_n|} \sum_{(C_1,\dots,C_{s+p})\in\Omega_n} \ \sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}} F(l_1,\dots,l_{s+p}) \Big\|_2 = C^{-1} \|\mathbb{E}(G)\|_2 \\
&\qquad \le C \sup_{(C_1,\dots,C_{s+p})\in\Omega_n} \|G((C_1,\dots,C_{s+p}))\|_2 \\
&\qquad = C \sup_{(C_1,\dots,C_{s+p})\in\Omega_n} \Big\| \sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}} F(l_1,\dots,l_{s+p}) \Big\|_2,
\end{aligned}$$

where we define $G : \Omega_n \to L^\infty(X) \,\bar{\otimes}\, D$ by

$$G((C_1,\dots,C_{s+p})) = \sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}} F(l_1,\dots,l_{s+p}).$$

<u>Step 2</u>. It suffices thus to estimate $\big\|\sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}} F(l_1,\dots,l_{s+p})\big\|_2$ for a fixed nondegenerate partition $C_1,\dots,C_{s+p}$ of $\{1,\dots,n\}$. Fix such an arbitrary partition. We define the sets

$$I_l = C_1 \cup \cdots \cup C_{s+p-1} \cup (\{1,\dots,l\} \cap C_{s+p})$$

and for $l \in C_{s+p}$

$$\begin{aligned}
d_l = \sum_{\substack{l_1\in C_1,\dots,l_s\in C_s \\ l_{s+1}\in C_{s+1},\dots,l_{p-1}\in C_{s+p-1}}} & \varepsilon_{l_1}\cdots\varepsilon_{l_s} \otimes \big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_l(x_{k_p'})\cdots\pi_l(x_{k_p''})\cdots\pi_{l_{\phi(m)}}(x_m) \\
& - E_{I_{l-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_l(x_{k_p'})\cdots\pi_l(x_{k_p''})\cdots\pi_{l_{\phi(m)}}(x_m))\big).
\end{aligned}$$

Note that $D_l = L^\infty(X) \,\bar{\otimes}\, A_{I_l}$, $l \in C_{s+p}$, form an increasing finite sequence of von Neumann subalgebras of $L^\infty(X) \,\bar{\otimes}\, D$. Now $d_l \in D_l$ and $E_{D_{l-1}}(d_l) = 0$ for all $l \in C_{s+p}$. The orthogonality together with the Cauchy–Schwarz inequality yields

$$\Big\| \sum_{l\in C_{s+p}} d_l \Big\|_2 \le n^{\frac{1}{2}} \sup_{l\in C_{s+p}} \|d_l\|_2.$$

On the other hand, since the products $\varepsilon_{l_1}\cdots\varepsilon_{l_s}$ are mutually orthogonal for different s-tuples $(l_1,\dots,l_s)$, we see that

$$\begin{aligned}
\|d_l\|_2 &= \Big\| \sum_{\substack{l_1\in C_1,\dots,l_s\in C_s\\ l_{s+1}\in C_{s+1},\dots,l_{s+p-1}\in C_{s+p-1}}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_l(x_{k_p'})\cdots\pi_l(x_{k_p''})\cdots\pi_{l_{\phi(m)}}(x_m) \\
&\qquad - E_{I_{l-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_l(x_{k_p'})\cdots\pi_l(x_{k_p''})\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2 \\
&= \Big\| \sum_{l_1\in C_1,\dots,l_s\in C_s} \varepsilon_{l_1}\cdots\varepsilon_{l_s} \\
&\qquad \otimes\Big(\sum_{l_{s+1}\in C_{s+1},\dots,l_{s+p-1}\in C_{s+p-1}} \big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_l(x_{k_p'})\cdots\pi_l(x_{k_p''})\cdots\pi_{l_{\phi(m)}}(x_m) \\
&\qquad - E_{I_{l-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_l(x_{k_p'})\cdots\pi_l(x_{k_p''})\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big)\Big\|_2 \\
&\le n^{\frac{s}{2}} \Big\| \sum_{l_{s+1}\in C_{s+1},\dots,l_{s+p-1}\in C_{s+p-1}} \big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m) - E_{I_{l-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2 \\
&\le n^{\frac{s}{2}} n^{p-1} \|\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)\|_\infty \le n^{\frac{m-2}{2}} \|x_1\|_\infty\cdots\|x_m\|_\infty.
\end{aligned}$$

According to axiom (3) we have

$$\begin{aligned}
&E_{I_{l-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_l(x_{k_p'})\cdots\pi_l(x_{k_p''})\cdots\pi_{l_{\phi(m)}}(x_m)) \\
&\qquad = E_{C_1,\dots,C_{s+p-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_l(x_{k_p'})\cdots\pi_l(x_{k_p''})\cdots\pi_{l_{\phi(m)}}(x_m));
\end{aligned}$$

hence

$$\begin{aligned}
&\Big\| \sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m) - E_{C_1\cup\cdots\cup C_{s+p-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2 \\
&\qquad = \Big\| \sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m) - E_{I_{l-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2 \\
&\qquad = \Big\| \sum_{l\in C_{s+p}} d_l \Big\|_2 \\
&\qquad \le \|x_1\|_\infty\cdots\|x_m\|_\infty n^{\frac{m-1}{2}} = C'(x_1,\dots,x_m) n^{\frac{m-1}{2}}.
\end{aligned}$$

Steps 1 and 2 so far imply that

$$\begin{aligned}
&\Big\| \sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m) \\
&\qquad - E_{C_1\cup\cdots\cup C_{s+p-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2 \le C' n^{\frac{m-1}{2}}.
\end{aligned}$$

Step 3. Now we may proceed inductively. Set $y=\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)$. Then, using axiom (4) and because the conditional expectations commute, we see that

$$\begin{aligned}
&y-E_{C_1\cup\cdots\cup C_{s+p-2}}(y)\\
&\quad= y-E_{C_1\cup\cdots\cup C_{s+p-1}}(y)+E_{C_1\cup\cdots\cup C_{s+p-1}}(y)-E_{C_1\cup\cdots\cup C_{s+p-2}}(y)\\
&\quad= y-E_{C_1\cup\cdots\cup C_{s+p-1}}(y)+E_{C_1\cup\cdots\cup C_{s+p-1}}(y)-E_{C_1\cup\cdots\cup C_{s+p-2}\cup C_{s+p}}(E_{C_1\cup\cdots\cup C_{s+p-1}}(y))\\
&\quad= y-E_{C_1\cup\cdots\cup C_{s+p-1}}(y)+E_{C_1\cup\cdots\cup C_{s+p-1}}(y-E_{C_1\cup\cdots\cup C_{s+p-2}\cup C_{s+p}}(y)).
\end{aligned}$$

Using the previous steps and the fact that the conditional expectations are $\|\cdot\|_2$-contractive, we obtain

$$\begin{aligned}
&\Big\|\sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)-E_{C_1\cup\cdots\cup C_{s+p-2}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2\\
&\le\Big\|\sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\\
&\qquad\qquad\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)-E_{C_1\cup\cdots\cup C_{s+p-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2\\
&\quad+\Big\|(\mathrm{id}\otimes E_{C_1\cup\cdots\cup C_{s+p-1}})\\
&\qquad\Big(\sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)\\
&\qquad\qquad-E_{C_1\cup\cdots\cup C_{s+p-2}\cup C_{s+p}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big)\Big\|_2\\
&\le\Big\|\sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\\
&\qquad\qquad\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)-E_{C_1\cup\cdots\cup C_{s+p-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2\\
&\quad+\Big\|\sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)\\
&\qquad\qquad-E_{C_1\cup\cdots\cup C_{s+p-2}\cup C_{s+p}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2\\
&\le 2C'n^{\frac{m-1}{2}}.
\end{aligned}$$

After using the triangle inequality p times, we get

$$\begin{aligned}
&\Big\|\sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)\\
&\qquad\qquad-E_{C_1\cup\cdots\cup C_s}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2\le pC'n^{\frac{m-1}{2}}=C''n^{\frac{m-1}{2}}.
\end{aligned}$$

Now we claim that

$$E_{C_1\cup\cdots\cup C_s}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))=E_{\{l_1,\dots,l_s\}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)).$$

This can be established using axioms (3) and (4). Indeed, since

$$l_{s+p}\notin C_1\cup\cdots\cup C_{s+p-1}\supset\{l_1,\dots,l_{s+p-1}\},$$

by applying axiom (3) we see that

$$\begin{aligned}
&E_{\{l_1,\dots,l_{s+p-1}\}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\\
&\qquad= \pi_{l_{\phi(1)}}(x_1)\cdots E_{\{l_1,\dots,l_{s+p-1}\}}(\pi_{l_{s+p}}(x_{k'_p})\cdots\pi_{l_{s+p}}(x_{k''_p}))\cdots\pi_{l_{\phi(m)}}(x_m)\\
&\qquad= \pi_{l_{\phi(1)}}(x_1)\cdots E_{C_1\cup\cdots\cup C_{s+p-1}}(\pi_{l_{s+p}}(x_{k'_p})\cdots\pi_{l_{s+p}}(x_{k''_p}))\cdots\pi_{l_{\phi(m)}}(x_m)\\
&\qquad= E_{C_1\cup\cdots\cup C_{s+p-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)),
\end{aligned}$$

and then

$$\begin{aligned}
E_{C_1\cup\cdots\cup C_s}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)) &= E_{C_1\cup\cdots\cup C_s}(E_{C_1\cup\cdots\cup C_{s+p-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)))\\
&= E_{C_1\cup\cdots\cup C_s}(E_{l_1,\dots,l_{s+p-1}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)))\\
&= E_{(C_1\cup\cdots\cup C_s)\cap\{l_1,\dots,l_{s+p-1}\}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\\
&= E_{\{l_1,\dots,l_s\}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)),
\end{aligned}$$

which proves the claim. Now the claim, together with the last inequality, gives

$$\Big\| \sum_{l_1\in C_1,\dots,l_{s+p}\in C_{s+p}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)-E_{\{l_1,\dots,l_s\}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2 \le C''n^{\frac{m-1}{2}}.$$

Step 1 now implies

$$\Big\| \sum_{(l_1,\dots,l_{s+p})=\dot{0}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes\big(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)-E_{\{l_1,\dots,l_s\}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\big)\Big\|_2 \le CC''n^{\frac{m-1}{2}},$$

which proves the first statement in the lemma. For the second statement, we begin by noting that $E_{\{l_1,\dots,l_s\}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))$ only depends on $l_1,\dots,l_s$, and not on $l_{s+1},\dots,l_{s+p}$. Indeed, let $(l_1,\dots,l_s,l'_{s+1},\dots,l'_{s+p})=\dot{0}$ be another $(s+p)$-tuple with the same first s entries. Take a finite permutation σ such that $\sigma(l_i)=l_i,\ i\le s$, and $\sigma(l_{s+i})=l'_{s+i},\ i\le p$. Then α_σ is the identity on $A_{l_1,\dots,l_s}$; hence

$$\begin{aligned}
&E_{\{l_1,\dots,l_s\}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{s+i}}(x_{k'_i})\cdots\pi_{l_{s+i}}(x_{k''_i})\cdots\pi_{l_{\phi(m)}}(x_m))\\
&\qquad= (E_{\{l_1,\dots,l_s\}}\circ\alpha_\sigma)(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{s+i}}(x_{k'_i})\cdots\pi_{l_{s+i}}(x_{k''_i})\cdots\pi_{l_{\phi(m)}}(x_m))\\
&\qquad= E_{\{l_1,\dots,l_s\}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l'_{s+i}}(x_{k'_i})\cdots\pi_{l'_{s+i}}(x_{k''_i})\cdots\pi_{l_{\phi(m)}}(x_m)),
\end{aligned}$$

which proves the claim. Now the first statement of the lemma together with an easy counting argument shows that

$$\begin{aligned}
&\Big(n^{-\frac{m}{2}}\sum_{(l_1,\dots,l_{s+p})=\dot{0}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)\Big)\\
&\qquad= \Big(n^{-\frac{m}{2}}\sum_{(l_1,\dots,l_{s+p})=\dot{0}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes E_{\{l_1,\dots,l_s\}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\Big)\\
&\qquad= \Big(n^{-\frac{s}{2}}\sum_{(l_1,\dots,l_s)=\dot{0}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes E_{\{l_1,\dots,l_s\}}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\Big).
\end{aligned}$$

Finally, let's note that

$$E_{l_1,\dots,l_s} \circ \alpha_{l_1,\dots,l_{s+p}} = \alpha_{l_1,\dots,l_s} \circ E_{1,\dots,s},$$

which implies

$$\begin{aligned} E_{l_1,\dots,l_s}(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)) &= (E_{l_1,\dots,l_s} \circ \alpha_{l_1,\dots,l_{s+p}})(\pi_{\phi(1)}(x_1)\cdots\pi_{\phi(m)}(x_m)) \\ &= (\alpha_{l_1,\dots,l_s} \circ E_{1,\dots,s})(\pi_{\phi(1)}(x_1)\cdots\pi_{\phi(m)}(x_m)) \\ &= \alpha_{l_1,\dots,l_s}(F_\sigma(x_1,\dots,x_m)). \qquad \square \end{aligned}$$

Theorem 3.11. *Let (π_j, B, A, D) be a sequence of symmetric independent copies, $x_1,\dots,x_m \in A$, $\sigma \in P_{1,2}(m)$ having s singletons and p pairs and $\phi : \{1,\dots,m\} \to \{1,\dots,s+p\}$ which encodes σ. Then*

$$\begin{aligned} &x_\sigma(x_1,h_1,\dots,x_m,h_m) \\ &\quad = \Bigl(n^{-\frac{m}{2}} \sum_{(j_1,\dots,j_m)=\sigma} s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\otimes \pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)\Bigr) \\ &\quad = f_\sigma(h_1,\dots,h_m)\Bigl(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} s(e_{l_1}\otimes h_{k_1})\cdots s(e_{l_s}\otimes h_{k_s})\otimes \alpha_{\{l_1,\dots,l_s\}}(F_\sigma(x_1,\dots,x_m))\Bigr) \\ &\quad = f_\sigma(h_1,\dots,h_m) W_\sigma(x_1,h_1,\dots,x_m,h_m), \end{aligned}$$

where

$$\begin{aligned} F_\sigma(x_1,\dots,x_m) &= E_{\{1,\dots,s\}}(\pi_{\phi(1)}(x_1)\cdots\pi_{\phi(m)}(x_m)), \\ f_\sigma(h_1,\dots,h_m) &= q^{\mathrm{cr}(\sigma)} \prod_{\{k,l\}\in\sigma} \langle h_k, h_l\rangle \end{aligned}$$

and $\{k_1,\dots,k_s\}$ are the singletons of σ. The elements

$$W_\sigma(x_1,h_1,\dots,x_m,h_m) = \Bigl(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} s(e_{l_1}\otimes h_{k_1})\cdots s(e_{l_s}\otimes h_{k_s})\otimes \alpha_{\{l_1,\dots,l_s\}}(F_\sigma(x_1,\dots,x_m))\Bigr)$$

will be called reduced Wick words.

Proof. We will use the previous lemma. Let

$$\widehat{B} = B, \quad \hat{A} = \Gamma_q(H)\,\bar{\otimes}\, A, \quad \widehat{D} = \Gamma_q(\ell^2\otimes H)\,\bar{\otimes}\, D$$

and $\hat{\pi}_j : \hat{A} \to \Gamma_q(\ell^2\otimes H)\,\bar{\otimes}\, D$ be the $*$-homomorphisms given by

$$\hat{\pi}_j(s(h)\otimes x) = s(e_j\otimes h)\otimes \pi_j(x).$$

Then $(\hat{\pi}_j, B, \hat{A}, \widehat{D})$ represents a sequence of independent symmetric copies. Moreover, it is easy to see that $\hat{A}_I = \Gamma_q(\ell^2(I)\otimes H)\,\bar{\otimes}\, A_I$. Now according to the previous lemma we have

$$
\begin{aligned}
&\left(n^{-\frac{m}{2}} \sum_{(j_1,\dots,j_m)=\sigma} \varepsilon_{j_{k_1}} \cdots \varepsilon_{j_{k_s}} \otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\otimes \pi_{j_1}(x_1)\cdots \pi_{j_m}(x_m)\right)\\
&\quad= \left(n^{-\frac{m}{2}} \sum_{(l_1,\dots,l_{s+p})=\dot{0}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes s(e_{l_{\phi(1)}}\otimes h_1)\cdots s(e_{l_{\phi(m)}}\otimes h_m)\otimes \pi_{l_{\phi(1)}}(x_1)\cdots \pi_{l_{\phi(m)}}(x_m)\right)\\
&\quad= \left(n^{-\frac{m}{2}} \sum_{(l_1,\dots,l_{s+p})=\dot{0}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes \hat{\pi}_{l_{\phi(1)}}(s(h_1)\otimes x_1)\cdots \hat{\pi}_{l_{\phi(m)}}(s(h_m)\otimes x_m)\right)\\
&\quad= \left(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes \hat{\alpha}_{l_1,\dots,l_s}(\widehat{F}_\sigma(s(h_1)\otimes x_1,\dots,s(h_m)\otimes x_m))\right)\\
&\quad= \left(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes \hat{\alpha}_{l_1,\dots,l_s}\big(E_{\hat{A}_{1,\dots,s}}(\hat{\pi}_{\phi(1)}(s(h_1)\otimes x_1)\cdots \hat{\pi}_{\phi(m)}(s(h_m)\otimes x_m))\big)\right)\\
&\quad= \Bigg(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes \hat{\alpha}_{l_1,\dots,l_s}\big(E_{\Gamma_q(\ell_s^2\otimes H)\bar{\otimes} A_{1,\dots,s}}\big(s(e_{\phi(1)}\otimes h_1)\cdots s(e_{\phi(m)}\otimes h_m)\\
&\qquad\qquad \otimes \pi_{\phi(1)}(x_1)\cdots \pi_{\phi(m)}(x_m)\big)\big)\Bigg)\\
&\quad= f_\sigma(h_1,\dots,h_m)\Bigg(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes \hat{\alpha}_{l_1,\dots,l_s}\big(s(e_1\otimes h_{k_1})\cdots s(e_s\otimes h_{k_s})\\
&\qquad\qquad \otimes E_{1,\dots,s}(\pi_{\phi(1)}(x_1)\cdots \pi_{\phi(m)}(x_m))\big)\Bigg)\\
&\quad= f_\sigma(h_1,\dots,h_m)\Bigg(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes \big(s(e_{l_1}\otimes h_{k_1})\cdots s(e_{l_s}\otimes h_{k_s})\\
&\qquad\qquad \otimes \alpha_{l_1,\dots,l_s}(E_{1,\dots,s}(\pi_{\phi(1)}(x_1)\cdots \pi_{\phi(m)}(x_m)))\big)\Bigg)\\
&\quad= f_\sigma(h_1,\dots,h_m)\left(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} \varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes s(e_{l_1}\otimes h_{k_1})\cdots s(e_{l_s}\otimes h_{k_s})\otimes \alpha_{l_1,\dots,l_s}(F_\sigma(x_1,\dots,x_m))\right).
\end{aligned}
$$

To see why the fifth equality is true, note that

$$
\begin{aligned}
&s(e_{\phi(1)}\otimes h_1)\cdots s(e_{\phi(m)}\otimes h_m)\\
&\qquad= \sum_{\theta\in P_{1,2}(m)} f_\theta(e_{\phi(1)}\otimes h_1\otimes\cdots\otimes e_{\phi(m)}\otimes h_m) W((e_{\phi(1)}\otimes h_1\otimes\cdots\otimes e_{\phi(m)}\otimes h_m)_\theta),
\end{aligned}
$$

where the notation $(\cdot)_\theta$ means that the pair positions of θ have been removed. After the application of $E_{\Gamma_q(\ell_s^2\otimes H)}$, we see that the only surviving partition is $\theta=\sigma$ and

$$
\begin{aligned}
E_{\Gamma_q(\ell_s^2\otimes H)}(s(e_{\phi(1)}\otimes h_1)\cdots s(e_{\phi(m)}\otimes h_m)) &= f_\sigma(h_1,\dots,h_m) W(e_1\otimes h_{k_1}\cdots e_s\otimes h_{k_s})\\
&= f_\sigma(h_1,\dots,h_m) s(e_1\otimes h_{k_1})\cdots s(e_s\otimes h_{k_s}).
\end{aligned}
$$

Now, let's define

$$
\tilde{s}(x,h) = \left(n^{-\frac{1}{2}} \sum_{j=1}^{n} \varepsilon_j\otimes s(e_j\otimes h)\otimes \pi_j(x)\right) \in (L^\infty(X)\,\bar{\otimes}\,\Gamma_q(\ell^2\otimes H)\,\bar{\otimes}\, D)^\omega.
$$

We claim that the new Wick words $\tilde{x}_\sigma$ associated to the variables $\tilde{s}(x,h)$ have the same moments as x_σ and hence they generate an isomorphic von Neumann algebra. Indeed, fix $\sigma\in P_{1,2}(m)$. Note that for

$(l_1,\dots,l_s)=\dot{0}$, we have $\mu(\varepsilon_{l_1}\cdots\varepsilon_{l_s})=\mu(\varepsilon_{l_1})\cdots\mu(\varepsilon_{l_s})=\delta_{s=0}$, due to the fact that ε_j are mean-zero, independent random variables. Then

$$\begin{aligned}
&\tau_\omega(\tilde{x}_\sigma(x_1,h_1,\dots,x_m,h_m))\\
&=\tau_\omega\Bigg(\Bigg(n^{-\frac{m}{2}}\sum_{(l_1,\dots,l_{s+p})=\dot{0}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes s(e_{l_{\phi(1)}}\otimes h_1)\cdots s(e_{l_{\phi(m)}}\otimes h_m)\otimes\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)\Bigg)\Bigg)\\
&=\lim_n\Bigg(n^{-\frac{m}{2}}\sum_{(l_1,\dots,l_{s+p})=\dot{0}}\mu(\varepsilon_{l_1}\cdots\varepsilon_{l_s})\,\tau(s(e_{l_{\phi(1)}}\otimes h_1)\cdots s(e_{l_{\phi(m)}}\otimes h_m))\,\tau_D(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\Bigg)\\
&=\delta_{s=0}\lim_n\Bigg(n^{-\frac{m}{2}}\sum_{(l_1,\dots,l_{s+p})=\dot{0}}\tau(s(e_{l_{\phi(1)}}\otimes h_1)\cdots s(e_{l_{\phi(m)}}\otimes h_m))\,\tau_D(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\Bigg)\\
&=\delta_{\sigma\in P_2(m)}\lim_n\Bigg(n^{-\frac{m}{2}}\sum_{(l_1,\dots,l_{s+p})=\dot{0}}\tau(s(e_{l_{\phi(1)}}\otimes h_1)\cdots s(e_{l_{\phi(m)}}\otimes h_m))\,\tau_D(\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m))\Bigg)\\
&=\tau_\omega(x_\sigma(x_1,h_1,\dots,x_m,h_m)).
\end{aligned}$$

Define $\mathcal{M}\subset(\Gamma_q(\ell^2\otimes H)\,\bar{\otimes}\,D)^\omega$ to be the von Neumann algebra generated by all the Wick words x_σ. Also define $\widetilde{\mathcal{M}}\subset(L^\infty(X)\,\bar{\otimes}\,\Gamma_q(\ell^2\otimes H)\,\bar{\otimes}\,D)^\omega$ to be the von Neumann algebra generated by the elements $\tilde{x}_\sigma$. Using the claim, the convolution formula and Proposition 3.1 we see that the map

$$\mathcal{M}\to\widetilde{\mathcal{M}},\quad \sum x_\sigma\mapsto\sum\tilde{x}_\sigma,$$

is a $*$-isomorphism. Applying the inverse of this isomorphism to the equality

$$\begin{aligned}
&\Bigg(n^{-\frac{m}{2}}\sum_{(l_1,\dots,l_{s+p})=\dot{0}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes s(e_{l_{\phi(1)}}\otimes h_1)\cdots s(e_{l_{\phi(m)}}\otimes h_m)\otimes\pi_{l_{\phi(1)}}(x_1)\cdots\pi_{l_{\phi(m)}}(x_m)\Bigg)\\
&=f_\sigma(h_1,\dots,h_m)\Bigg(n^{-\frac{s}{2}}\sum_{(l_1,\dots,l_s)=\dot{0}}\varepsilon_{l_1}\cdots\varepsilon_{l_s}\otimes s(e_{l_1}\otimes h_{k_1})\cdots s(e_{l_s}\otimes h_{k_s})\otimes\alpha_{l_1,\dots,l_s}(F_\sigma(x_1,\dots,x_m))\Bigg),
\end{aligned}$$

we obtain the desired identity. □

Proposition 3.12. *Let $x_1,\dots,x_m\in A$ and $h_1,\dots,h_m\in H$. Then we have the moment formula*

$$\tau(s(x_1,h_1)\cdots s(x_m,h_m))=\delta_{m\in 2\mathbb{N}}\sum_{\sigma\in P_2(m)}q^{\mathrm{cr}(\sigma)}\prod_{\{l,r\}\in\sigma}\langle h_l,h_r\rangle\tau(\pi_{j_1^\sigma}(x_1)\cdots\pi_{j_m^\sigma}(x_m)),$$

as well as the B-valued moment formula

$$E_B(s(x_1,h_1)\cdots s(x_m,h_m))=\delta_{m\in 2\mathbb{N}}\sum_{\sigma\in P_2(m)}q^{\mathrm{cr}(\sigma)}\prod_{\{l,r\}\in\sigma}\langle h_l,h_r\rangle E_B(\pi_{j_1^\sigma}(x_1)\cdots\pi_{j_m^\sigma}(x_m)),$$

where for every $\sigma\in P_2(m)$, the $j_1^\sigma,\dots,j_m^\sigma$ are chosen such that $(j_1^\sigma,\dots,j_m^\sigma)=\sigma$.

Proof. It's a straightforward application of the reduction formula. □

Remark 3.13. Proposition 3.1 shows that $M = \Gamma_q^0(B, S \otimes H)$ could be introduced abstractly as the tracial von Neumann algebra (M, τ) generated by elements $s(x, h)$, $x \in BSB$, $h \in H$, which satisfy the above moment formula.

Proposition 3.14. *Let K be infinite-dimensional and $x_1, \dots, x_m \in BSB$, $h_1, \dots, h_m \in K$, $\sigma \in P_{1,2}(m)$. Then $x_\sigma(x_1, h_1, \dots, x_m, h_m) \in \Gamma_q^0(B, S \otimes K)$. For every Hilbert space H, all the Wick words $x_\sigma(x_1, h_1, \dots, x_m, h_m)$, $x_i \in BSB$, $h_i \in H$, are in $M = \Gamma_q(B, S \otimes H)$. In particular, M is the ultraweakly closed linear span of the* (*reduced*) *Wick words and $L^2(M)$ is the $\|\cdot\|_2$-closed span of the* (*reduced*) *Wick words.*

Proof. We need a basic fact about infinite-dimensional Hilbert spaces.

Fact. Let K be an infinite-dimensional Hilbert space and $\lambda_1, \dots, \lambda_p \in \mathbb{C}$. Then there exist norm-bounded sequences $\xi_n^k, \eta_n^k \in K$ for $1 \le k \le p$ such that $\xi_n^k \to 0$, $\eta_n^k \to 0$ weakly and $\langle \xi_n^k, \eta_n^k \rangle = \lambda_k$ for all $1 \le k \le p$, and moreover $\xi_n^k, \eta_n^k \perp \xi_n^j, \eta_n^j$ for $k \ne j$. Indeed, let (e_n) be an orthonormal infinite sequence in K. Define

$$\xi_n^1 = \lambda_1 e_n, \quad \eta_n^1 = e_n, \quad \xi_n^2 = \lambda_2 e_{n+1}, \quad \eta_n^2 = e_{n+1}, \quad \dots, \quad \xi_n^p = \lambda_p e_{n+p-1}, \quad \eta_n^p = e_{n+p-1}.$$

To prove the proposition we will use induction on s, the numbers of singletons in σ. For $s = 0$, $x_\sigma(x_1, h_1, \dots, x_m, h_m) \in B$ due to the Wick word reduction formula, so the statement is trivial. For a given σ with pairs $B_1, \dots, B_p$ and $B = \{l, r\}$ we use the fact above to find uniformly norm-bounded vectors $h_{l,B}(k), h_{r,B}(k) \in K$ which converge to 0 weakly and such that $\langle h_{l,B}(k), h_{r,B}(k) \rangle = \langle h_l, h_r \rangle$ for all pairs $B = \{l, r\}$, and such that the $h_{l/r,B}(k)$'s are orthogonal for different pairs B. Let us define $\tilde{h}_i(k) = h_i$ for any singleton $\{i\} \in \sigma$ and $\tilde{h}_i(k) = h_{l/r,B}(k)$ if $i \in B$ and $i = l$ or $i = r$. For every other Wick word $x'_{\sigma'}(y_1, f_1, \dots, y_{m'}, f_{m'})$, with $y_j \in BSB$, $f_j \in K$, we have

$$\begin{aligned}
&\lim_{k\to\infty} \tau(s(x_1, \tilde{h}_1(k)) \cdots s(x_m, \tilde{h}_m(k)) x'_{\sigma'}) \\
&\qquad = \lim_{k\to\infty} \sum_{\theta \in P_{1,2}(m)} \tau(x_\theta(x_1, \tilde{h}_1(k), \dots, x_m, \tilde{h}_m(k)) x'_{\sigma'}) \\
&\qquad = \tau(x_\sigma x'_{\sigma'}) + \lim_{k\to\infty} \sum_{\theta_p \supset \sigma_p,\ |\theta_s| < s} \tau(x_\theta(x_1, \tilde{h}_1(k), \dots, x_m, \tilde{h}_m(k)) x'_{\sigma'}).
\end{aligned}$$

Indeed, for every $\theta \in P_{1,2}(m)$ which does not contain all the pairs of σ, we use the convolution and the moment formulas to obtain

$$\begin{aligned}
\tau(x_\theta x'_{\sigma'}) &= \sum_{\nu \in P_2(m+m')} \tau(x_\nu(x_1, \tilde{h}_1(k), \dots, y_{m'}, f_{m'})) \\
&= \sum_{\nu \in P_2(m+m')} f_\nu(\tilde{h}_1(k), \dots, \tilde{h}_m(k), f_1, \dots, f_{m'})\, \tau(W_\nu(x_1, \tilde{h}_1(k), \dots, y_{m'}, f_{m'})),
\end{aligned}$$

where the sum is taken over all ν that preserve the pairs of θ and σ' and additionally pair all the singletons of θ and σ'. Now since θ does not contain all the pairs of σ, there must be a leg l of a pair $\{l, r\} = B \in \sigma$ which is connected by θ to something other than its other leg in σ. There are three possibilities:

(1) θ connects l to a leg l' of another pair $B' = \{l', r'\} \in \sigma$. Then $\langle \tilde{h}_l(k), \tilde{h}_{l'}(k) \rangle = 0$; hence for every ν in the sum above we have $f_\nu(h_1, \dots, f_{m'}) = 0$.

(2) θ connects l to a singleton $\{i\} \in \sigma$. Then, since $\tilde{h}_l(k) \to 0$ weakly, we have $\langle \tilde{h}_l(k), h_i \rangle \to 0$; hence for every ν we also have that $f_\nu(h_1, \dots, f_{m'}) \to 0$ as $k \to \infty$.

(3) $\{l\}$ is a singleton of θ. In this case, every $\nu \in P_{1,2}(m+m')$ which appears in the sum has to connect l to a singleton $j \in \{1, \dots, m'\}$. Thus, $\langle \tilde{h}_l(k), f_j \rangle \to 0$ and again $f_\nu(h_1, \dots, f_{m'}) \to 0$ as $k \to \infty$.

Summing up, we see that for every θ such that $\sigma_p \not\subseteq \theta_p$, we have $\tau(x_\theta(k) x'_{\sigma'}) \to 0$ as $k \to \infty$. Thus, when letting $k \to \infty$, only those θ's containing the pairs of σ make a nonzero contribution. Among them, there is exactly one which has s singletons, namely σ; all the others have more pairs and hence less than s singletons. We deduce that

$$x_\sigma = w - \lim_{k\to\infty} \Big(s(x_1, \tilde{h}_1(k)) \cdots s(x_m, \tilde{h}_m(k)) - \sum_{\theta_p \supset \sigma_p,\ |\theta_s| < s} x_\theta(x_1, \tilde{h}_1(k), \dots, x_m, \tilde{h}_m(k)) \Big).$$

Since by the induction hypothesis all the x_θ's, with $|\theta_s| < s$, are in $\Gamma_q^0(B, S \otimes K)$, this proves the statement. For the second statement, let H be any Hilbert space and K an infinite-dimensional Hilbert space containing H. Let $x_i \in BSB$, $h_i \in H$ and $\sigma \in P_{1,2}(m)$. Then, by the first part, $x_\sigma(x_1, h_1, \dots, x_m, h_m) \in \Gamma_q^0(B, S \otimes K)$. But $x_\sigma = (E_{\Gamma_q(\ell_n^2 \otimes H)} \otimes \mathrm{id})_n (x_\sigma)$; hence $x_\sigma \in \Gamma_q(B, S \otimes H)$. $\square$

Remark 3.15. The reader can now better appreciate why we needed the "closure operation" in the definition of $\Gamma_q(B, S \otimes H)$. Indeed, Definition 3.4 ensures that the Wick words belong to $M = \Gamma_q(B, S \otimes H)$ for every Hilbert space H, finite- or infinite-dimensional. Also, Proposition 3.14 shows that $M = \Gamma_q(B, S \otimes H)$ could have been defined as the ultraweakly closed span of the Wick words.

In the following we use the notation $L_k^2(M)$ for the $\|\cdot\|_2$-closed span of the Wick words of degree k and $W_k(M)$ for the linear span of the Wick words of degree k.

Theorem 3.16. *Let (π_j, B, A, D) be a sequence of symmetric independent copies, $1 \in S = S^* \subset A$, H be a Hilbert space and $M = \Gamma_q(B, S \otimes H)$. Set $\widetilde{H} = H \oplus H$. Take an infinite-dimensional Hilbert space $K \supset H$ and set $\widetilde{K} = K \oplus K$:*

(1) *For every angle θ, let o_θ be the canonical rotation on $\widetilde{K}$. Then*

$$\theta \mapsto \alpha_\theta = (\Gamma_q(\mathrm{id} \otimes o_\theta) \otimes \mathrm{id})_n \in \mathrm{Aut}((\Gamma_q(\ell^2 \otimes \widetilde{K}) \mathbin{\overline{\otimes}} D)^\omega)$$

defines by restriction a one-parameter group of automorphisms of $\widetilde{M} = \Gamma_q(B, S \otimes \widetilde{H})$. Moreover, for every Wick word $x_\sigma(x_1, \tilde{h}_1, \dots, x_m, \tilde{h}_m) \in \widetilde{M}$ we have

$$\alpha_\theta(x_\sigma(x_1, \tilde{h}_1, \dots, x_m, \tilde{h}_m)) = x_\sigma(x_1, o_\theta(\tilde{h}_1), \dots, x_m, o_\theta(\tilde{h}_m)).$$

(2) *For every Wick word $x_\sigma(x_1, h_1, \dots, x_m, h_m) \in M$, the following formula holds:*

$$(E_M \circ \alpha_\theta)(x_\sigma(x_1, h_1, \dots, x_m, h_m)) = (\cos(\theta))^s x_\sigma(x_1, h_1, \dots, x_m, h_m),$$

where $E_M : \widetilde{M} \to M$ is the conditional expectation and s is the number of singletons of σ.

(3) *For every* $\theta \in \left[0, \frac{\pi}{2}\right)$, *let* $t = -ln(\cos(\theta))$. *Then* $t \mapsto T_t = E_M \circ \alpha_\theta|_M$ *defines a one-parameter semigroup of normal, trace-preserving, u.c.p. maps on* M. *Moreover, for every Wick word* $x_\sigma \in M$ *we have* $T_t(x_\sigma) = e^{-ts} x_\sigma$, *where* s *is the number of singletons of* σ. *Hence, when viewed as a contraction on* $L^2(M)$, *we have* $T_t = \sum_{s \geq 0} e^{-ts} P_s$, *where* P_s *is the orthogonal projection of* $L^2(M)$ *on* $L^2_s(M)$ *and the series is* $\|\cdot\|_\infty$*-convergent for every* $t > 0$. *In particular, if* $L^2_s(M)$ *is finitely generated as a right* B*-module for every* s, *then* T_t *is compact over* B *for every* $t > 0$.

(4) *The generator* N *of* T_t *is a positive, self-adjoint, densely defined operator in* $L^2(M) = \bigoplus_{k=0}^{\infty} L^2_k(M)$, *acting by*

$$N(x_\sigma(x_1, h_1, \ldots, x_m, h_m)) = k x_\sigma(x_1, h_1, \ldots, x_m, h_m)$$

for every $x_\sigma(x_1, h_1, \ldots, x_m, h_m) \in L^2_k(M)$. *The spectrum of N is the set of nonnegative integers* $\mathbb{N}$, *all of which are eigenvalues. N is called the number operator.*

Proof. The formula $\alpha_\theta(x_\sigma(x_1, \tilde{h}_1, \ldots, x_m, \tilde{h}_m)) = x_\sigma(x_1, o_\theta(\tilde{h}_1), \ldots, x_m, o_\theta(\tilde{h}_m))$ for $x_i \in BSB$, $\tilde{h}_i \in \widetilde{H}$ is easily checked, due to entrywise functoriality, and it shows that α_θ restricts to a one-parameter group of automorphisms on $\widetilde{M} = \Gamma_q(B, S \otimes \widetilde{H})$. This proves (1). Then, using the reduction formula and the functoriality in each entry, we see that

$$\begin{aligned}
&(E_M \circ \alpha_\theta)(x_\sigma(x_1, h_1, \ldots, x_m, h_m)) \\
&= f_\sigma(h_1, \ldots, h_m)(E_M \circ \alpha_\theta)\Bigg(\Bigg(n^{-\frac{s}{2}} \sum_{(l_1,\ldots,l_s)=\dot{0}} s(e_{l_1} \otimes h_{k_1}) \cdots s(e_{l_s} \otimes h_{k_s}) \otimes \alpha_{l_1,\ldots,l_s}(F_\sigma(x_1, \ldots, x_m))\Bigg)\Bigg) \\
&= f_\sigma(h_1, \ldots, h_m)(E_M \circ \alpha_\theta)\Bigg(\Bigg(n^{-\frac{s}{2}} \sum_{(l_1,\ldots,l_s)=\dot{0}} W(e_{l_1} \otimes h_{k_1} \cdots e_{l_s} \otimes h_{k_s}) \otimes \alpha_{l_1,\ldots,l_s}(F_\sigma(x_1, \ldots, x_m))\Bigg)\Bigg) \\
&= f_\sigma(h_1, \ldots, h_m)\Bigg(n^{-\frac{s}{2}} \sum_{(l_1,\ldots,l_s)=\dot{0}} W(e_{l_1} \otimes P_H \alpha_\theta(h_{k_1}) \cdots e_{l_s} \otimes P_H \alpha_\theta(h_{k_s})) \otimes \alpha_{l_1,\ldots,l_s}(F_\sigma(x_1, \ldots, x_m))\Bigg) \\
&= (\cos(\theta))^s f_\sigma(h_1, \ldots, h_m)\Bigg(n^{-\frac{s}{2}} \sum_{(l_1,\ldots,l_s)=\dot{0}} s(e_{l_1} \otimes h_{k_1}) \cdots s(e_{l_s} \otimes h_{k_s}) \otimes \alpha_{l_1,\ldots,l_s}(F_\sigma(x_1, \ldots, x_m))\Bigg) \\
&= (\cos(\theta))^s x_\sigma(x_1, h_1, \ldots, x_m, h_m),
\end{aligned}$$

which establishes (2). Part (3) is straightforward using (2). To obtain (4), we calculate

$$\lim_{t \to 0} \frac{1}{t}(T_t(x_\sigma) - x_\sigma) = \lim_{t \to 0} \frac{e^{-st} - 1}{t} x_\sigma = -s x_\sigma$$

for any Wick word x_σ of degree s. The rest of the statements are straightforward. □

Remark 3.17. Due to (4), we have that for every $x \in M$, the function

$$\theta \mapsto \|\alpha_\theta(x) - x\|_2, \quad \theta \in \left[0, \tfrac{\pi}{2}\right),$$

is increasing.

Definition 3.18. We denote by $D_k(S) \subset L^2(D)$ the $\|\cdot\|_2$-closed linear span of the expressions

$$F_\sigma(x_1, \dots, x_m) = E_{1,\dots,k}(\pi_{\phi(1)}(x_1) \cdots \pi_{\phi(m)}(x_m))$$

for all $m \geq 1,\ x_1, \dots, x_m \in BSB,\ \sigma \in P_{1,2}(m)$ having k singletons and ϕ which encodes σ.

Lemma 3.19. *Let $y(j_1, \dots, j_k) \in L^p(D)$ be such that*

$$\sup_{j_1,\dots,j_k} \|y(j_1, \dots, j_k)\|_p < \infty$$

and $h_1, \dots, h_k \in H$. Then

$$\sup_n \left\| n^{-\frac{k}{2}} \sum_{(l_1,\dots,l_k)=\dot{0}} s_{l_1}(h_1) \cdots s_{l_k}(h_k) \otimes y(l_1, \dots, l_k) \right\|_p < \infty.$$

Proof. It suffices to consider

$$\left\| \sum_{l_1 \in C_1,\dots,l_k \in C_k} s_{l_1}(h_1) \cdots s_{l_k}(h_1) \otimes y(l_1, \dots, l_k) \right\|,$$

with $C_1 \cup \cdots \cup C_k = \{1, \dots, n\}$. Using the martingale decomposition from Lemma 3.10 we deduce

$$\left\| \sum_{l_1 \in C_1,\dots,l_k \in C_k} s_{l_1}(h_1) \cdots s_{l_k}(h_k) \otimes y(l_1, \dots, l_k) \right\|_p \leq c(p)\sqrt{n} \sup_{l \in C_k} \left\| \sum_{l_1,\dots,l_{k-1}} s_{l_1}(h_1) \cdots s_{l_k}(h_k) \otimes y(l_1, \dots, l_k) \right\|_p.$$

Iterating this procedure we get

$$\left\| \sum_{l_1 \in C_1,\dots,l_k \in C_k} s_{l_1}(h_1) \cdots s_{l_k}(h_k) \otimes y(l_1, \dots, l_k) \right\|_p \leq c(p)^k n^{\frac{k}{2}} \sup_{l_1,\dots,l_k} \|s_{l_1}(h_1) \cdots s_{l_k}(h_k)\|_p \|y(l_1, \dots, l_k)\|_p.$$

Since the products $s_{l_1}(h_1) \cdots s_{l_k}(h_k)$ are uniformly bounded in the p-norm, we obtain the assertion. □

Proposition 3.20. *Let (π_j, B, A, D) be a sequence of independent symmetric copies, let H be a finite-dimensional Hilbert space and $1 \in S = S^* \subset A$, and assume that $D_s(S)$ is finitely generated as a right B-module. Then $L^2_s(M)$ is finitely generated as a right B-module. In particular, when $D_s(S)$ is finitely generated over B for every s, the maps T_t are compact over B for every $t > 0$.*

Proof. Let N be the dimension of D_s as a right B-module, and let $\{\xi_1, \dots, \xi_N\}$ be a basis of D_s over B. Then, for every $\sigma \in P_{1,2}(m)$ having s singletons, and every $x_1, \dots, x_m \in BSB$, we can find coefficients $b_k(\sigma, x_1, \dots, x_m) \in B$ such that

$$F_\sigma(x_1, \dots, x_m) = \sum_{k=1}^{N} \xi_k b_k(\sigma, x_1, \dots, x_m).$$

For every $(l_1, \dots, l_s) = \dot{0}$ we have

$$\alpha_{l_1,\dots,l_s}(F_\sigma(x_1,\dots,x_m)) = \sum_{k=1}^{N} \alpha_{l_1,\dots,l_s}(\xi_k)\, b_k(\sigma, x_1,\dots,x_m).$$

Fix a finite basis $\mathcal{B}$ of H. Then, for every σ having s singletons, every $x_1,\dots,x_m \in BSB$ and every $h_1,\dots,h_m \in \mathcal{B}$ we have, due to the reduction formula,

$$\begin{aligned} x_\sigma(x_1,h_1,\dots,x_m,h_m) &= \Bigl(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} s_{l_1}(h_{i_1})\cdots s_{l_s}(h_{i_s}) \otimes \alpha_{l_1,\dots,l_s}(F_\sigma(x_1,\dots,x_m))\Bigr) \\ &= \sum_{k=1}^{N} \Bigl(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} s_{l_1}(h_{i_1})\cdots s_{l_s}(h_{i_s}) \otimes \alpha_{l_1,\dots,l_s}(\xi_k)\Bigr) b_k(\sigma, x_1,\dots,x_m). \end{aligned}$$

Thus $L^2_s(M)$ is spanned over B by at most $N|\mathcal{B}|^s = N(\dim(H))^s$ elements, namely

$$\Bigl(n^{-\frac{s}{2}} \sum_{(l_1,\dots,l_s)=\dot{0}} s_{l_1}(h_{i_1})\cdots s_{l_s}(h_{i_s}) \otimes \alpha_{l_1,\dots,l_s}(\xi_k)\Bigr),$$

with $h_i \in \mathcal{B}$ and $1 \le k \le N$. These elements belong to $L^2_s(M)$ by the previous lemma, and this finishes the proof. □

Remark 3.21. Since the dimension of $D_s(S)$ over B is finite, the basis elements $\xi_k \in D_s \subset L^2(D)$ could be chosen in fact to be bounded, i.e., $\xi_k \in D$, due to [Paschke 1973; 1974]. This implies that $L^2_s(M)$ admits a basis over B consisting of elements in M.

Corollary 3.22. *Assume moreover that the dimension N_s of $D_s(S)$ over B has polynomial growth*; *i.e., there exist constants $d, C > 0$ such that $N_s \le C d^s$ for all s. Then the dimension of $L^2_s(M)$ over B is less than $C(\dim(H)d)^s$ for all s*; *i.e., the dimension of $L^2_s(M)$ over B also has polynomial growth.*

The following argument is essentially due to Śniady [2004] and Królak [2006].

Proposition 3.23. *Let $M = \Gamma_q(B, S \otimes H)$. There exists $d = d(q)$ such that for* $\dim(H) \ge d$ *we have $\mathcal{Z}(M) \subset \mathcal{Z}(B)$. In particular*, *$M$ is a factor whenever B is.*

Proof. Let $\{e_i\}_{1\le i\le k}$ be an orthonormal set in H. We consider the operator $T : L^2(M) \to L^2(M)$ given by

$$T = \sum_{i=1}^{k} (L_{s(1,e_i)} - R_{s(1,e_i)})^2.$$

Here L_x and R_x, where $x \in M$, are the canonical left and right multiplication operators, respectively, on $L^2(M)$. We see that

$$T - 2k\,\mathrm{id} = \sum_{i=1}^{k} (L_{s(1,e_i)^2-1} - R_{s(1,e_i)^2-1}) - 2\sum_{i=1}^{k} L_{s(1,e_i)} R_{s(1,e_i)}.$$

Since $s(1,e_i)^2-1$ is a mean-zero element, we deduce from [Nou 2004] that

$$\Big\|\sum_{i=1}^{k} s(1,e_i)^2-1\Big\|_\infty \le c_q\sqrt{k}.$$

Let us set $V=\sum_{i=1}^k L_{s(1,e_i)}R_{s(1,e_i)}$ and denote by $\iota : L^2(M)\to(\mathcal{F}_q(\ell^2\otimes H)\otimes L^2(D))^\omega$ the natural embedding given by the definition. Then we see that

$$\iota(V\xi)=(V_n\iota(\xi)_n)_n,\quad \xi\in L^2(M),$$

where

$$V_n=\frac{1}{n}\sum_{\substack{1\le i\le k\\ 1\le j,j'\le n}} L_{s(1,e_i\otimes e_j)}R_{s(1,e_i\otimes e_{j'})}.$$

Now we can easily modify the argument from [Królak 2006] to show that

(1) $\big\|\sum_{k,j,j'} l^+(e_i\otimes e_j)R_{s(1,e_i\otimes e_{j'})}\big\|\le c_q\sqrt{kn^2}$;

(2) $\|\sum_{k,j,j'} r^+(e_i\otimes e_j)L_{s(1,e_i\otimes e_{j'})}\big\|\le c_q\sqrt{kn^2}$;

(3) $\big\|\sum_{k,j\neq j'} l^-(e_i\otimes e_j)R_{s(1,e_i\otimes e_{j'})}\big\|\le c_q\sqrt{kn^2}$;

(4) $\big\|\sum_{k,j} l^-(e_i\otimes e_j)r^+(e_i\otimes e_j)|_{\mathbb{C}^\perp}\big\|\le q+c_q\sqrt{kn}$.

Here l^+,l^-,r^+,r^- are the left and right creation operators on the q-Fock space coming from the decomposition $L_{s(h)}=l^+(h)+l^-(h),\ R_{s(h)}=r^+(h)+r^-(h)$. The main estimate is derived from

$$l^-(h)r^+(k)(\xi)=q^{|\xi|}\xi+l^-(h)(\xi)\otimes k.$$

The second part can then be estimated via the second item above. This yields

$$\|(T-2k\,\mathrm{id})((\mathrm{id}-E_B)(\xi))\|\le 2qk\|(\mathrm{id}-E_B)(\xi)\|+2c_q\sqrt{k}\|(\mathrm{id}-E_B)(\xi)\|.$$

Now take $z\in\mathcal{Z}(M)$ with $E_B(z)=0$. Thus $T(z)=0$ and also

$$0=\|T(z)\|=\|2kz-(T(z)-2kz)\|\ge 2k\|z\|-2qk\|z\|-C_q\sqrt{k}\|z\|=(2k(1-q)-C_q\sqrt{k})\|z\|.$$

Thus for $2k(1-q)-C_q\sqrt{k}>0$, i.e., $k>\sqrt{C_q/(2(1-q))}$, we have that $z=0$. This implies $z=E_B(z)$ for all $z\in\mathcal{Z}(M)$; hence $\mathcal{Z}(M)\subset B$ and also $\mathcal{Z}(M)\subset\mathcal{Z}(B)$. □

3A. *H-less generalized q-gaussians.* Finally, let us mention that there is an H-less version of the generalized q-gaussians, which can be described as follows: let (π_j,B,A,D) be a sequence of symmetric independent copies. For $1\in S=S^*\subset A$, define the von Neumann algebra $\Gamma_q(B,S)\subset(\Gamma_q(\ell^2)\bar{\otimes}D)^\omega$ as being generated by the elements $s_q(x)=\big(n^{-\frac{1}{2}}\sum_{j=1}^n s_q(e_j)\otimes\pi_j(x)\big)_n$ for $x\in BSB$. This is equivalent to taking H to be 1-dimensional in Definition 3.4 above; hence the H-less q-gaussians are a particular case of Definition 3.4. Surprisingly, the H generalized q-gaussians can also be obtained as a particular case of this construction. Indeed, let H be a (real) Hilbert space and (π_j,B,A,D) a sequence of symmetric independent copies. Let (X,μ) be a standard probability measure space and define a new sequence of symmetric independent copies $(\tilde{\pi}_j, \widetilde{B}, \widetilde{A}, \widetilde{D})$ by taking $\widetilde{B} = B$, $\widetilde{A} = A \bar{\otimes} L^\infty(X)$, $\widetilde{D} = D \bar{\otimes} \big(\bar{\bigotimes}_1^\infty L^\infty(X)\big)$ and $\tilde{\pi}_j : \widetilde{A} \to \widetilde{D}$ by

$$\tilde{\pi}_j(a \otimes f) = \pi_j(a) \otimes (1 \otimes 1 \otimes \cdots \otimes \underbrace{f}_{j\text{-th position}} \otimes \cdots \otimes 1 \otimes \cdots), \quad \text{where } a \in A,\ f \in L^\infty(X).$$

Using Rademacher variables, we see that there exists a dense subspace $H_0 \subset H$ and an isometric embedding $\iota : H_0 \to L^\infty(X) \subset L^2(X)$. Take $\widetilde{S} = S \otimes \iota(H_0) = \{a \otimes \iota(h) : a \in S, h \in H_0\} \subset \widetilde{A}$. The reader can check that

$$\Gamma_q(B, S \otimes H) = \Gamma_q(\widetilde{B}, \widetilde{S}).$$

4. Examples

We will discuss several types of examples of generalized q-gaussian von Neumann algebras. The underlying idea in all these cases is that whenever we have a finite von Neumann algebra on which the symmetric group acts, we can construct a sequence of symmetric copies. In particular, countable tensor or (amalgamated) free products von Neumann algebras or the pure q-gaussian von Neumann algebras $\Gamma_q(H)$, for an infinite-dimensional H, constitute obvious candidates, since the symmetric group acts naturally on them.

4A. ***Tensor products.*** Let B and C be finite von Neumann algebras. Define $A = B \mathbin{\bar{\otimes}} C$ and $D = B \mathbin{\bar{\otimes}} C^{\mathbb{N}} = B \mathbin{\bar{\otimes}} \big(\bar{\bigotimes}_{\mathbb{N}} C\big)$. Define $\pi_j : A \to D$ by the formula

$$\pi_j(b \otimes a) = b \otimes 1 \otimes 1 \otimes \cdots \otimes \underbrace{a}_{j\text{-th position}} \otimes \cdots \otimes 1 \otimes \cdots .$$

Then it's easy to check that (π_j, B, A, D) is a sequence of symmetric independent copies. It's likewise easy to see that

$$\Gamma_q(B, A \otimes H) = B \mathbin{\bar{\otimes}} \Gamma_q(L^2_{\text{sa}}(C) \otimes H).$$

For any finite subset $S \subset L^2_{\text{sa}}(C) \otimes H$, the space $D_k(S)$ has finite dimension over B.

4B. ***Free products with amalgamation.*** Let $B \subset A$ be an inclusion of finite tracial von Neumann algebras. Take $D = \mathop{\ast}_B A_j$, the amalgamated free product of a countable number of copies A_j, $j \in \mathbb{N}$, of A. Define $\pi_j : A \to D$ by the formula

$$\pi_j(a) = 1 * 1 * \cdots * \underbrace{a}_{j\text{-th position}} * \cdots * 1 * \cdots .$$

Then (π_j, B, A, D) represents a sequence of independent symmetric copies. To see why this is true it suffices to consider elements a_i such that $E_B(a_i) = 0$. Then we have to calculate

$$\tau_\sigma(a_1, \ldots, a_m) = \tau(\pi_{j_1}(a_1) \cdots \pi_{j_m}(a_m))$$

such that $(j_1, \ldots, j_m) = \sigma$. If σ has no crossings, we can inductively replace neighboring pairs by $E_B(\pi_{j_i}(a_i)\pi_{j_i}(a_{i+1})) = E_B(a_i a_{i+1})$ and finally find an element in B. For a noncrossing pair partition we can also join all the pairs, but then we find an expression of the form

$$\tau(b_1\pi_{j_{i_1}}(a_{i_1})b_2\pi_{j_{i_2}}(a_{i_2})\cdots\pi_{j_{i_k}}(a_{i_k}))=0.$$

Thus in the moment formula we only have to expand over noncrossing pair partitions. Now take $S=\{1,u,u^*\}$ for $u\in A$ a Haar unitary such that $E_B(u^n)=0$ for all $n\neq 0$. It's easy to see that $D_k(S)$ is the closed linear span of all the expressions $b_1\pi_1(u^{\varepsilon_1})b_2\cdots\pi_k(u^{\varepsilon_k})b_{k+1}$, with $b_i\in B$ and $\varepsilon_i\in\{0,1,*\}$. In particular, when $B=\mathbb{C}$ or is finite-dimensional, we have $\dim_B(D_k(S))\leq C2^{2k}$.

4C. *Group actions.*

4C1. *Second quantization.* Let $G\curvearrowright_\alpha C$ be a trace-preserving action of the discrete group G on the finite von Neumann algebra C. Also let $\nu: G\to\mathcal{O}(H_\mathbb{R})$ be an orthogonal representation of G on a real Hilbert space $H_\mathbb{R}$. Let (Ω,μ) be the gaussian construction associated to ν; see, e.g., [Peterson and Sinclair 2012]. We also denote the corresponding action $G\curvearrowright L^\infty(\Omega)$ by ν. Then define $B=C\rtimes_\alpha G$, $A=(C\mathbin{\bar\otimes} L^\infty(\Omega))\rtimes_\rho G$, $D=(C\mathbin{\bar\otimes} L^\infty(\Omega^\mathbb{N}))\rtimes_\rho G$, where the action ρ is given by $\rho_g(d\otimes f)=\alpha_g(d)\otimes\nu_g(f)$. Define the $*$-homomorphisms $\pi_j: A\to D$ by

$$\pi_j((d\otimes f)u_g)=(d\otimes 1\otimes 1\otimes\cdots\otimes\underbrace{f}_{j\text{-th position}}\otimes\cdots\otimes 1\otimes\cdots)u_g.$$

Then it is easy to see that the fixed points algebra is $C\rtimes_\alpha G$. Again the moments only depend on the inner product. Moreover, the gaussian functor yields a map $\mathrm{Br}: H\to L^2(\Omega)$. Then we find

$$M=\Gamma_q(C\rtimes G,\mathrm{Br}(H))=(C\mathbin{\bar\otimes}\Gamma_q(H))\rtimes G.$$

The spaces $D_k(S)$ are finite-dimensional modules over $B=C\rtimes G$ if $L^2_k(H)\rtimes G$ has a finite basis over G. For $k=1$ this means that H is finite-dimensional. In a forthcoming paper we will also analyze the case of profinite actions and/or representations, i.e., when H can be written as $H=\overline{\bigcup_i H_i}$ such that every H_i is a finite-dimensional G-invariant Hilbert subspace. However, discrete subgroups of $\mathcal{O}_n=\mathcal{O}(\mathbb{R}^n)$ provide a large class of nontrivial, nonamenable examples. The examples in [Junge, Longfield and Udrea 2014] are subalgebras of M.

4C2. *Symmetric group action.* Throughout this subsection Σ will denote the group of finite permutations on $\mathbb{N}$. Let us consider a countable discrete group G on which Σ acts by automorphisms. Examples for such a symmetric action are given by the natural action of Σ on the free group with countably many generators, or by the natural action of Σ on the direct product groups $\prod_{n\in\mathbb{N}}G$. More generally, let $R\subset\mathbb{F}_\infty$ be a set of generators which is invariant under the action of Σ, and assume that $\langle R\rangle\subset\mathbb{F}_\infty$ is a normal subgroup. Then $G=\mathbb{F}_\infty/\langle R\rangle$ is a group on which Σ acts. A perfect example is given by an amalgamated free product $\ast_H G_j$, where $G_j=G$. To make things more concrete, we may consider the discrete Heisenberg group $\mathcal{H}=\langle\mathbb{Z},\mathbb{Z}^\infty\rangle$ with generators $\{g_k\}_{k\geq 0}$ such that $\mathbb{Z}=\langle g_0\rangle$, $\mathbb{Z}^\infty=\langle g_k, k\geq 1\rangle$ and the following relations hold:

$$g_k^{-1}g_jg_k=g_0g_j,\quad k\neq j.$$

Then Σ acts on $\mathcal{H}$ by permuting the generators g_k for $k \geq 1$, and leaving g_0 fixed. Now we assume that such a G, with action $\Sigma \curvearrowright_\beta G$, acts trace-preservingly on a finite von Neumann algebra A and B is the fixed points algebra of this action α. Let $g \in G$ be an arbitrary element and $g_j = \beta_{(1j)}(g)$. We can then construct a sequence of symmetric copies (π_j, B, A, D) by defining $\pi_j : A \to A$, via $\pi_j(x) = \alpha_{g_j}(x)$. Working in the crossed product $(A \rtimes_\alpha G) \rtimes_\beta \Sigma$ it is easy to see that the π_j's are symmetric copies, and that B is the fixed points algebra for these symmetric copies. In fact we may and will always assume that G is generated by the g_j's and then $\pi_j(x) = x$ for all j is exactly the fixed points algebra of the action. In general $\pi_j(A) = A$ and hence we find an example of symmetric, but not necessarily independent, copies. In general independent copies are obtained from considering a suitable subalgebra $B \subset A_1 \subset A$. More generally for a subset $S \subset A$ we may however consider the algebras

$$A_j(S) = \{\pi_j(x) : x \in S,\ j \in A\}.$$

This is particularly interesting for a single self-adjoint x. Then independence depends on the mixing properties of the sequence $\pi_j(x)$, and has to be analyzed on a case by case basis. A more specific example can be constructed starting from a trace-preserving action α of $\mathbb{Z}$ on a finite von Neumann algebra N. Take $D = N \rtimes_\beta \mathcal{H}$ where the action β is obtained by lifting the action of $\mathbb{Z}$ via the group homomorphism $\pi : \mathcal{H} \to \mathbb{Z}$ given by $\pi(g_0) = 0$ and $\pi(g_j) = 1$ for $j \geq 1$. In other words,

$$\beta_g(x) = \alpha_{\pi(g)}(x), \quad g \in \mathcal{H},\ x \in N.$$

Let $\mathcal{H}_1$ be the group generated by g_0 and g_1 and take $B = N \rtimes \mathbb{Z} = N \bar{\otimes} L(\mathbb{Z})$ and $A = N \rtimes \mathcal{H}_1$. Define $\pi_j : A \to D$ by

$$\pi_j(xu_{g_1}) = \alpha_{\pi(g_j)}(x)u_{g_j}, \quad \pi_j(xu_{g_0}) = xu_{g_0}, \quad x \in N,\ j,k \in \mathbb{N}.$$

Then (π_j, B, A, D) is a sequence of symmetric independent copies. In full generality the dimensions of the spaces $D_k(S)$ or $L^2_k(M)$, where $M = \Gamma_q(B, A \otimes H)$, cannot be controlled. If we restrict ourselves to a small set of generators, e.g., $S = \{1, g_1, g_1^{-1}\}$, then we get a more well-behaved example. The space $D_k(S)$ is the closed linear span of the expressions of the form

$$\pi_{j_1}(u_{g_1}) \cdots \pi_{j_k}(u_{g_k}) u_{g_0}^{l(\sigma)} \alpha_{n(\sigma)}(x).$$

Thus $\dim_B(D_k(S)) \leq (2\dim(H))^{2k}$. For more general group actions and $S \subset L(G)$, we find coefficients in $B = L([G, G]) \bar{\otimes} N$ and finite dimension over B as long as we have finite generating sets. Note however, that $L([G, G])$ is in general not invariant under the action of Σ, and hence a more detailed case by case analysis is required. Again a particularly nice class of examples comes from one step nilpotent groups with commutators in the center, such as the Heisenberg groups.

4D. *Colored Brownian motion.*

4D1. *Top up q-gaussians.* Let H be a Hilbert space and $q_0 \in [-1, 1]$. Symmetric independent copies can be obtained from second quantization, or simply by defining $\pi(s_{q_0}(h)) = s_{q_0}(e_j \otimes h)$. This provides symmetric copies of $A = \Gamma_{q_0}(H)$ into $D = \Gamma_q(\ell_2(H))$. By looking at Wick words it is easy to see that the fixed points algebra is $\mathbb{C}$. Moreover, independence follows from the moment formula for q_0-gaussian random variables. Let $S = \{x_1, \dots, x_p\}$ be a finite, self-adjoint subset, where $x_i = s_{q_0}(h_1(i)) \cdots s_{q_0}(h_{l(i)}(i))$ for $1 \le i \le p$. Then we see that

$$\tau(s_q(k_1, x_1) \cdots s_q(k_m, x_m)) = \sum_{\sigma \in P_2(m)} q^{\mathrm{cr}(\sigma)} f_\sigma(k_1, \dots, k_m)\, \tau(\pi_{j_1^\sigma}(x_1) \cdots \pi_{j_m^\sigma}(x_m)),$$

where for every σ, we choose an m-tuple $(j_1^\sigma, \dots, j_m^\sigma)$ depending on σ such that $(j_1^\sigma, \dots, j_m^\sigma) = \sigma$. Now we may use the formula for q_0-gaussians and find for $L = \sum_{i=1}^p l(i)$ that

$$\tau(\pi_{j_1^\sigma}(x_1) \cdots \pi_{j_m^\sigma}(x_m)) = \sum_{\sigma' \in P_2(L),\ \sigma' \le \phi(\sigma)} q_0^{\mathrm{cr}(\sigma')} f_{\sigma'}(h_1(1), \dots, h_{l(1)}(1), \dots, h_1(m), \dots, h_{l(m)}(m)).$$

Here $\phi(\sigma)$ is the block partition which gives the same color to the union of two blocks in σ connected via pairs in σ'. This means

$$\tau(s_q(k_1, x_1) \cdots s_q(k_m, x_m)) = \sum_{\sigma \in P_2(m),\ \sigma' \le \phi(\sigma)} q^{\mathrm{cr}(\sigma')} q_0^{\mathrm{cr}(\sigma')} f_\sigma(k_1, \dots, k_m) f_{\sigma'}(h_1, \dots, h_L),$$

where σ' runs over the partitions of $\{1, \dots, L\}$ and $\{h_1, \dots, h_L\}$ is a relabeling of $\{h_j(i) : 1 \le i \le p, 1 \le j \le l(i)\}$. Note that $\Gamma_q(\mathbb{C}, \Gamma_{q_0}(H) \otimes K)$ contains both $\Gamma_q(K)$ and $\Gamma_{q_0}(H)$ if $s_{q_0}(H) \subset S$. Using a decomposition into minimal links, we deduce that the space $D_k(S)$ is the closed linear span of the elements

$$c(\sigma, x_1, \dots, x_r)\, \pi_{j_{i_1}}(x_{i_1}) \cdots \pi_{j_{i_r}}(x_{i_r}),$$

where $c(\sigma, x_1, \dots, x_r)$ is a scalar. This means for a finite set S of generators, the dimension of $D_k(S)$ over $B = \mathbb{C}$ is less than $(|S| \dim(K))^{2k}$. One could call these algebras "mixed" gaussian algebras, but the reader should not mistake them for the mixed q-gaussian algebras, introduced in [Junge and Zeng 2015], which we use in Section 6.

4D2. *Actions of Σ by conjugation.* Let us consider the finite permutations group $\Sigma_{\mathbb{Z}}$ acting on $\mathbb{Z}$ instead of $\mathbb{N} \setminus \{0\}$. For every subset $F \subset \mathbb{Z}$ we can identify Σ_F, the permutations group on F, with a subgroup of $\Sigma_{\mathbb{Z}}$ by viewing the elements of Σ_F as acting nontrivially only on F and acting as the identity on $\mathbb{Z} \setminus F$. For convenience, we use interval notation for the subsets of $\mathbb{Z}$. In particular we have $\Sigma = \Sigma_{[1,\infty)} \subset \Sigma_{\mathbb{Z}}$ in this way. Let Σ act on $\Sigma_{\mathbb{Z}}$ by conjugation. This gives rise to an action α of Σ on the von Neumann algebra $L(\Sigma_{\mathbb{Z}})$ (which is in fact isomorphic to the hyperfinite factor). We denote the canonical unitaries generating $L(\Sigma_{\mathbb{Z}})$ by u_σ, $\sigma \in \Sigma_{\mathbb{Z}}$. The fixed points algebra of this action is $B = L(\Sigma_{(-\infty,0]})$. Take $A = L(\Sigma_{(-\infty,1]}) = B \vee \{u_{(01)}\}''$, $D = L(\Sigma_{\mathbb{Z}})$ and define $\pi_j : A \to D$ by $\pi_j(a) = \alpha_{(j1)}(a)$ for $a \in A$ and $j \ge 2$, where $(j1)$ is the transposition interchanging j and 1, and $\pi_1 = \mathrm{id}$. Then (π_j, B, A, D) is a sequence of symmetric independent copies. Indeed, we recall that A is generated by transpositions $(k1)$, $k \le 0$, and that for $j \ge 2$ we have

$$(j1)(k1)(j1) = (kj).$$

This means $A_j = B \vee \{u_{(0j)}\}''$ and $A_{1,\dots,j} = L(\Sigma_{(-\infty,j]})$. In particular, we have a coset representation $\sigma = \sigma'(j1)$ with $\sigma' \in \Sigma_{(-\infty,0]}$. The algebras A_I are generated by Σ_I, $\Sigma_{(-\infty,0]}$ and one generator $(j1)$ for

$j \in I$. This easily implies independence. We take $S = \{1, u_{(01)}\} \subset A$ and define $M = \Gamma_q(B, S \otimes H)$. Fix $\sigma \in P_{12}(m)$ having k singletons and p pairs, and take $\phi : \{1, \ldots, m\} \to \{1, \ldots, k+p\}$ which encodes σ. This means $\phi(j_t) = t$, where $\{j_t\}$, $1 \le t \le k$, are the singletons of σ, and $\phi(j'_t) = \phi(j''_t) = k + t$, where $\{j'_t, j''_t\}$, $1 \le t \le p$, are the pairs of σ. Then $D_k(S)$ is the closed span of elements of the form

$$\begin{aligned} E_{1,\ldots,k}(u_{(\phi(1)0)}u_{\gamma_1}u_{(\phi(2)0)}u_{\gamma_2}\cdots u_{\gamma_m}u_{(\phi(m)0)}u_{\gamma_{m+1}}) \\ &= E_{1,\ldots,k}\big(u_{(\phi(1)0)}\mathrm{ad}(u_{\gamma_1})(u_{(\phi(2)0)})\cdots \mathrm{ad}(u_{\gamma_1\cdots\gamma_m})(u_{(\phi(m)0)})u_{\gamma_1\cdots\gamma_{m+1}}\big) \\ &= E_{1,\ldots,k}(u_{(\phi(1)\gamma_1(0))}u_{(\phi(2)\gamma_1\gamma_2(0))}\cdots u_{(\phi(m)(\gamma_1\cdots\gamma_m)(0))}u_{\gamma_1\cdots\gamma_{m+1}}) \\ &= E_{1,\ldots,k}(u_{(\phi(1)s_1)}u_{(\phi(2)s_2)}\cdots u_{(\phi(m)s_m)})u_{\gamma_1\cdots\gamma_{m+1}}, \end{aligned}$$

where $\gamma_1, \ldots, \gamma_{m+1} \in \Sigma_{(-\infty,0]}$ are arbitrary. Here $s_1 = \gamma_1(0)$, $s_2 = \gamma_1\gamma_2(0), \ldots, s_m = \gamma_1\gamma_2\cdots\gamma_m(0)$ in $(-\infty, 0]$ depend only the γ_i's. In full generality the modules $D_k(S)$ do not have finite dimensions over B. If we however replace B by $B_d = L(\Sigma_{[-d,0]}) \cong L(\mathbb{S}_{d+1})$, A by $A_d = L(\Sigma_{[-d,1]}) = B_d \vee \{u_{(01)}\}''$ and D by $D_d = L(\Sigma_{[-d,\infty)})$ for a fixed $d \in \mathbb{N} \setminus \{0\}$, then we obtain a new sequence of symmetric independent copies (π_j, B_d, A_d, D_d) and in this case we have at most $(d+1)^k$ different choices for the s_j's. After repeated conjugation with the unitaries on the pair positions, the above expression becomes

$$u_{(s'_{j_1}1)}\cdots u_{(s'_{j_k}k)}E_{1,\ldots,k}(u_{(s'_{j_{k+1}}k+1)}\cdots u_{(s'_{j_{k+p}}k+p)})$$

for some new indices $s'_i \in (-\infty, 0] \cap \mathbb{Z}$ which in general depend on the γ_i's and σ. Since for an inclusion of groups $H \subset G$ and $g \in G$ we have $E_{L(H)}(u_g) = \delta_{g \in H} u_g$, and the product $(s'_{j_{k+1}}k+1)\cdots(s'_{j_{k+p}}k+p)$ belongs to $\Sigma_{(-d,k]}$ only if it's equal to 1, we see that a spanning set of $D_k(S)$ over B_d is given by the elements

$$u_{(s'_{j_1}1)}u_{(s'_{j_2}2)}\cdots u_{(s'_{j_k}k)}$$

for all choices of $-d \le s'_i \le 0$, $1 \le i \le k$, which in particular implies that the dimension of $D_k(S)$ over B_d is at most $(d+1)^{2k}$. Note that B_d and A_d are finite-dimensional von Neumann algebras. Thus, for the von Neumann algebras $M(d) = \Gamma_q(B_d, S \otimes H)$, the spaces $D_k(S)$ have polynomial growth of their dimensions over B_d. This remains true for any finite subset $1 \in S = S^* \subset A_d$.

4E. ***Operator-valued gaussians.*** This example is motivated by Shlyahktenko's A-valued semicircular algebras and derived from the tensor product construction. Let $x_k \in N$ be self-adjoint operators and $X = \sum_k g_k x_k$. We consider $A_1 = L^\infty(\mathbb{R})$ and the independent symmetric copies over N given by

$$\pi_j(f) = f\bigg(\sum_k g_{k,j}x_k\bigg),$$

where $g_{k,j}$ are i.i.d. gaussians (we could also work with q-gaussians). The copies are independent over N. Let D be the von Neumann algebra generated by the $\pi_j(f)$'s and B be the tail algebra

$$B = \bigcap_{m\ge 0} \bigvee_{j\ge m} \pi_j(L^\infty(\mathbb{R})).$$

One can show that the copies π_j are independent symmetric in the sense of our Definition 3.2. Note that N is invariant under the shift from the tensor product construction and hence $B\subset N$. Thus $M=\Gamma_q(B,S\otimes H)$ is a legitimate example where $S=\sum_k g_k x_k$ is obtained by approximating X with bounded functions. Since $X\in\bigcap_{1\le p<\infty}L^p(\mathbb{R})$, one can actually directly work with one generator x. The dimension of $L^2_k(M)$ over B is in general hard to determine. The case of $N=M_m(\mathbb{C})$ and $X=\sum_{r,s}g_{rs}((e_{rs}+e_{sr})/2)$ has been considered by Avsec and Speicher.

Remark 4.1. The examples in Sections 4A, 4B, 4D1, 4D2 (for $d=0$) and 4E are all factors if B is a factor and $\dim(H)\ge d(q)$.

5. Weak amenability produces approximately invariant states

Let (π,B,A,D) be a sequence of symmetric independent copies, $1\in S=S^*\subset A$ and assume that $D_s(S)$ is finitely generated over B for all $s\ge 1$. Let $M=\Gamma_q(B,S\otimes H)$ for a finite-dimensional space H, $\mathcal{A}\subset M$ be a von Neumann subalgebra which is amenable relative to B inside M, and let $P=\mathcal{N}_M(\mathcal{A})''$. Define $\mathcal{M}=(\Gamma_q(\ell^2\otimes H)\mathbin{\bar\otimes}D)\vee M\subset(\Gamma_q(\ell^2\otimes H)\mathbin{\bar\otimes}D)^\omega$, where $\Gamma_q(\ell^2\otimes H)\mathbin{\bar\otimes}D$ is embedded as constant sequences. Let

$$\mathcal{H}\subset\big((L^2(\mathcal{M})\otimes_{\mathcal{A}}L^2(P))\otimes\mathcal{F}_q(\ell^2\otimes H)\big)^\omega$$

be the $\|\cdot\|$-closed span of the sequences

$$\Bigg(n^{-\frac{m}{2}}\sum_{(j_1,\dots,j_m)=\sigma}(\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)y\otimes_{\mathcal{A}}z)\otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\Bigg)$$

for all $m\ge 1$, $\sigma\in P_{1,2}(m)$, $x_i\in BSB$, $y\in M$, $z\in P$ and $h_1,\dots,h_m\in H$. Define two $*$-representations $\pi:M\to B(\mathcal{H})$, $\theta:P^{\mathrm{op}}\to B(\mathcal{H})$ by

$$\begin{aligned}\pi(x_{\sigma'})&\Bigg(n^{-\frac{m}{2}}\sum_{(j_1,\dots,j_m)=\sigma}(\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)y\otimes_{\mathcal{A}}z)\otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\Bigg)\\&=\Bigg(n^{-\frac{m+m'}{2}}\sum_{(i_k)=\sigma',(j_l)=\sigma}(\pi_{i_1}(y_1)\cdots\pi_{j_m}(x_m)y\otimes_{\mathcal{A}}z)\otimes s(e_{i_1}\otimes k_1)\cdots s(e_{j_m}\otimes h_m)\Bigg)\end{aligned}$$

and

$$\begin{aligned}\theta(w^{\mathrm{op}})&\Bigg(n^{-\frac{m}{2}}\sum_{(j_1,\dots,j_m)=\sigma}(\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)y\otimes_{\mathcal{A}}z)\otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\Bigg)\\&=\Bigg(n^{-\frac{m}{2}}\sum_{(j_1,\dots,j_m)=\sigma}(\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)y\otimes_{\mathcal{A}}zw)\otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\Bigg),\end{aligned}$$

where

$$x_{\sigma'}=\Bigg(n^{-\frac{m'}{2}}\sum_{(i_1,\dots,i_{m'})=\sigma'}\pi_{i_1}(y_1)\cdots\pi_{i_{m'}}(y_{m'})\otimes s(e_{i_1}\otimes k_1)\cdots s(e_{i_{m'}}\otimes k_{m'})\Bigg)\in M$$

is a Wick word in M and $w\in P$. Define $\mathcal{N}=\pi(M)\vee\theta(P^{\mathrm{op}})\subset B(\mathcal{H})$. Note that $\pi(M)$ and $\theta(P^{\mathrm{op}})$ commute.

Theorem 5.1. *There exists a sequence of normal states $\omega_n \in \mathcal{N}_*$ satisfying the following properties*:

(1) $\omega_n(\pi(x)) \to \tau(x),\ x \in M$.

(2) $\omega_n(\pi(a)\theta(\bar{a})) \to 1,\ a \in \mathcal{U}(\mathcal{A})$.

(3) $\|\omega_n \circ \mathrm{Ad}(\pi(u)\theta(\bar{u})) - \omega_n\| \to 0,\ u \in \mathcal{N}_M(\mathcal{A})$.

Proof. Throughout the proof m_n will be the completely contractive finite-rank multipliers on $\Gamma_q(\ell^2 \otimes H)$ given by multiplication with a positive finitely supported function f_n constructed in [Avsec 2012] and $\varphi_n := (m_n \otimes \mathrm{id}) : M \to M$ the corresponding cb map on M. Take

$$\mathcal{K} \subset (L^2(\mathcal{M}) \otimes_D L^2(\mathcal{M}))^\omega$$

to be the $\|\cdot\|$-closed span of the sequences

$$\left(n^{-\frac{m}{2}} \sum_{(j_1,\dots,j_m)=\sigma} (\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m) \otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)) \otimes_D y\right) = (x_\sigma^n \otimes_D y),$$

where $x_i \in BSB$ and $y \in M$. Note that $\mathcal{K}$ is naturally an M-M bimodule with the actions

$$x_{\sigma'} \cdot \left(n^{-\frac{m}{2}} \sum_{(j_1,\dots,j_m)=\sigma} (\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m) \otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)) \otimes_D y\right)\cdot z$$
$$= \left(n^{-\frac{m+m'}{2}} \sum_{(i_k)=\sigma',(j_l)=\sigma} (\pi_{i_1}(y_1)\cdots\pi_{j_m}(x_m) \otimes s(e_{i_1}\otimes k_1)\cdots s(e_{j_m}\otimes h_m)) \otimes_D yz\right),$$

where $x_{\sigma'} = x_{\sigma'}(y_1, k_1, \dots, y_{m'}, k_{m'}) \in M$ and $z \in M$. Define $\mathcal{S}_\mathcal{A} = \lambda(M) \vee \rho(\mathcal{A}^{\mathrm{op}}) \subset B(\mathcal{K})$, where λ and ρ are the representations of M and M^{op} canonically associated to the left and right actions on $\mathcal{K}$, respectively.

Step 1. There exists a normal, unital, completely positive map $\mathcal{E} : \mathcal{N} \to \mathcal{S}_\mathcal{A}$ such that

$$\mathcal{E}(\pi(x)\theta(y^{\mathrm{op}})) = \lambda(x)\rho(E_\mathcal{A}(y)^{\mathrm{op}}), \quad x \in M,\ y \in P.$$

Indeed, define an isometry $V : \mathcal{K} \to \mathcal{H}$ by

$$\left(n^{-\frac{m}{2}} \sum_{(j_1,\dots,j_m)=\sigma} (\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m) \otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)) \otimes_D y\right)$$
$$\mapsto \left(n^{-\frac{m}{2}} \sum_{(j_1,\dots,j_m)=\sigma} (\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)y \otimes_\mathcal{A} 1) \otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\right).$$

Then $\mathcal{E}$ can be defined by $\mathcal{E}(z) = V^* z V,\ z \in \mathcal{N}$.

Step 2. There exist normal functionals $\mu_n^\mathcal{A} : \mathcal{S}_\mathcal{A} \to \mathbb{C}$ such that

$$\mu_n^\mathcal{A}(\lambda(x)\rho(a^{\mathrm{op}})) = \tau(\varphi_n(x)a), \quad x \in M,\ a \in \mathcal{A}.$$

We need two lemmas. Recall the formulas for Wick words and reduced Wick words introduced in Theorem 3.11.

Lemma 5.2. *$L^2(M)\otimes_B L^2(M)$ embeds as an M-M bimodule into $\mathcal{K}$.*

Proof. The map

$$L^2(M)\otimes_B L^2(M)\to\mathcal{K},$$

$$\Bigl(n^{-\frac{m}{2}}\sum_{(j_1,\dots,j_m)=\sigma}\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)\otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\Bigr)\otimes_B y$$
$$\mapsto\Bigl(n^{-\frac{m}{2}}\sum_{(j_1,\dots,j_m)=\sigma}(\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)\otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m))\otimes_D y\Bigr),$$

or in other words $(x^n_\sigma)\otimes_B y\mapsto(x^n_\sigma\otimes_D y)$, is an M-M bimodular isometry. The bimodularity is obvious, so it remains to check that it preserves inner products, in other words that

$$\langle(x_n)\otimes_B y,(x'_n)\otimes_B y'\rangle=\langle(x_n\otimes_D y),(x'_n\otimes_D y')\rangle.$$

Let's denote by $E_D:\mathcal{M}\to D$ and by $E_{D\otimes 1}:\Gamma_q(\ell^2\otimes H)\mathbin{\bar{\otimes}}D\to D\otimes 1$ the canonical conditional expectations. Since $D=D\otimes 1\subset\mathcal{M}\subset(\Gamma_q(\ell^2\otimes H)\mathbin{\bar{\otimes}}D)^\omega$ is embedded as constant sequences, for every $(x_n)\in\mathcal{M}$ we have

$$E_D((x_n))=w-\lim_{n\to\omega}E_{D\otimes 1}(x_n).$$

We now claim that for any $(x_n)\in M\subset\mathcal{M}$ we have $E_B((x_n))=E_D((x_n))$. It suffices to prove this for $(x_n)=W_\sigma\in M$ a reduced Wick word. Let s be the number of singletons in σ. Let

$$W_\sigma=\Bigl(n^{-\frac{s}{2}}\sum_{(l_1,\dots,l_s)=\dot{0}}\alpha_{l_1,\dots,l_s}(F_\sigma(x_1,\dots,x_m))\otimes s(e_{l_1}\otimes h_{k_1})\cdots s(e_{l_s}\otimes h_{k_s})\Bigr).$$

We have two possibilities. If $s=0$, then $W_\sigma=F_\sigma(x_1,\dots,x_m)=E_B(\pi_{\phi(1)}(x_1)\cdots\pi_{\phi(m)}(x_m))\in B$; hence $E_D(W_\sigma)=W_\sigma=E_B(W_\sigma)$. If $s>0$, then $E_B(W_\sigma)=0$. On the other hand, according to our previous remark, we have

$$\begin{aligned}E_D(W_\sigma)&=w-\lim_n E_{D\otimes 1}\Bigl(n^{-\frac{s}{2}}\sum_{(l_1,\dots,l_s)=\dot{0}}\alpha_{l_1,\dots,l_s}(F_\sigma(x_1,\dots,x_m))\otimes s(e_{l_1}\otimes h_{k_1})\cdots s(e_{l_s}\otimes h_{k_s})\Bigr)\\&=w-\lim_n n^{-\frac{s}{2}}\sum_{(l_1,\dots,l_s)=\dot{0}}\tau(s(e_{l_1}\otimes h_{k_1})\cdots s(e_{l_s}\otimes h_{k_s}))\alpha_{l_1,\dots,l_s}(F_\sigma(x_1,\dots,x_m))\\&=w-\lim_n n^{-\frac{s}{2}}\sum_{(l_1,\dots,l_s)=\dot{0}}\tau(W(e_{l_1}\otimes h_{k_1}\cdots e_{l_s}\otimes h_{k_s}))\alpha_{l_1,\dots,l_s}(F_\sigma(x_1,\dots,x_m))=0.\end{aligned}$$

This proves our claim. Now, for $(x_n),(x'_n),y,y'\in M$ we have

$$\begin{aligned}\langle(x_n)\otimes_B y,(x'_n)\otimes_B y'\rangle&=\tau_M(E_B((x'^*_n x_n))yy'^*)=\tau_{\mathcal{M}}(E_D((x'^*_n x_n))yy'^*)\\&=\lim_n\tau_{\mathcal{M}}(E_{1\otimes D}(x'^*_n x_n)yy'^*)=\lim_n\langle x_n\otimes_D y,x'_n\otimes_D y'\rangle\\&=\langle(x_n\otimes_D y),(x'_n\otimes_D y')\rangle,\end{aligned}$$

which finishes the proof of the lemma. □

Lemma 5.3. *There exists an orthonormal basis Y_α of $L^2(M)$ over B such that for every n, $f_n(Y_\alpha)=0$ for all but finitely many α's, where we denote somewhat abusively $f_n(Y_\alpha)=f_n(s)$, $s=$ the degree of Y_α.*

Proof. Since D_s is finitely generated over B for all s, according to Proposition 3.20, for every $s\geq 0$ we can find a finite orthonormal basis (Y^s_β) of $L^2_s(M)$ over B. The union (Y_α) of all the Y^s_β's is a basis of $L^2(M)$ over B. For a fixed n, there exists $s=s(n)$ such that $f_n(\xi)=0$ for all $\xi\in H^{\otimes k}$, for $k>s(n)$. For any $t\geq 0$ and $Y_\alpha\in L^2_t(M)$ we have $f_n(Y_\alpha)=f_n(t)$ and also for every $Y_\alpha\in\bigoplus_{k>s(n)}L^2_k(M)$ we have $f_n(Y_\alpha)=0$, both due to the reduction formula. On the other hand, the set of those $Y_\alpha\in\bigoplus_{k=0}^{s(n)}L^2_k(M)$ is finite, which finishes the proof. □

Denote by ι the M-bimodular embedding in Lemma 5.2 and define

$$\mu_n^{\mathcal{A}}(T)=\sum_\alpha f_n(Y_\alpha)\langle T\iota(1\otimes_B 1),\iota(Y_\alpha^*\otimes_B Y_\alpha^*)\rangle,\quad T\in\mathcal{S}_{\mathcal{A}}.$$

Then $\mu_n^{\mathcal{A}}\in(\mathcal{S}_{\mathcal{A}})_*$ satisfies all the required properties.

Step 3. Set $\gamma_n=\mu_n^{\mathcal{A}}\circ\mathcal{E}\in\mathcal{N}_*$, and $\omega_n=\|\gamma_n\|^{-1}|\gamma_n|$. We will prove that the ω_n's satisfy all the required properties. First note that, by construction,

$$\gamma_n(\pi(x)\theta(y^{\mathrm{op}}))=\tau(\varphi_n(x)E_{\mathcal{A}}(y)),\quad x\in M,\ y\in P.$$

Toward proving the required properties of the ω_n's, we will first establish the following two claims:

Claim 1. $\limsup_n\|\mu_n^{\mathcal{A}}\|=1$.

Claim 2. $\lim_n\|\mu_n^{\mathcal{A}}\circ\mathrm{Ad}(\lambda(u)\rho(\bar{u}))-\mu_n^{\mathcal{A}}\|=0$, $u\in\mathcal{N}_M(\mathcal{A})$.

Proof of Claim 1. Fix a von Neumann subalgebra $Q\subset P$ which is amenable over B. Just as in Step 2 above one can construct normal functionals μ_n^Q on $\mathcal{S}_Q=\lambda(M)\vee\rho(Q^{\mathrm{op}})\subset B(\mathcal{K})$ satisfying $\mu_n^Q(\lambda(x)\rho(y^{\mathrm{op}}))=\tau(\varphi_n(x)y)$ for $x\in M$, $y\in Q$. We will show that $\limsup\|\mu_n^Q\|=1$, and this will help us establish both claims. Since μ_n^Q is normal, it suffices to estimate its norm on an ultraweakly dense C^*-subalgebra of $\mathcal{S}_Q$. Denote by S_Q the ultraweakly dense C^*-subalgebra of $\mathcal{S}_Q$ generated by $\lambda(x_\sigma)$ for $x_\sigma\in M$ the Wick words and $\rho(Q^{\mathrm{op}})$. First we note that there exist cb maps $\tilde{\varphi}_n:S_Q\to S_Q$ such that

$$\tilde{\varphi}_n(\lambda(x_\sigma)\rho(y^{\mathrm{op}}))=\lambda(\varphi_n(x_\sigma))\rho(y^{\mathrm{op}}),\quad x_\sigma\in M,\ y\in Q,$$

and $\|\tilde{\varphi}_n\|_{\mathrm{cb}}=\|\varphi_n\|_{\mathrm{cb}}$. To prove this take $\widetilde{\mathcal{K}}\subset L^2((\mathcal{M}\mathbin{\bar{\otimes}}\Gamma_q(\ell^2\otimes H))^\omega)$ to be the $\|\cdot\|_2$-closed linear span of the sequences

$$\left(n^{-\frac{m}{2}}\sum_{(j_1,\dots,j_m)=\sigma}\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)y\otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\right)=(x_\sigma^n(y\otimes 1))$$

for all $x_i\in BSB$, $h_i\in H$, $y\in M$. Now define an unitary operator $U:\mathcal{K}\to\widetilde{\mathcal{K}}$ by

$$(x_\sigma^n\otimes_D y)\mapsto(x_\sigma^n(y\otimes 1)).$$

We can then define

$$\tilde{\varphi}_n(z)=U^*(\mathrm{id}\otimes m_n)^\omega(UzU^*)U,\quad z\in S_Q.$$

Then the maps $\tilde{\varphi}_n$ satisfy all the required properties. The complete boundedness of the $\tilde{\varphi}_n$ is a delicate matter and it will be addressed in Section 5A below. On the other hand, since Q is amenable relative to B, we see that the M-Q bimodule $L^2(M)$ is weakly contained in $L^2(M) \otimes_B L^2(M)$, which in turn is contained in $\mathcal{K}$. This produces a $*$-homomorphism $\Theta : S_Q \to B(L^2(M))$ such that $\Theta(\lambda(x)\rho(y^{\mathrm{op}})) = \lambda_M(x)\rho_M(y^{\mathrm{op}})$, where λ_M, ρ_M are the natural actions of M on $L^2(M)$. But then

$$\mu_n^Q(z) = \langle \Theta(\tilde{\varphi}_n(z))1, 1\rangle, \quad z \in S_Q,$$

and this implies that $\limsup \|\mu_n^Q\| = 1$. Then by taking $Q = \mathcal{A}$ we get $\limsup \|\mu_n^{\mathcal{A}}\| = 1$, which finishes the proof of the first claim. □

Proof of Claim 2. Fix a unitary $u \in \mathcal{N}_M(\mathcal{A})$. The algebra $Q = \langle \mathcal{A}, u\rangle \subset P$ is amenable relative to B, so by the proof of Claim 1, $\limsup \|\mu_n^Q\| = 1$. Now since $\mu_n^Q(1) = \tau(\phi_n(1)) \to 1$ and $\mu_n^Q(\lambda(u)\rho(\bar{u})) = \tau(\phi_n(u)u^*) \to 1$, we see that $\|\mu_n^Q \circ \mathrm{Ad}(\lambda(u)\rho(\bar{u})) - \mu_n^Q\| \to 0$; hence by restricting to $S_{\mathcal{A}}$ we get $\|\mu_n^{\mathcal{A}} \circ \mathrm{Ad}(\lambda(u)\rho(\bar{u})) - \mu_n^{\mathcal{A}}\| \to 0$. Using the fact that $\mathrm{Ad}(\lambda(u)\rho(\bar{u})) \circ \mathcal{E} = \mathcal{E} \circ \mathrm{Ad}(\pi(u)\theta(\bar{u}))$ and the fact that $\gamma_n = \mu_n^{\mathcal{A}} \circ \mathcal{E}$, we see at once that $\|\gamma_n \circ \mathrm{Ad}(\pi(u)\theta(\bar{u})) - \gamma_n\| \to 0$. But since $\gamma_n(1) = \tau(\phi_n(1)) \to 1$ and $\limsup \|\gamma_n\| = 1$, we see that $\|\gamma_n - \omega_n\| \to 0$. This further implies $\|\omega_n \circ \mathrm{Ad}(\pi(u)\theta(\bar{u})) - \omega_n\| \to 0$, which establishes the third required property, and the other two follow easily. □

This completes the proof of Theorem 5.1. □

5A. ***cb-estimates for the multipliers.*** Here we will prove that some multipliers defined on certain C^*-algebras or von Neumann algebras are completely bounded. The first case is that of the maps $\tilde{\varphi}_n$ which were used in the proof of Theorem 5.1 above. In the second case we prove the complete boundedness of some normal multipliers on the von Neumann algebra $\mathcal{N}$ introduced above, which are needed to construct a concrete standard form for $\mathcal{N}$. We recall some notation.

Notation. $\mathcal{M} = (\Gamma_q(\ell^2 \otimes H) \bar{\otimes} D) \vee M \subset (\Gamma_q(\ell^2 \otimes H) \bar{\otimes} D)^\omega$, where we regard $\Gamma_q(\ell^2 \otimes H)$ and D as constant sequences. Let $K = L^2(\mathcal{M})$ or $K = L^2(\mathcal{M}) \otimes_A L^2(P)$. We introduce the subspace

$$\mathcal{L} \subset (K \otimes \mathcal{F}_q(\ell_2 \otimes H))^\omega$$

as the $\|\cdot\|$-closed linear span of the sequences

$$\Big(n^{-\frac{m}{2}} \sum_{(j_1,\dots,j_m)=\sigma} \pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)y \otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\Big) = (x_\sigma^n(y\otimes 1)) \in (K \bar{\otimes} \Gamma_q(\ell^2 \otimes H))^\omega,$$

for $m \geq 1,\ \sigma \in P_{12}(m),\ x_i \in BSB,\ h_i \in H,\ y \in M$. Let's define the extended Wick words $x_\sigma = x_\sigma(x_1, h_1, \dots, x_m, h_m, y^{\mathrm{op}})$ by

$$x_\sigma = \Big(n^{-\frac{m}{2}} \sum_{(j_1,\dots,j_m)=\sigma} \pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)y^{\mathrm{op}} \otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\Big),$$

where $m \geq 1,\ \sigma \in P_{1,2}(m),\ x_i \in BSB,\ h_i \in H,\ y \in P$, viewed as operators in $B(\mathcal{K})$, i.e., acting naturally on sequences in $\mathcal{L}$. The reader can check that:

(a) $\mathcal{L}$ is invariant to the natural action of the extended Wick words.

(b) $\mathcal{L} = \overline{\text{span}}\{\lambda(x_\sigma)\rho(y^{\text{op}})(1\otimes 1) : x_\sigma \in M,\ y \in M\}$ when $K = L^2(\mathcal{M})$ and

$$\mathcal{L} = \overline{\text{span}}\{\pi(x_\sigma)(1\otimes y)\theta(z^{\text{op}})((1\otimes_{\mathcal{A}} 1)\otimes 1) : x_\sigma \in M,\ y\in M,\ z \in P\}$$

when $K = L^2(\mathcal{M})\otimes_{\mathcal{A}} L^2(P)$.

(c) $\mathcal{L}$ is invariant to the natural action by orthogonal transformations of H given by

$$\mathcal{O}(H) \to \text{Aut}((\mathcal{M}\otimes\Gamma_q(\ell^2\otimes H))^\omega), \quad o \to \alpha_o = (\text{id}\otimes\Gamma_q(\text{id}\otimes o)).$$

Let $C(H)\subset B(\mathcal{L})$ be the C^*-algebra generated by the elements

$$\left(n^{-\frac{m}{2}} \sum_{(j_1,\ldots,j_m)=\sigma} \pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)y^{\text{op}}\otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\right) = (x_\sigma^n(y^{\text{op}}\otimes 1)),$$

where $x_i \in BSB,\ h_i\in H,\ y\in M,\ \sigma\in P(m)$. Also let $\widehat{C}(H)\subset (B(K)\otimes_{\min}\Gamma_q(\ell^2\otimes H))^\omega$ be the C^*-algebra generated by the elements

$$\left(n^{-\frac{m}{2}} \sum_{(j_1,\ldots,j_m)=\sigma} \pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m)y^{\text{op}}\otimes s(e_{j_1}\otimes h_1)\cdots s(e_{j_m}\otimes h_m)\right) = (x_\sigma^n(y^{\text{op}}\otimes 1)),$$

where $x_i\in BSB,\ y\in M,\ h_i \in H,\ \sigma\in P(m)$, the ultraproduct being the C^*-algebra ultraproduct.

Remark 5.4. Let m_α be the multipliers on $\Gamma_q(H)$ associated to the nonnegative finite support functions $f_\alpha : \mathbb{N}\to\mathbb{R}$:

(1) One may assume that for every k, $f_\alpha(k) = 1$ for α large enough and that $\limsup_\alpha \|m_\alpha\|_{\text{cb}} = 1$.

(2) $(\text{id}\otimes m_\alpha) : \widehat{C}(H)\to (B(K)\otimes_{\min}\Gamma_q(\ell^2\otimes H))^\omega$ are completely bounded, and are the restrictions of normal maps.

Lemma 5.5. *Let $\widehat{C}(H)$, $C(H)$ and m_α be defined as above*:

(1) *Let $\rho : (B(K)\otimes_{\min}\Gamma_q(\ell^2\otimes H))^\omega \to B((K\otimes\mathcal{F}_q(\ell^2\otimes H))^\omega)$ be the $*$-homomorphism defined by $\rho((T_n))(\xi_n) = (T_n\xi_n)$. Then $\rho(\widehat{C}(H))(\mathcal{L})\subset\mathcal{L}$, so $[\rho(\widehat{C}(H)), P_{\mathcal{L}}] = 0$.*

(2) *The map $\Phi : \widehat{C}(H)\to C(H)$ defined by $\Phi(T) = \rho(T)P_{\mathcal{L}}$ is a surjective $*$-homomorphism.*

(3) *If $\sigma\notin P_{1,2}(m)$, then $\Phi(x_\sigma^n(y^{\text{op}}\otimes 1)) = 0$. In particular, $C(H) = \Phi(\widehat{C}(H))$ is spanned by the elements $\Phi((x_\sigma^n(y^{\text{op}}\otimes 1)))$ for $m\geq 1,\ \sigma\in P_{1,2}(m)$.*

(4) *If $(x_n) = (x_n')\in M$, then $\Phi((x_n(y^{\text{op}}\otimes 1))) = \Phi((x_n'(y^{\text{op}}\otimes 1)))$. In particular, $C(H)$ is spanned by the elements $\Phi((W_\sigma(y^{\text{op}}\otimes 1)))$, where $W_\sigma\in M,\ \sigma\in P_{1,2}(m)$ are the reduced Wick words.*

Proof. Take $(x_\sigma^n(y^{\text{op}}\otimes 1))\in\widehat{C}(H),\ (x_{\sigma'}^n(z\otimes 1))\in\mathcal{L}$. Due to the convolution rule we have

$$\Phi((x_\sigma^n(y^{\text{op}}\otimes 1)))(x_{\sigma'}^n(z\otimes 1)) = (x_\sigma^n x_{\sigma'}^n(zy\otimes 1)) = \sum_{\gamma\in P(m+m')} (x_\gamma^n(zy\otimes 1)),$$

where the summation is taken over all those γ's which preserve the connections of both σ and σ'; i.e., if some indices are connected by σ or σ', they will remain connected in γ. Now for all $\gamma\notin P_{1,2}(m+m')$, the corresponding term vanishes, because $\|x^n_\gamma(zy\otimes 1)\|_2 \le \|zy\|_\infty \|x^n_\gamma\|_2 \to 0$. Thus

$$\Phi((x^n_\sigma(y^{\mathrm{op}}\otimes 1)))(x^n_{\sigma'}(z\otimes 1)) = \sum_{\gamma\in P_{1,2}(m+m')} (x^n_\gamma(zy\otimes 1)) \in \mathcal{L},$$

which proves the first statement. Also, if $\sigma \notin P_{1,2}(m)$ to begin with, every γ in the sum will also not be in $P_{1,2}(m+m')$; hence the whole sum vanishes, which proves the third statement. The second statement is trivial. If $(x_n), (x'_n) \in M$ such that $\lim \|x_n - x'_n\|_2 = 0$, then for every $(y^n_\sigma(z\otimes 1)) \in \mathcal{K}$, we have $\|x_n y^n_\sigma(zy\otimes 1) - x'_n y^n_\sigma(zy\otimes 1)\|_2 \le \|y^n_\sigma\|_\infty \|zy\|_\infty \|x_n - x'_n\|_2 \to 0$, i.e., $\Phi((x_n(y^{\mathrm{op}}\otimes 1))) = \Phi((x'_n(y^{\mathrm{op}}\otimes 1)))$. The last statement then follows from the reduction formula. □

Our goal is to prove that under certain conditions the maps $(\mathrm{id}\otimes m_\alpha)$ descend to a multiplier on the quotient algebra, namely $C(H)$. This is done via a careful analysis of Φ_*.

Lemma 5.6. *There exists a complete contraction*

$$\psi : \overline{(K\otimes L^2(\Gamma_q(\ell^2\otimes H)))}^r \otimes_h (K\otimes L^2(\Gamma_q(\ell^2\otimes H)))^c \to \overline{L^1(B(K)\otimes\Gamma_q(\ell^2\otimes H))}$$

such that

$$\psi((h\otimes a)\otimes(k\otimes b)) = (h\otimes\bar{k})\otimes ab^*$$

and $\mathrm{tr}((S\otimes T)(\psi((h\otimes a)\otimes(k\otimes b))^*) = ((S\otimes T)(k\otimes b), h\otimes a)$. *Here* $(k\otimes h)$ *is the rank-*1 *operator with entries* $(k_i h_j)$ *in a given basis and* $\otimes_h$ *denotes the Haagerup tensor product of operator spaces.*

Proof. We recall that for a semifinite von Neumann algebra M the space $M = \overline{L^1(M,\mathrm{tr})}^*$ is the antilinear dual with respect to the trace $\langle T,\rho\rangle = \mathrm{tr}(T\rho^*)$. Moreover, for $M = B(H)$ one usually considers linear duality with respect to the transposed ρ^t of a density ρ:

$$\langle\langle T,\rho\rangle\rangle_{B(H),S_1(H)} = \mathrm{tr}(T\rho^t) = \mathrm{tr}(T\dot{\rho}^*) = \langle T,\bar{\rho}\rangle_{B(H),\overline{S_1(H)}}.$$

Using the description of $S_1(H) = H^r \otimes_h H^c$ as a Haagerup tensor product, we find a natural map $\omega : H^r\otimes_h H^c \to B(H)^*$ given by

$$\omega(h\otimes k)(T) = \mathrm{tr}\Big(T\Big(\sum_{ij} h_i k_j e_{ij}\Big)^t\Big) = \mathrm{tr}\Big(T\Big(\sum_{ij} h_i k_j e_{ji}\Big)\Big) = \sum_{ij} T_{ij} h_i k_j = (T(k),\bar{h}).$$

Let M be a semifinite von Neumann algebra and (ξ_j) be an orthonormal basis. Then we may define the antilinear map $v(a) = \sum_j \langle \xi_j, a\rangle \xi_j$ and observe that

$$(b,\overline{v(a)}) = \sum_j \langle b,\xi_j\rangle\langle v(a),\xi_j\rangle = \sum_j \langle b,\xi_j\rangle\langle\xi_j,a\rangle = \tau(ba^*).$$

Therefore $\bar{m} = \omega(v\otimes\mathrm{id}) : \overline{L^2_r(M)}\otimes_h L^2_c(M) \to B(L^2(M))^*$ satisfies

$$m(a\otimes b)(T) = (T(b),\overline{v(a)}) = \tau(Tba^*) = \tau(T(ab^*)^*)) = \langle T,(ab^*)\rangle$$

for all $T\in M$. This shows that $m(a\otimes b) = ab^*$ is a complete contraction from $\overline{L^2_r}(M)\otimes_h L^2_c(M) \to \overline{L^1(M)}$. Now we repeat the argument for $H = K\otimes L^2(M)$ and $V(h\otimes b) = \bar{k}\otimes v(b)$. Then we obtain a complete contraction $\psi = \omega(V \otimes \mathrm{id}) : \overline{(K \otimes L^2(M))}^r \otimes_h K \otimes L^2(M) \to B(K \otimes L^2(M))^*$ such that for $S \in B(K)$ and $T \in M$

$$\psi((h \otimes a) \otimes (k \otimes b))(S \otimes T)(S(k), \bar{\bar{h}})\tau(Tba^*) = (S(k), h)(Tb, a) = ((S \otimes T)(k \otimes b), h \otimes a)$$
$$= (tr \otimes \tau)(S \otimes T)(k \otimes \bar{h}).$$

Here $(\alpha \otimes \beta) = \sum_{ij} e_{ij}\alpha_i\beta_j$ is the density of the corresponding rank-1 operator. Therefore the map $\psi((k \otimes a) \otimes (h \otimes b)) = (h \otimes \bar{k}) \otimes ab^*$ does the job. □

Corollary 5.7. *Let $\mathcal{L} \subset (K \otimes L^2(\Gamma_q(\ell^2 \otimes H)))^\omega$ be defined as above. Then there exists a completely contractive map*

$$\Psi : \bar{\mathcal{L}}^r \otimes_h \mathcal{L}^c \to \overline{(L_1(B(K) \otimes \Gamma_q(\ell^2 \otimes H)))}^\omega$$

and a complete contraction $q : \overline{(L_1(B(K) \otimes \Gamma_q(\ell^2 \otimes H)))}^\omega \to [(B(K) \bar{\otimes} \Gamma_q(\ell^2 \otimes H))^\omega]^*$ such that

$$(q \circ \Psi)(k \otimes h)(T) = \langle T(k), h\rangle.$$

In particular $\Psi^*|_{\widehat{C}(H)} = \Phi$.

Proof. For $\xi, \eta \in \mathcal{L}$ given by $\xi = (\xi_n)_n,\ \eta = (\eta_n)_n$ we may define

$$\Psi(\xi \otimes \eta) = (\psi(\xi_n \otimes \eta_n))_n,$$

where ψ is the map from Lemma 5.6. Now Ψ obviously extends by linearity, thanks to the definition of the Haagerup tensor product and the well-known fact that $M_m((X_n)^\omega) = (M_m(X_n))^\omega$; see [Pisier 2003]. The map q is given by the limit

$$q((\bar{\zeta}_n)_n)(T_n)_n = \lim_{n\to\omega} (\mathrm{tr} \otimes \tau)(T_n \zeta_n^*).$$

Now the assertion follows from Lemma 5.6 and the fact that the duality pairing is given by the limit along the ultraproduct. □

Remark 5.8. Let H be an infinite Hilbert space and $H \subset H'$. Thanks to the definition of the C^*-algebra $\widehat{C}(H)$ as a subalgebra of the ultraproduct, we clearly have an isometric inclusion $\widehat{C}(H) \subset \widehat{C}(H')$. The C^*-algebra $C(H) \subset B(\mathcal{L}(H))$ depends on our minimalistic definition of $\mathcal{L}(H)$. Certainly, $\mathcal{L}(H) \subset \mathcal{L}(H')$ and hence the tautological map $\iota(x_\sigma) = x_\sigma,\ \iota(y^{\mathrm{op}}) = y$ produces a larger norm on $\mathcal{L}(H')$ than on $\mathcal{L}(H)$. Let us consider a noncommutative polynomial p in a finite number of x_σ's and y^{op}'s, and we may assume that the x_σ only contain vectors from a finite-dimensional subspace $H_0 \subset H$. Then we can find norm attaining vectors $\xi, \eta \in \mathcal{L}(H')$ for p. Then we write $H' = H_0 \oplus H_0^\perp$ and may also assume that the ξ and η are linear combination of elements in $\mathcal{L}(H_0)$ and $\mathcal{L}(H_1)$, where $H_1 \subset H_0^\perp$ is a finite-dimensional subspace. Using the moment formula, we see that the inner product remains unchanged after applying an orthogonal transformation o which sends H_1 to a finite-dimensional subspace of H orthogonal to H_0 and leaves H_0 invariant. This implies

$$\|p\|_{C(H')} = \sup_{\|\xi\|\le 1, \|\eta\|\le 1} |\langle \xi, p\eta\rangle| = \sup_{\|\xi\|\le 1, \|\eta\|\le 1} |\langle \alpha_o(\xi), p\alpha_o(\eta)\rangle| \le \|p\|_{C(H)}.$$

Let us denote by $q_H = \Phi|_{\widehat{C}(H)} : \widehat{C}(H) \to C(H)$ the quotient map. Then we obtain a commutative diagram

$$\begin{array}{ccc} \widehat{C}(H) & \xrightarrow{q_H} & C(H) \\ \downarrow & & \downarrow \\ \widehat{C}(H') & \xrightarrow{q_{H'}} & C(H') \end{array}$$

where the left-hand downward arrow is the natural ultraproduct inclusion and the right-hand downward arrow is the tautological inclusion (which is well-defined and injective). This allows us to identify elements in the kernel of q_H by considering $q_{H'}$.

We recall that thanks to Avsec's result, the orthogonal projection $P_k : \Gamma_q(H) \to \Gamma_q(H)$ onto Wick words of length k is a normal completely bounded map. We use the same notation

$$P_k : L^1(\Gamma_q(H)) \to L^1(\Gamma_q(H)),$$
$$\mathrm{id} \otimes P_k : \overline{L^1(B(K) \otimes \Gamma_q(H))} \to \overline{L^1(B(K) \otimes \Gamma_q(H))}.$$

Let us note that one can take $P_k^\omega : \prod \overline{L^1(B(K) \otimes \Gamma_q(H))}$, the extension to the ultraproduct of L^1 spaces, which satisfies

$$\langle (\mathrm{id} \otimes P_k)((T_n)), (\xi_n) \rangle = \langle (T_n), (\mathrm{id} \otimes P_k)((\xi_n)) \rangle$$

with respect to the antilinear bracket given by the ultraproduct trace; see also [Raynaud 2002].

Lemma 5.9. *The kernel of* $\Phi \circ P_k^\omega$ *contains the kernel of* q_H.

Proof. The map $\Phi \circ P_k^\omega$ is normal. According to Remark 5.8 it therefore suffices to show that for $\xi, \eta \in \mathcal{L}$ we have

$$(\mathrm{id} \otimes P_k^\omega)(\Psi(\xi \otimes \eta)) \in \mathrm{Im}(\psi_{H'})$$

for some potentially larger Hilbert space H'. Let us now consider Wick words $(x_\sigma^n)_n$, $\tilde{x}_\sigma^n$, and y^{op}, $\tilde{y}^{\mathrm{op}}$. We have to consider

$$\Psi((\tilde{x}_\sigma \tilde{y}^{\mathrm{op}})_n \otimes (x_\sigma y^{\mathrm{op}})_n) = (\psi(\tilde{x}_\sigma^n y \tilde{y}^{\mathrm{op}} \otimes x_\sigma^n y^{\mathrm{op}}))_n.$$

For fixed $n \in \mathbb{N}$ we see that

$$\begin{aligned} \Psi((\tilde{x}_{\tilde{\sigma}}^n y) \otimes (x_\sigma^n y^{\mathrm{op}})) &= n^{-\frac{m+\tilde{m}}{2}} \sum_{(\tilde{j}_k)=\tilde{\sigma},\ (j_k)=\sigma} (\overline{\vec{\pi}(a) y^{\mathrm{op}}} \otimes \vec{\pi}_{\tilde{j}_1}(\tilde{a}) \tilde{y}^{\mathrm{op}}) \otimes \vec{s}_{\tilde{j}}(\tilde{h}) \vec{s}_j^{\,*} \\ &= \sum_{\sigma' \in P(m+m')} \Psi^{\sigma'}(\tilde{x}_{\tilde{\sigma}} \tilde{y}^{\mathrm{op}} \otimes x_\sigma y^{\mathrm{op}}), \end{aligned}$$

where

$$\Psi^{\sigma'}(\tilde{x}_{\tilde{\sigma}} \tilde{y}^{\mathrm{op}} \otimes x_\sigma y^{\mathrm{op}}) = \sum_{(\tilde{j}_1,\dots,\tilde{j}_{\tilde{m}}, j_m,\dots,j_1)=\sigma'} (\overline{\vec{\pi}(a) y^{\mathrm{op}}} \otimes \vec{\pi}_{\tilde{j}_1}(\tilde{a}) \tilde{y}^{\mathrm{op}}) \otimes \vec{s}_{\tilde{j}}(\tilde{h}) \vec{s}_j^{\,*}.$$

Note also that σ' has to be obtained by joining singletons from $\tilde{\sigma}$ and σ. In this context we observe again that is enough to consider $\sigma' \in P_{1,2}(\tilde{m}+m)$. In the following example we see that

$$\Big\| \sum_{j_1} (\overline{\pi_{j_1}(a_1)\pi_{j_1}(a_2)y^{\mathrm{op}}} \otimes \pi_{j_1}(\tilde{a}_1)\tilde{y}^{\mathrm{op}}) \otimes s_{j_1}^2 s_{j_1} \Big\|_1$$

$$\leq \Big\| \sum_{j_1} \pi_{j_1}(a_1)\pi_{j_1}(a_2)y^{\mathrm{op}} \otimes s_{j_1}^2 \otimes e_{1,j_1} \Big\| \Big\| \sum_{j_1} \pi_{j_1}(\tilde{a}_1)\tilde{y}^{\mathrm{op}} \otimes s_{j_1} e_{j_1,1} \Big\|$$

$$\leq c_q(a,\tilde{a})n \ll n^{\frac{3}{2}}$$

is much smaller than $n^{\frac{3}{2}}$ and hence vanishes in the limit. For more complicated configurations, we may assume that σ and $\tilde{\sigma}$ are pair/singleton partitions, and that new links in $\sigma' \in P_{1,2}(m+\tilde{m})$ are obtained from joining pairs or singletons in σ with pairs in $\tilde{\sigma}$ (or the other way around). All the joint pairings can be estimated using the definition of the Haagerup tensor product as above which yields the bound

$$\Big\| \sum_{\substack{(\tilde{j}_k)=\tilde{\sigma},\ (j_k)=\sigma \\ (\tilde{j}_k, j_k)=\sigma'}} (\overline{\vec{\pi}_j(a)y^{\mathrm{op}}} \otimes \vec{\pi}_j(\tilde{a})\tilde{y}^{\mathrm{op}}) \otimes \vec{s}_{\tilde{j}}(\tilde{h})^* \vec{s}_j(h) \Big\| \leq c_q n^{f(\sigma,\tilde{\sigma},\sigma')} \sup_j \|a_j\| \sup_j \|\tilde{a}_j\| \|y^{\mathrm{op}}\| \|\tilde{y}^{\mathrm{op}}\|.$$

The function f is obtained as follows. Let α be the number of pairs in σ being linked to either a pair or singleton in $\tilde{\sigma}$, and similarly let β be the number of linked pairs. Then we find

$$f(\sigma,\tilde{\sigma},\sigma') = \frac{|\sigma_s|}{2} + |\sigma_p| - \alpha + \frac{\alpha}{2} + \frac{|\tilde{\sigma}_s|}{2} + |\tilde{\sigma}_p| - \beta + \frac{\beta}{2} = \frac{m+\tilde{m}}{2} - \frac{\alpha+\beta}{2}$$

using row and column vectors $e_{1,i_1,\ldots,i_l}$, $e_{i_1,\ldots,i_l,1}$ for the number l of links in σ'. Thus for $\alpha+\beta > 0$ we obtain 0 in the limit and therefore only those σ' which link singletons to singletons give a contribution in the limit. Now we use Pisier's version [2000, Sublemma 3.3] of the Möbius transform. Let σ' be a fixed partition with pairs $\{\{l_1,r_1\},\ldots,\{l_p,r_p\}\}$. Then there are unitaries $\lambda_j^{\sigma'}$ in a product of free group factors such that

$$S_j(h) = s_j \otimes \lambda_j^{\sigma'}$$

satisfies

$$a(\sigma') := \sum_{\sigma'' \geq \sigma'} \Psi^{\sigma'}(\tilde{x}_{\tilde{\sigma}}\tilde{y}^{\mathrm{op}} \otimes x_\sigma y^{\mathrm{op}}) = \sum_{\tilde{j}_k, j_k} (\overline{\vec{\pi}_j(a)y^{\mathrm{op}}} \otimes \vec{\pi}_j(\tilde{a})\tilde{y}^{\mathrm{op}}) \otimes (\mathrm{id} \otimes E)(\vec{S}_j(\tilde{h})^* \vec{S}(h))$$
$$= (\mathrm{id} \otimes E)\psi(\widetilde{X}_{\tilde{\sigma}}^{\sigma'} \otimes X_\sigma^{\sigma'}).$$

Here

$$X_\sigma^{\sigma'} = \Big(n^{-\frac{m}{2}} \sum_{(j_1,\ldots,j_m) \leq \sigma} \pi_{j_1}(a_1) \cdots \pi_{j_m}(a_m) \otimes S_{j_1}(h_1) \cdots S_{j_m}(h_m) \Big)$$

and the corresponding expression for $\widetilde{X}_{\tilde{\sigma}}^{\sigma'}$ depends on σ'. Moreover, there exists a Möbius function $\mu(\cdot,\cdot)$ such that, see [Pisier 2000, Proposition 1],

$$\Psi^{\sigma'}(\tilde{x}_{\tilde{\sigma}}\tilde{y}^{\mathrm{op}} \otimes x_\sigma y^{\mathrm{op}}) = \sum_{\pi \geq \sigma'} \mu(\sigma',\pi)a(\pi).$$

The advantage of this representation comes from the fact that we can actually calculate P_k for such a fixed σ'. Recall that we may assume that σ' is a pair/singleton partition. For fixed $n \in \mathbb{N}$ and an element $\eta_n = S \otimes W_n$, W_n a Wick word of length k, we obtain

$$\operatorname{tr}(S(\overline{(\vec{\pi}_j(a)y^{\mathrm{op}} \otimes \vec{\pi}_j(\tilde{a})\tilde{y}^{\mathrm{op}})})) \, \tau(W_n s_{\tilde{j}_1} \cdots s_{\tilde{j}_{\tilde{m}}} s_{j_m} \cdots s_{j_1})$$

such that $(\tilde{j}_1, \ldots, \tilde{j}_{\tilde{m}}, j_m, \ldots, j_1) = \sigma'$. Then we obtain a nonzero term only if $|\sigma'_s| = k$ has exactly k singletons. Hence we find that

$$(\mathrm{id} \otimes P_k)(\psi(\xi \otimes \eta)) = \sum_{|\sigma'_s|=k} \Psi^{\sigma'}(\tilde{x}_{\tilde{\sigma}} \tilde{y}^{\mathrm{op}} \otimes x_\sigma y^{\mathrm{op}}) = \sum_{|\sigma'_s|=k} \sum_{\pi \geq \sigma'} \mu(\sigma, \pi) a(\pi).$$

Therefore we are left to consider

$$a(\pi) = (\mathrm{id} \otimes E)(\Psi(\widetilde{X}^\pi_{\tilde{\sigma}} \otimes X^\pi_\sigma)).$$

In order to use Remark 5.8 we have to modify the variables X^π_σ. Indeed, for every pair $p = \{l, r\}$ in π we introduce a label e_p and replace $s(e_{j_l} \otimes h_l)$ by $s(e_{j_l} \otimes e_p \otimes h_l)$, and $s(e_{\tilde{j}_r} \otimes \tilde{h}_r)$ by $s(e_{\tilde{j}_r} \otimes e_p \otimes h_r)$. For the remaining singletons we replace $s(e_j \otimes h_j)$ by $S_j = s(e_j \otimes e_0)$ and work in the Hilbert space $H' = H \otimes \ell_2$. Using the so modified X^π_σ's we still have

$$a(\pi) = (\mathrm{id} \otimes E_{\Gamma_q(\ell_2 \otimes H \otimes e_0)})_n \Psi(\widetilde{X}^\pi_{\tilde{\sigma}} \otimes X^\pi_\sigma) = \lim_{j \to \infty} \Psi(\alpha_{o_j}(\widetilde{X}^\pi_{\tilde{\sigma}}) \otimes \alpha_{o_j}(X^\pi_\sigma))$$

for any sequence (o_j) of orthogonal transformations such that $o_j(e_0) = e_0$, which converges weakly to $e_0^\perp$. For elements in $\widehat{C}(H)$ the limit for $j \to \infty$ converges, and hence this remains true for the norm closure. Thus for an element $x \in \widehat{C}(H)$ in the kernel of q_H we find $q_{H'}(x) = 0$ and hence

$$\langle x, a(\pi) \rangle = \lim_{j \to \infty} \langle x, \Psi(\alpha_{o_j}(\widetilde{X}^{\sigma'}_{\tilde{\sigma}}) \otimes \alpha_{o_j}(X^{\sigma'}_\sigma)\rangle = 0.$$

Using linear combinations we deduce indeed that $\langle P_k(x), \Psi(\tilde{x}_\sigma \tilde{y}^{\mathrm{op}} \otimes x_\sigma y^{\mathrm{op}}) \rangle = 0$. □

Corollary 5.10. *Let m_α be multipliers given by the cb-approximation property for $\Gamma_q(H)$:*

(i) *Then* $(\mathrm{id} \otimes m_\alpha)_n$ *extend to completely bounded maps on* $C(H)$ *with* $\limsup_\alpha \|(\mathrm{id} \otimes m_\alpha)_n\|_{\mathrm{cb}} = 1$, *and* $\lim_\alpha f_\alpha(k) = 1$, *where f_α are the associated scalar finitely supported functions. In particular, the maps $\tilde{\varphi}_n$ used in the proof of Theorem 5.1 above are completely bounded with* $\limsup_n \|\tilde{\varphi}_n\|_{\mathrm{cb}} = 1$.

(ii) *Let* $L(H) = \overline{C(H)}^{\mathrm{so}} \subset B(\mathcal{L})$ *and note that $L(H)$ is spanned by "extended Wick words" (i.e., images of extended Wick words through Φ) such that $L^2_k(L(H))$ (i.e., the $\|\cdot\|_2$-closed linear span of the extended Wick words of degree k) is finitely generated over B. Then there exists a modified family $f_\alpha(N)^* : L(H)_* \to L(H)_*$ converging in the point-norm topology.*

Proof. Since $(\mathrm{id} \otimes m_\alpha)(T) = \sum_k f_\alpha(k) P_k(T)$, we see that $\|\Phi_H \circ (\mathrm{id} \otimes m_\alpha)\|_{\mathrm{cb}} \leq 1 + \varepsilon_\alpha$ and also $\ker(q_H) \subset \ker(\Phi_H \circ (\mathrm{id} \otimes m_\alpha))$. But that means that there is a unique map $\tilde{m}_\alpha : \widehat{C}(H)/\ker(q_H) \to B(\mathcal{K})$ such that

$$\|\tilde{m}_\alpha\|_{\mathrm{cb}} = \|m_\alpha\|_{\mathrm{cb}} \leq 1 + \varepsilon_\alpha.$$

However, $\widehat{C}(H)/\ker(q_H) = C(H)$ completely isometrically, and hence $\tilde{m}_\alpha = (\mathrm{id} \otimes m_\alpha)$ coincides with the densely defined map $(\mathrm{id} \otimes m_\alpha)W(\sigma, \xi, a, y) = f_\alpha(|\sigma_s|)W(\sigma, \xi, a, y)$. Let us now consider a finite-dimensional subspace $H_0 \subset H$. Since $L_k^2(L(H))$ is finitely generated over B, we deduce that the projection P_d is normal on $L(H_0)$. Hence the maps m_α are also normal and restricted to the weakly dense subspace $C(H_0)$ we know that

$$\|m_\alpha\|_{\mathrm{cb}} \leq (1 + \varepsilon_\alpha).$$

Since a weakly dense subspace is norming for $L(H_0)_*$ we deduce that $\|(m_\alpha)_* : L(H_0)_* \to L(H_0)_*\|_{\mathrm{cb}} \leq (1 + \varepsilon_\alpha)$. Hence the normal map m_α coincides with the normal map $((m_\alpha)_*)^*$ and satisfies the same cb-norm estimate. Moreover, since we have normal conditional expectations $\mathcal{E}_{H_0} : L(H) \to L(H_0)$ so that $\bigcup_{H_i} \mathcal{E}_{H_i}(L(H_i))_*$ is norm dense in $L(H)_*$, we deduce that $(m_\alpha)_*$ extends to a completely bounded map of cb-norm at most $1 + \varepsilon_\alpha$ and hence $m_\alpha = ((m_\alpha)_*)^*$ is indeed a normal extension of the map $m_\alpha : C(H) \to C(H)$ with the same cb-norm estimate. This concludes the proof of (ii). □

The remainder of the subsection is devoted to proving some auxiliary results which will help us construct a standard form for the von Neumann algebra $\mathcal{N}$ which was used in the proof of Theorem 5.1. This standard form will be crucial in the proof of the main technical theorem.

Lemma 5.11. *There exists an action by $*$-automorphisms $\alpha : \mathcal{O}(H) \to \mathrm{Aut}(\mathcal{N})$ such that*

$$\alpha_o(\pi(x)\theta(y^{\mathrm{op}})) = \pi(\alpha_o(x))\theta(y^{\mathrm{op}}), \quad o \in \mathcal{O}(H),\ x \in M,\ y^{\mathrm{op}} \in P^{\mathrm{op}}.$$

Moreover, let E_0 be the orthogonal projection of $\mathcal{L}$ onto the closed linear span of the extended Wick words of degree zero. For $T \in \mathcal{N}$ the condition

$$\alpha_o(T) = T \quad \text{for all } o \in \mathcal{O}(H)$$

implies that $[T, E_0] = 0$.

Proof. Let us recall that $\mathcal{N}$ acts on

$$\begin{aligned}\mathcal{H} &= \mathrm{span}\{\pi(x_\sigma)(y \otimes 1)\theta(z^{\mathrm{op}})((1 \otimes_{\mathcal{A}} 1) \otimes 1) : x_\sigma \in M,\ y \in M,\ z \in P\} \\ &\subset \big((L^2(\mathcal{M}) \otimes_{\mathcal{A}} L^2(P)) \otimes L_2(\Gamma_q(\ell^2 \otimes H))\big)^\omega.\end{aligned}$$

Recall here that H is infinite-dimensional, and thanks to second quantization, $u_o = (\mathrm{id} \otimes \alpha_o)_n$ acts on $\mathcal{H}$ as a unitary. By normality, we deduce that $\alpha_o(x) = u_o x u_o^*$ extends to a $*$-automorphism of $\mathcal{N}$ and moreover, $\alpha_o(\theta(y^{\mathrm{op}})) = \theta(y^{\mathrm{op}})$. Let $o_i \in \mathcal{O}(H)$ be a family of orthogonal transformations of H such that $o_i(h)$ goes to 0 weakly in H. Let $\xi = \pi(x_\sigma)(y \otimes_{\mathcal{A}} z \otimes 1)$ and $\eta = \pi(x'_{\sigma'})(y' \otimes_{\mathcal{A}} z' \otimes 1)$. Then we obtain

$$\begin{aligned}\lim_i (u_{o_i}(\xi), \eta) &= \lim_i \lim_{n \to \omega} n^{-\frac{m+m'}{2}} \sum_{(j_k)=\sigma,\ (j'_{k'})=\sigma'} (\vec{\pi}_j(x)(y \otimes_{\mathcal{A}} z), \vec{\pi}_{j'}(x')(y' \otimes_{\mathcal{A}} z')) \\ &\qquad\qquad \tau(s_{j_1}(o_i(h_1)) \cdots s_{j_m}(o_i(h_m)) s_{j'_{m'}}(h'_{m'}) \cdots s_{j'_1}(h'_1)) \\ &= 0.\end{aligned}$$

Indeed, we expand the sum into the summation over $\sigma'' \in P_{1,2}(m+m')$ and execute the limit over n. Then we observe that the coefficients remain uniformly bounded. However, $o_i(h_k)$ is eventually orthogonal to every $h'_{k'}$ and then the moment formula for the q-gaussian yields 0 in the limit. We have therefore shown that u_{o_i} converges weakly to E_0, the projection onto words of length 0 in the second component. By taking convex combinations we find a net such that

$$\mathrm{SOT}-\lim_s \sum_i \alpha_i^s u_{o_i} = E_0.$$

Thus for $T \in N$ with $\alpha_o(T) = T$ for all o, we deduce that $[u_o, T] = 0$ and hence

$$E_0(T(\xi)) = \lim_s \sum_i \alpha_i^s u_{o_i} T(\xi)) = T\Big(\lim_s \sum_i \alpha_i^s u_{o_i}(\xi)\Big) = T(E_0(\xi)).$$

This means $E_0 T = T E_0$ as desired. □

Lemma 5.12. *Let $B \vee P^{\mathrm{op}} \subset B(L^2(\mathcal{M}) \otimes_{\mathcal{A}} L^2(P))$. Then the natural inclusion map*

$$\pi : B \vee P^{\mathrm{op}} \to \mathcal{N}$$

is normal.

Proof. By density it suffices to consider $\xi_n = \pi(x_\sigma^n)(y \otimes_A z)$ and $\eta_n = \pi(\tilde{x}_{\tilde{\sigma}}^n)(\tilde{y} \otimes_A \tilde{z})$. We may assume that x_σ and $\tilde{x}_\sigma$ is a Wick word. Our goal is to analyze

$$\phi(T) = \lim_{n\to\omega} \langle T\xi_n, \eta_n\rangle.$$

Let us first fix $n \in \mathbb{N}$. Then $\omega_n(T) = \langle T\xi_n, \eta_n\rangle$ is normal, and hence it suffices to assume $T = b\theta(p^{\mathrm{op}})$. It turns out that we need $|\sigma| = |\tilde{\sigma}| = k$ and then

$$\omega_n(T) = \frac{n\cdots(n-k+1)}{n^k} \sum_{\gamma\in S_k} q^{\mathrm{inv}(\gamma)} \tau\big(\tilde{z}^* E_A(\tilde{y}^* \pi_{\gamma(k)}(\tilde{x}_k)\cdots\pi_{\gamma(1)}(\tilde{x}_1) b \pi_1(x_1)\cdots\pi_k(x_k) y) z p\big).$$

Thanks to Lemma 5.2, we may replace $L^2(\mathcal{M})$ by $L^2(D) \otimes_B L^2(\mathcal{M})$ in the definition of $\mathcal{H}$. For fixed γ we may now define

$$x_\gamma = \alpha_{1,\dots,k}(x) \otimes_B y \otimes_{\mathcal{A}} z, \quad \tilde{x}_\gamma = \alpha_{\gamma(1),\dots,\gamma(k)}(\tilde{x}) \otimes_B \otimes \tilde{y} \otimes_{\mathcal{A}} \tilde{z}.$$

Since ω_n is normal we deduce that

$$\omega_n(T) = \sum_\gamma q^{\mathrm{inv}(\gamma)} \frac{n\cdots(n-k+1)}{n^k} \langle T(x_\gamma), \tilde{x}_\gamma\rangle$$

for all $T \in B \vee P^{\mathrm{op}}$. Since the summation is finite and the scalar coefficients converge, the limit exists for all $T \in B \vee P^{\mathrm{op}}$ and results in a normal functional $\phi(T)$ given by the same sum but with coefficient 1 instead of $n\cdots(n-k+1)/n^k$. □

Proposition 5.13. *Assume that for every finite-dimensional Hilbert space H, $L_k^2(M(H))$ is finitely generated as a right B-module* (*note that in particular this is the case if* $\dim_B(D_k(S)) < \infty$ *for all k*).

Then:

(i) *There exists a faithful normal conditional expectation* $\mathcal{E} : \mathcal{N} \to B_P = \pi(B) \vee \theta(P^{\mathrm{op}})$.

(ii) *The action* α *is implemented by an sot-continuous family of unitary operators* $(V_o)_{o\in\mathcal{O}(H)}$ *on* $L^2(\mathcal{N})$.

(iii) $L^2(\mathcal{N}) = \overline{\bigoplus_{k\geq 0} W_k(M)L^2(B_P)}$ *and* $V_o(\pi(x_\sigma)\xi) = \pi(\alpha_o(x_\sigma))\xi$ *for* $x_\sigma \in M$, $\xi \in L^2(B_P)$. *Moreover*, $\mathcal{E}|_{\pi(M)} = E_B$, *where* $E_B : \pi(M) \to \pi(B)$ *is the conditional expectation.*

Proof. For a subspace $H' \subset H$ we use the notation

$$\mathcal{H}(H') = \{\pi(x_\sigma)((y \otimes_{\mathcal{A}} z) \otimes 1) : y \in M,\ z \in P,\ x_\sigma = x_\sigma(x_1, \dots, x_m, h_1, \dots, h_m),\ h_i \in H'\}$$

for the subspace generated by H'-Wick words. Let $\iota_{H'} : \mathcal{H}(H') \subset \mathcal{H}$ be the canonical inclusion map and $F_{H'}(T) = \iota_{H'}^* T \iota_{H'}$ the induced completely positive map. Certainly, we have $F_{H'}(\theta(y^{\mathrm{op}})) = \theta(y^{\mathrm{op}})$ and

$$F_{H'}(x_\sigma) = E_{H'}(x_\sigma).$$

Indeed, if a Wick word x_σ contains a singleton $h_i \in (H')^\perp$, then $F_{H'}(x_\sigma) = 0$. Using $h_i \in H' \cup (H')^\perp$ we deduce the assertion by linearity. Thus $F_{H'}(\mathcal{N}(H)) = \mathcal{N}(H_i) \subset B(\mathcal{H}(H'))$ defines a normal surjective conditional expectation $F_{H'}$. Let $e_{H'}$ be the support of $F_{H'}$. We observe that $\pi(M(H'))$ and $\theta(P^{\mathrm{op}})$ belong to the multiplicative domain of $F_{H'}$. Let $\widetilde{N}(H') \subset \mathcal{N}(H)$ be the von Neumann algebra generated by $\pi(M(H'))$ and $\theta(P^{\mathrm{op}})$ inside $\mathcal{N}_P(H)$. According to Remark 5.8 and Kaplansky's density theorem, we deduce that $F_{H'}$ induces the same weak* topology on the unit ball of $\widetilde{N}(H')$. This means that the tautological embedding $\sigma_{H'H} : \mathcal{N}(H_i) \to \mathcal{N}(H)$ given by $\sigma_{H'H}(x_\sigma) = x_\sigma$ and $\sigma_{H'H}(\theta(y^{\mathrm{op}})) = \theta(y^{\mathrm{op}})$ satisfies $F_{H'}\sigma = \mathrm{id}_{\mathcal{N}(H_i)}$ and $F_{H'}$ is an isomorphism when restricted to $\widetilde{N}(H')$. We denote by $\mathcal{E}_{H'} = \sigma_{H_i H} F_{H'} : \mathcal{N} \to \mathcal{N}$ the resulting, not necessarily faithful, conditional expectation. Let H_i be an increasing net of finite-dimensional spaces whose union is dense. Since $\bigcup_i \mathcal{H}(H_i)$ is norm dense, we deduce that $\widehat{\mathcal{E}}_{H_i}(x)$ converges weakly to x as i goes to infinity along the net of finite-dimensional subspaces. Recall that the multiplier maps m_α are normal and commute with every $\mathcal{E}_{H_i}$. Adding convex combination we may find a new completely contractive net, still denoted by m_α, converging in the strong, strong* operator topology. Thus we may assume that

$$\lim_i \lim_\alpha (\widehat{E}_{H_i}(m_\alpha x)) = x$$

converges strongly for all $x \in \mathcal{N}$. In our next step we consider $H' = 0$, i.e., the map $\iota : L^2(M) \otimes_{\mathcal{A}} L^2(P) \to \mathcal{H}$, given by $\iota(y \otimes_{\mathcal{A}} z) = (y \otimes_{\mathcal{A}} z) \otimes 1$. This yields a completely positive map $\Phi(T) = \iota^* T \iota$ such that $\Phi(\theta(y^{\mathrm{op}})) = \theta(y^{\mathrm{op}})$ and $\Phi(\pi(b)) = \pi(b)$. On the other hand, for a Wick word $x = W_\sigma$, we see that

$$\langle \pi(x)\iota(y \otimes_{\mathcal{A}} z), \iota(y' \otimes_{\mathcal{A}} z')\rangle = \lim_{n\to\omega} n^{-\frac{m}{2}} \sum_{(i_j)=\sigma} \langle \vec{\pi}(x)(y \otimes_{\mathcal{A}} z), y' \otimes_{\mathcal{A}} z'\rangle \tau(s_{j_1}(h_1) \cdots s_{j_m}(h_m)) = 0.$$

By normality, we deduce that $\Phi(\mathcal{N}) = B \vee P^{\mathrm{op}} \subset B(L^2(\mathcal{M}) \otimes_{\mathcal{A}} L^2(P))$. Let us denote by $B_P = \Phi(\mathcal{N})$ the resulting von Neumann algebra and by e_{B_P} the support of $\mathcal{E} = \Phi|_{\mathcal{N}_P}$. Since the Wick words of order 0 are obviously invariant under α_o for all $o \in \mathcal{O}(H)$ and

$$\mathcal{E}\alpha_o(x) = \alpha_o(\mathcal{E}(x)) = \mathcal{E}(x),$$

we must have $\alpha_o(e_{B_P}) = e_{B_P}$ for every $o \in \mathcal{O}(H)$. More precisely, $1 - e_{B_P}$ is the projection of the ideal $I = \{x : \mathcal{E}(x^*x) = 0\}$ and we certainly have $\alpha_o(I) = I$. This implies $\alpha_o(1 - e_{B_P}) = 1 - \alpha_o(e_{B_P})$. We deduce that for all α we have $\alpha_o(m_\alpha e_{B_P}) = m_\alpha e_{B_P}$ and hence, thanks to Lemma 5.11 we know that $[E_0, m_\alpha(e_{B_P}))] = 0$. Now, we fix α and consider $x_{i,\alpha} = \mathcal{F}_{H_i}(m_\alpha(e_{B_P})) = m_\alpha F_{H_i}(e_{B_P})$. This means

$$x_{i,\alpha} = \sum_{k \le k(\alpha)} x_k,$$

where $x_k = P_k(x)$. However, we have a finite basis $\xi_{k,s}$ of $L^2_k(M(H))$ over B made of elements in $W_k(H_i)$ and hence for all $z = \pi(x_{\sigma'})\theta(y^{\mathrm{op}})$ we find

$$P_k(z) = \sum_s \pi(\xi_{k,s}) E_B(\xi_s^* x_{\sigma'})\theta(y^{\mathrm{op}}).$$

Since P_k is normal we deduce that there are coefficients $a_s \in \pi(B) \vee \theta(P^{\mathrm{op}})$ such that

$$x_k = \sum_s \pi(\xi_{k,s}) a_{s,k} \in \mathcal{N}(H_i).$$

Note here that we have rewritten m_α as normal map, because the maps $T_{k,s}(x) = \pi(\xi_{ks})\sigma(\mathcal{E}(\xi_{ks}^* x))$ are normal, thanks to Lemma 5.12. Note also that due to Lemma 5.12, $\sigma(B \vee P^{\mathrm{op}}) = \pi(B) \vee \theta(P^{\mathrm{op}}) \subset \mathcal{N}$. On the other hand the projection P_{H_i} onto the range of ι_{H_i} contains the range of ι and hence

$$[E_0, \iota_{H_i}^* \hat{m}_\alpha(e_{B_P})\iota_{H_i}] = \iota_{H_i}^*[E_0, \hat{m}_\alpha(e_{B_P})]\iota_{H_i} = 0.$$

Thus we have $[E_0, x_{i,\alpha}] = 0$. Let us consider $\eta = (y \otimes_{\mathcal{A}} z) \otimes 1$. We deduce that

$$x_{i,\alpha}(\eta) = \sum_{k \le k(\alpha)} \sum_s \pi(\xi_{k,s}) a_{k,s}(\eta).$$

Moreover, we see that

$$E_B(\xi_{s,k}^* x_{i,\alpha}(\eta)) = E_B(\xi_{s,k}^* \xi_{s,k}) a_{k,s}(\eta).$$

We may assume that $f_{k,s} = E_B(\xi_{s,k}^* \xi_{s,k})$ is a projection in B and $a_{k,s} = f_{ks} a_{ks}$. Since the conditional expectation can be calculated using vectors in the Hilbert space, we deduce that

$$a_{k,s}(\eta) = E_B(\xi_{s,k}^* x_{i,\alpha}(\eta)) = E_B(\xi_{s,k}^* E_0(x_{i,\alpha}(\eta))) = 0$$

for all $k > 0$. Thus only the coefficient for $k = 0$ survives and hence $x_{i,\alpha} \in \sigma(B \vee P^{\mathrm{op}})$. This remains true for the limit along α; i.e., $x_i = F_{H_i}(e_{B_P}) \in \sigma(B \vee P^{\mathrm{op}})$. Since $\bigcup_i \iota_{H_i}$ is norm dense we find that

$$e_{B_P} = w^* - \lim_i F_{H_i}(e_{B_P}) \in \sigma(B \vee P^{\mathrm{op}}).$$

The restriction of the normal map $\sigma \circ \mathcal{E}$ to $\sigma(B \vee P^{\mathrm{op}})$ is the identity. This implies

$$1 - e_{B_P} = \sigma \circ \mathcal{E}(1 - e_{B_P}) = \sigma \circ \mathcal{E}(e_{B_P}(1 - e_{B_P})e_{B_P}) = 0.$$

Thus $e_{B_P} = 1$ and $\mathcal{E}$ is indeed a faithful normal expectation. Now it is easy to conclude the proof of the crucial assertion (iii). Indeed, we may assume that $\pi(B)$ and $\theta(P^{\mathrm{op}})$ both admit weakly dense separable C^*-subalgebras and hence fix a faithful normal state ϕ on B_P. Then $\psi = \phi \circ \mathcal{E}$ satisfies Connes' commutativity relation for the modular group $\mathcal{E}(\sigma_t^{\psi}(x))=\sigma_t^{\phi}(\mathcal{E}(x))$. We refer to [Haagerup, Junge and Xu 2010] for the fact that we have a natural embedding of the Haagerup spaces $L^p(B_P)\to L^p(\mathcal{N})$ given by

$$\iota_p(xd_{\phi}^{\frac{1}{p}})=xd_{\psi}^{\frac{1}{p}}$$

for the densities $d_\phi\in L^1(B\vee P^{\mathrm{op}})$, $d_\psi\in L^1(\mathcal{N}_P)$ associated with the states. Moreover, the support of d_ψ is 1. This implies that $L^2(\mathcal{N})=\mathcal{N}L^2(B_P)$. By approximation in the C^*-algebra generated by $\pi(M)$ and $\theta(P^{\mathrm{op}})$, we see that the span of elements of the form

$$\pi(x_\sigma)\theta(y^{\mathrm{op}})d_{\psi}^{\frac{1}{2}}$$

is dense in $L^2(\mathcal{N})$. However, we have

$$\begin{aligned}\mathrm{tr}\big((\pi(x_\sigma)\theta(y^{\mathrm{op}})d_{\psi}^{\frac{1}{2}})^*\pi(x_\nu)\theta(z^{\mathrm{op}})d_{\psi}^{\frac{1}{2}}\big)&=\mathrm{tr}\big(\theta(y^{\mathrm{op}})^*\pi(x_\sigma)^*\pi(x_\nu)\theta(z^{\mathrm{op}})d_\psi\big)\\&=\mathrm{tr}\big(\theta(y^{\mathrm{op}})^*\theta(z^{\mathrm{op}})\pi(x_\sigma)^*\pi(x_\nu)d_\psi\big)\\&=\psi\big(\theta(y^{\mathrm{op}})^*\theta(z^{\mathrm{op}})\pi(x_\sigma)^*\pi(x_\nu)\big)\\&=\phi\big(\mathcal{E}(\theta(y^{\mathrm{op}})^*\theta(z^{\mathrm{op}})\pi(x_\sigma)^*\pi(x_\nu))\big)\\&=\phi\big(\theta(y^{\mathrm{op}})^*\theta(z^{\mathrm{op}})\mathcal{E}(\pi(x_\sigma)^*\pi(x_\nu))\big)\\&=\phi\big(\theta(y^{\mathrm{op}})^*\theta(z^{\mathrm{op}})E_B\pi(x_\sigma)^*\pi(x_\nu)\big).\end{aligned}\tag{5-1}$$

For the proof of the last equality, we may assume that x_σ and x_ν are reduced Wick words. As in Lemma 5.12, we see that

$$\begin{aligned}&\langle\pi(x_\xi)(y\otimes_{\mathcal{A}}z),\pi(x_\nu)(\tilde{y}\otimes_{\mathcal{A}}\tilde{z})\rangle\\&=\lim_n n^{-\frac{|\sigma|+|\nu|}{2}}\sum_{(j_k)=\sigma,\ (\tilde{j}_{\tilde{k}})=\tilde{\sigma}}\tau\big(\tilde{z}^*E_{\mathcal{A}}(\tilde{y}^*\vec{\pi}_{\tilde{j}}(\tilde{x})^*\vec{\pi}_j(x)y)z\big)\,\tau(s_{\tilde{j}_m}(\tilde{h}_m)\cdots s_{\tilde{j}_1}(\tilde{h}_1)s_{j_1}(h_1)\cdots s_{j_m}(h_m))\\&=\delta_{|\sigma|,|\nu|}\sum_{\gamma\in S_k}q^{\mathrm{inv}(\sigma)}n^{-|\sigma|}\sum_{(j_1,\dots,j_k)}\tau\big((\alpha_{j_{\gamma(1)},\dots,j_{\gamma(k)}}(\tilde{x}))^*\alpha_{j_1,\dots,j_k}(x)yE_{\mathcal{A}}(z\tilde{z}^*)\tilde{y}^*\big)\\&=\delta_{|\sigma|,|\nu|}\sum_{\gamma\in S_k}q^{\mathrm{inv}(\sigma)}\tau(b(x,\tilde{x},\gamma)yE_{\mathcal{A}}(z\tilde{z}^*)\tilde{y}^*).\end{aligned}$$

The limit $b(x,\tilde{x},\gamma)\in B$ only depends on x and $\tilde{x}$ and the permutation γ. Placing the summation inside we find indeed $E_B(x_\nu^*x_\sigma)$. Thus we have shown that $\mathcal{E}|_{\pi(M)}=E_B$. We deduce that the spaces $W_k(M)L_2(B_P)$ are mutually orthogonal. Finally, we have to discuss the action $\alpha:\mathcal{O}(H)\to\mathrm{Aut}(\mathcal{N})$. For an arbitrary $*$-automorphism α of $\mathcal{N}$, we may define the action on $L^2(\mathcal{N})$ via

$$\alpha(xd_{\psi}^{\frac{1}{2}})=\alpha(x)(d_\psi\circ\alpha^{-1})^{\frac{1}{2}}.$$

It is easy to show that this action is independent of the choice of a normal faithful density d associated with state ψ. Here $d\circ\alpha^{-1}$ is the density of $\psi\circ\alpha^{-1}$. Thus we deduce from $\alpha_o(\theta(y^{\mathrm{op}}))=\theta(y^{\mathrm{op}})$ and the fact that $\psi \circ \alpha_o = \psi$, that

$$\alpha_o(\pi(x_\sigma)\theta(y^{\mathrm{op}})d_\psi^{\frac{1}{2}}) = \alpha_o(\pi(x_\sigma))\theta(y^{\mathrm{op}})d_\psi^{\frac{1}{2}},$$

as expected. □

Remark 5.14. A posteriori, we deduce that under the assumptions above, $F_{H'}$ is faithful for every subspace $H' \subset H$ because $\mathcal{E} = \mathcal{E} F_{H'}$.

6. The deformation bimodules are weakly contained in $L^2(M) \otimes_B L^2(M)$ for polynomial dimensions of $D_k(S)$ over B

6A. *Norm estimates for decomposable maps.* Let H be an M-N bimodule over finite von Neumann algebras M and N. We will introduce some norms which will enable us to show that the M-N bimodules associated to certain maps $\Phi : M \to L^1(N) = N_*^{\mathrm{op}}$ are weakly contained in H. To be more precise define

$$\|\Phi\|_H = \inf\left\{ \sum_j \|\xi_j\| \, \|\eta_j\| : \tau(\Phi(x)y) = \sum_j \langle (x \otimes y^{\mathrm{op}})\xi_j, \eta_j \rangle \right\}.$$

The infimum is taken over elements $\xi_j, \eta_j \in H$.

Lemma 6.1. *Let K be an M-N bimodule such that for a total set of vectors $\xi \in K$ the map $\Phi_\xi : M \to L^1(N)$ defined by*

$$\tau(\Phi_\xi(x)(y)) = \langle (x \otimes y^{\mathrm{op}})\xi, \xi \rangle = \langle x\xi y, \xi \rangle$$

satisfies $\|\Phi_\xi\|_H < \infty$. Then K is weakly contained in H.

Proof. Let us recall that $K \prec H$ if and only if we have the relation between the kernels

$$\ker(\pi_H) \subset \ker(\pi_K),$$

where $\pi_H : M \otimes_{\mathrm{bin}} N^{\mathrm{op}} \to B(H)$, respectively $\pi_K : M \otimes_{\mathrm{bin}} N^{\mathrm{op}} \to B(K)$ are the canonical representations. Let $z = \lim z_j$ be a limit of norm-1 elementary tensors which converges to an element $z \in \ker(\pi_H)$ with respect to the max norm. Let $\xi \in K$ such that $\|\Phi_\xi\|_H < \infty$. This means we may assume that

$$\tau(\Phi_\xi(x)y) = \sum_l \alpha_l \langle \xi_l, x\eta_l y \rangle, \quad \|\xi_l\| \, \|\eta_l\| \le 1,$$

and $\sum_l |\alpha_l|$ is finite. Using $\|z_j\|_{\mathrm{bin}} \le 1$ and uniform convergence, we may interchange limits and deduce

$$\langle z\xi, z\xi \rangle = \lim_j \langle \xi, z_j^* z_j \xi \rangle = \sum_l \alpha_l \lim_j \langle \xi_l, z_j^* z_j \eta_l \rangle = \sum_l \alpha_l \langle z\xi_l, z\eta_l \rangle = 0.$$

Thus for any linear combination $\xi = \sum_k \xi_k$ of elements such that the Φ_{ξ_k}'s have finite H norm, we still have $\pi_K(z)\xi = 0$. By density this holds for all $\xi \in K$. □

As an illustration for the norm estimates let us prove the following result.

Lemma 6.2. *Let $H_B = L_2(M) \otimes_B L_2(M)$, and assume that $L^2_k(M)$ has dimension d_k over B. Let $P_k : L^2(M) \to L^2_k(M)$ be the orthogonal projection. Then*

$$\|P_k\|_{H_B} \le d_k.$$

Proof. We recall that

$$\langle x \otimes y^{\mathrm{op}}(c \otimes d), a \otimes b\rangle = \tau(b^* E_B(a^*xc)dy) = \tau(E_B(a^*xc)E_B(dyb^*)). \tag{6-1}$$

Assuming that ξ_j is a basis with $E_B(\xi_j^*\xi_i) = \delta_{ij}e_i$, e_i a projection, we see that

$$\tau(yP_k(x)) = \sum_j \tau(y\xi_j E_B(\xi_j^* x)) = \sum_j \langle x \otimes y^{\mathrm{op}}(1 \otimes 1), \xi_j \otimes \xi_j^*\rangle.$$

Since $\langle \xi_j \otimes \xi_j^*, \xi_j \otimes \xi_j^*\rangle = \tau(E(\xi_j^*\xi_j)E(\xi_j^*\xi_j)) \le \tau(e_j)$, we deduce the assertion. ☐

6B. ***Configurations.*** Our main goal here is to analyze the operators $\Phi_{\xi,\eta} : M \to L_1(M)$ given by $\Phi_{\xi,\eta}(x) = E_M(\xi x \eta)$, where ξ, η are elements in $\Gamma_q(B, A \otimes (H \oplus H))$. We will start with monomials

$$\xi = s(x_1, h_1) \cdots s(x_m, h_m), \quad \eta = s(x'_{m'}, h'_{m'}) \cdots s(x'_1, h'_1),$$

where $h_i, h'_{i'} \in H \times \{0\} \cup \{0\} \times H$. Although our goal is to obtain estimates for arbitrary x, we will first assume that $x = \zeta$ is a reduced Wick word from M and only contains singletons from $H \times \{0\}$. By considering the moment formula we can reorganize the trace using configurations

$$\tau(\zeta'\xi\zeta\eta) = \sum_{\alpha \text{ configuration}} \tau(\zeta'\Phi_\alpha(\zeta))$$

whenever ζ' is another reduced Wick word. Here a configuration $\alpha = (\sigma_{0\times H}, \sigma_{H\times 0}, I_{\xi,\zeta}, I_{\zeta,\eta})$ is given by

(i) a pair partition $\sigma_{0\times H}$ of $\{1, \dots, m\}\dot\cup\{m', \dots, 1\}$ so that all the pairs $\{l, r\}$ have indices in $0 \times H$;

(ii) a pair partition $\sigma_{H\times 0}$ of $\{1, \dots, m\}\dot\cup\{m', \dots, 1\}$ so that all the pairs $\{l, r\}$ have indices from $H \times 0$;

(iii) subsets $I_{\xi,\zeta} \subset \{1, \dots, m\}$, $I_{\zeta,\eta} \subset \{m', \dots, 1\}$ disjoint from the support $\bigcup \sigma_{0\times H} \cup \bigcup \sigma_{H\times 0}$ of the partitions above.

Indeed, using the moment formula for $\tau(\zeta'\xi\zeta\eta)$ we know that we have to take the sum over all pair partitions of length $m + m' + k + k'$, $k = |\zeta|$, $k' = |\zeta'|$. Every such pair partition has to respect the pairs of $0 \times H$ and that defines our $\sigma_{0\times H}$. Some pairs can combine elements from ξ and η with coefficients in $H \times 0$. This defines $\sigma_{H\times 0}$. Some partitions connect ξ and ζ and some ζ with η. The left-hand sides of the pairs between ξ and η define the set $I_{\xi,\zeta}$ and the right-hand sides of the pairs from ζ, η define $I_{\zeta,\eta}$. All the remaining pairs will connect ζ' and ζ. Since ζ and ζ' are themselves Wick words, there are no pairs connecting elements from ζ (ζ') with itself. We see that indeed, the sum over all partitions can be regrouped into first summing over all configurations (which only depend on ξ and η), and then summing over all partitions supported by these configurations. Let us note that once a configuration α is known we can determine exactly how many crossings will be produced by pairs in $0 \times H$. Indeed, we know that

$|I_{\xi,\zeta}|+|I_{\zeta,\eta}|$ many singletons will be removed from ζ. According to the position of the left legs in $\sigma_{0\times H}$ some extra crossing will be produced from the set $I_{\xi,\zeta}$. The same applies for $I_{\xi,\zeta}$. Here is an example:

$$a_1\ a_2\ b_1\ a_3\ b_3 \big| c_1\ c_2\ a_3\ d_1\ c_3\ c_4 \big| d_2\ b_3\ d_1\ a_1\ b_1\ .$$

Here $\sigma_{0\times H}$ are given by the positions of b_1 and b_3. The set $I_{\xi,\zeta}$ is given by the position of a_3 and $\sigma_{H\times 0}$ is given by the positions of a_1. The b's are responsible for $8+1+1+1$ crossings: eight crossings with c's, one crossing among themselves, one crossing coming from a's and b's, one crossing from the b's and d's. Thus $k(\alpha)=2\times(6-2)+1+2$.

In our next step we replace the monomials ξ and η by Wick words. This means we only have to sum over those configurations such that $\sigma_{0\times H}$ and $\sigma_{H\times 0}$ connect ξ and η and no pair ξ and η with itself. In addition the reduction procedure produces scalars and new operator-valued expression $\alpha_{j_1,\dots,j_l}(\beta)$ with $\beta\in D_k(S)$. We have proved the following simple combinatorial fact:

Lemma 6.3. *Let ξ and η be Wick words obtained by reduction and $\zeta\in M$ be a Wick word of length $k=|\zeta|$. For a fixed configuration α there is a number $k(\alpha)$ such that for all $-1\le q\le 1$*

$$\Phi_\alpha(\zeta)=q^{k(\alpha)}\tilde{\zeta},$$

where $\tilde{\zeta}$ is a linear combination of reduced words with smaller length $k-|I_{\xi,\zeta}|-|I_{\zeta,\eta}|$. Moreover, if $k\ge m+m'$ is the length of ζ, and L is the cardinality of $\sigma_{0\times H}$, then

$$k(\alpha)\ge(k-m-m')L.$$

We will give more precise information about $\tilde{\zeta}$ in the next paragraph.

6C. *Generalized q-gaussians.* As a tool we will use a slight generalization of the von Neumann algebra $\Gamma_q(B,A\otimes H)$. This generalization is based on matrix models of the ordinary q-gaussian von Neumann algebras. This approach was invented by Speicher [1992; 1993] and has been applied in many situations; see, e.g., [Biane 1997; Junge 2006; Junge, Palazuelos, Parcet, Perrin and Ricard 2015; Junge and Zeng 2015; Nou 2004; 2006]. Let $\mathrm{Br}:H\to\bigcap_p L^p(\Omega,\Sigma,\mu)$ the standard brownian motion so that $\mathrm{Br}(h)$ is a normal random variable and $(\mathrm{Br}(h),\mathrm{Br}(h'))=(h,h')$. The σ-algebra is chosen minimal. This construction is well known as the gaussian measure space construction. Given a self-adjoint matrix ε_{ij} with values $\{-1,1\}$, there are symmetries $v_j\in M_{2^n}(\mathbb{C})$ such that

$$v_iv_j=\varepsilon_{ij}v_jv_i.$$

Speicher's important idea is to choose the matrix ε_{ij} independently at random for all pairs. We will work with double indices $\varepsilon_{(j,t),(k,s)}$, which are independent as functions of the pairs $\{(j,t),(k,s)\}$ whenever $t\ne s$ or $j\ne k$ and satisfy

$$P(\varepsilon_{(jt),(ks)}=1)=\tfrac{1}{2}(1-Q_{s,t})$$

as long as $(j,t)\ne(ks)$ for a given matrix $Q_{s,t}$. This allows us to construct matrix models

$$u(t,h)=\Big(\frac{1}{\sqrt{n}}\sum_{j=1}^n v_{j,t}\otimes g_j(h)\Big)_n\in\prod_{n,\omega}(M_{2^n}(\mathbb{C})\otimes L^\infty(\Omega))_n$$

which satisfy

$$\tau(u(t_1,h_1)\cdots u(t_m,h_m)) = \sum_{\sigma\in P_2(m)} \prod_{\substack{\{a,b\}\in\sigma,\ \{c,d\}\in\sigma\\ a<c<b<d}} Q_{t_a t_c} \prod_{\{a,b\}\in\sigma} \langle h_a, h_b\rangle.$$

In other words the constant term $q^{\mathrm{inv}(\sigma)}$ is replaced by the product of the crossing inversions weighted according to Q. Indeed, by independence

$$\prod_{\substack{\{a,b\}\in\sigma,\ \{c,d\}\in\sigma\\ a<c<b<d}} Q_{t_a t_c} = \mathbb{E}\tau(v_{j_1,t_1}\cdots v_{j_m,t_m})$$

for $(j_1,\dots,j_m)\le\sigma$. In particular for a fixed t and $\|h\|=1$ the random variable $u(t,h)$ is just an ordinary q-gaussian. This central limit theorem is well known and goes back to [Speicher 1992; 1993]; see also [Junge 2006; Junge and Zeng 2015].

We may easily generalize this to the A-valued situation by considering a sequence of symmetric independent copies (π_j, B, A, D) and defining

$$u(t,h,a) = \Bigl(\frac{1}{\sqrt{n}}\sum_{j=1}^{n} v_{j,t}\otimes g_j(h)\otimes \pi_j(a)\Bigr)_n \in \prod_{n,\omega}(M_{2^n}(\mathbb{C})\otimes L^\infty(\Omega)\mathbin{\bar\otimes} D)_n.$$

For a subset $1\in S=S^*\subset A$, we denote the von Neumann algebra generated by the elements $u(t,h,a)$, $t\in Q,\ h\in H,\ a\in S$ by $\Gamma_Q^0(B,S\otimes H)$. Then define the von Neumann algebra $\Gamma_Q(B,S\otimes H)$ by the same procedure as in Definition 3.4. A look at the moment formula allows us to state the following fact.

Lemma 6.4. *Let $T_0\subset T$ be a nonempty subset such that $Q_{st}=q$ for all $s,t\in T_0$. Then $\Gamma_q(B,S\otimes H)$ embeds into $\Gamma_Q(B,S\otimes H)$ in a trace-preserving way.*

Remark 6.5. As observed in [Junge and Zeng 2015] the reduction procedure still works in the generalized q-gaussian setting.

Let us return to a configuration α as in Section 6B above. We replace the Wick words $W_q(\vec{a},\vec{h})$ and $W_q(\vec{a}',\vec{h}')$ by new Wick words $W_Q(\vec{a},\vec{h})\ W_Q(\vec{a}',\vec{h}')$ as follows. For a configuration α with a partition $\sigma_{0\times H}$ of the indices labeled with $0\times H$, we define a new matrix

$$Q_{st}(\tilde q) = \begin{cases}\tilde q & \text{if } h_s \text{ and } k_s \text{ are both in } 0\times H,\\ q & \text{else.}\end{cases}$$

Note that the matrix only depends on the first component $\sigma_{0\times H}$ of a configuration. For every pair $p=\{l,r\}\in\sigma_{0\times H}$ we introduce a label e_p and replace h_l and h'_r by $h_l\otimes e_p$ and $h'_r\otimes e_p$ to avoid over-counting. We denote by H_s, H'_t the modified vectors. Starting from

$$\xi_{Q(\tilde q)} = s_{Q(\tilde q)}(H_1,a_1)\cdots s_{Q(\tilde q)}(H_m),\quad \eta_{Q(\tilde q)} = s_{Q(\tilde q)}(H'_{m'},a'_{m'})\cdots s_{Q(\tilde q)}(H'_1,a'_1)$$

we apply the same reduction procedure (eliminating all the pairs from the nonreduced words $X_\sigma(\vec{h},\vec{a})$) for the W_q's and obtain the reduced Wick words $W_{Q(\tilde q)}(\vec{H},\vec{a}),\ W_{Q(\tilde q)}(\vec{H}',\vec{a}')$.

Lemma 6.6. *Fix $\sigma_{0\times H}$. The function*

$$F(\tilde{q}) = \sum_{\alpha,\alpha_1=\sigma_{0\times H}} E_M(W_{Q(\tilde{q})}(\vec{H},\vec{a})\zeta W_{Q(\tilde{q})}(\vec{H}',\vec{a}'))$$

is a polynomial in $\tilde{q}$ with lowest degree at least $(|\zeta|-m+m')L$ and largest degree at most $(|\zeta|+m+m')L$.

Proof. Let α be a configuration which contains $\sigma_{0\times H}$. Comparing the terms in the moment formula for

$$\tau(\zeta' W_q(\vec{h},\vec{a})\zeta W_q(\vec{h}',\vec{a}')) \quad \text{and} \quad \tau(\zeta' W_{Q(\tilde{q})}(\vec{h},\vec{a})\zeta W_{Q(\tilde{q})}(\vec{h}',\vec{a}'))$$

we see that they differ by the factor $(\tilde{q}/q)^{k(\alpha)}$ number of pairs. Note however, that $k(\alpha)$ only depends on α. This implies the assertion. □

6D. *Weak containment.* We need a simple fact about polynomials:

Lemma 6.7. *Let $[a,b]$ be an interval, $\mathcal{P}_d(a,b)$ the set of polynomials of degree d, and $a < t_0 < t_1 < \cdots < t_d < b$ distinct points. Then the map $\Phi : P_d(a,b) \to \mathbb{C}^{d+1}$, $\phi(f) = f(t_j)$, is injective. Moreover, there exists a matrix $a_{i,j}$ such that for every polynomial*

$$p(t) = \sum_{0\le k\le d} \alpha_k t^k$$

of degree $\le d$ we have

$$\alpha_k = \sum_j a_{k,j} f(t_j).$$

Proof. For $0 \le j \le d$ we define the polynomial $p_j(t) = \big(\prod_{i\neq j}(t_j - t_i)\big)^{-1} \prod_{i\neq j}(t-t_i)$, which has degree d. Then we see that $p_j(t_j) = 1$ and $p_j(t_i) = 0$ for $i \neq j$. In particular, the polynomials $(p_j)_{0\le j\le d}$ are linearly independent and hence $P_d(a,b) = \text{span}\{p_j : 0 \le j \le d\}$. This implies

$$p(t) = \sum_{0\le j\le d} p(t_j) p_j(t)$$

and in particular Φ is injective. Since moreover, the monomials are linearly independent in $C_\infty(a,b)$, we see that the linear map $\Psi(\alpha_0,\dots,\alpha_d) = \Phi\big(\sum_k \alpha_k t_j^k\big)$ is invertible and can be represented by the matrix $C_{j,k} = t_j^k$, the well-known Vandermonde matrix. Then $A = C^{-1}$ does the job. □

From now on we fix $\sigma = \sigma_{0\times H}$, and Wick words $\xi = W_q(\vec{H},\vec{a})$, $\eta = W_q(\vec{H}',\vec{a}')$ which are obtained after reduction from possible longer terms $s_q(h_1,a_1)\cdots s_q(h_m,a_m)$ and $s_q(h'_{m'},a'_{m'})\cdots s_q(h'_1,a'_1)$. This allows us to define

$$F_\sigma(t) = E_M\big(W_{Q(t)}(\vec{H},\vec{a})\zeta W_{Q(t)}(\vec{H},\vec{a})\big).$$

As in Section 6B, we assume that at least L labels of ξ and η are of the form $(0,h_i)$.

Corollary 6.8. *Fix m, m' and L. Then there exists a degree $D = D(m,m',L)$ such that for $q \in [a,b]$ and $a \le t_1 < \cdots < t_D \le b < 1$ there are coefficients γ_l such that*

$$E_M(\xi\zeta\eta) = \sum_\sigma \sum_l \left(\frac{q}{t_l}\right)^{(k-m-m')L} \gamma_l F_\sigma(t_l)$$

holds for $k = |\zeta| \geq 2(m+m')$. *Moreover, for some possibly different coefficients* $\tilde{\gamma}_l$

$$E_M(\xi\zeta\eta) = \sum_{\sigma}\sum_{l} \tilde{\gamma}_l F_\sigma(t_l)$$

holds for $|\zeta| \leq 2(m+m')$.

Proof. We fix σ and $k \geq m+m'$. Let $[a,b] \subset (-1,1)$ be an interval and $a = q$. The t_i's are all chosen bigger than a. We define the polynomial $p_k(t) = t^{-(k-m-m')L}F(t)$ which has degree at most $(k+m'+m-(k-m-m'))L \leq (2m+2m')L$ and hence

$$p_k(t) = \sum_{0\leq j\leq (2m+2m')L} a_j t^j \quad \text{and} \quad a_j = \sum_i c_{ij}\, p_k(t_i)$$

hold for mutually different points $a \leq t_1, \ldots, t_d \leq b$, where $d \leq (2m+2m')L+1$ are independent of k. Hence we get

$$\begin{aligned} F_\sigma(q) &= q^{(k-m-m')L} p_k(q) = q^{(k-m-m')L} \sum_{j,i} c_{ij} q^j\, p_k(t_i) \\ &= q^{(k-m-m')L} \sum_{j,i} c_{ij} q^j t_i^{-(k-m-m')L} F_\sigma(t_i) = \sum_i \Bigl(\sum_j c_{ij} q^j\Bigr)\Bigl(\frac{q}{t_i}\Bigr)^{(k-m-m')L} F_\sigma(t_i). \end{aligned}$$

This defines the coefficients γ_i. For $k \leq 2(m+m')$ we work directly with the polynomial $F(t)$ of degree at most $2(m+m')L$. □

Let $M = \Gamma_q(B, S\otimes H)$ and $\widetilde{M} = \Gamma_q(B, S\otimes(H\oplus H))$. Define the M-M bimodule $\mathcal{F}_m \subset L^2(\widetilde{M})$ as the $\|\cdot\|_2$-closed linear span of the reduced Wick words $W_\sigma(x_1,\ldots,x_t,h_1,\ldots,h_N)$, $N \geq 1$, such that $h_i \in H\times\{0\}\cup\{0\}\times H$ for all i and at least m of the vectors h_i belong to $\{0\}\times H$. This bimodule will play a crucial role in our deformation-rigidity arguments in the next section.

Theorem 6.9. *Let* $M = \Gamma_q(B, S\otimes H)$ *and let* $C > 0$, $d > 0$ *be two constants such that the dimension of* $L^2_k(M)$ *over* B *is smaller than* Cd^k *for all* k. *Let* $|q| < 1$. *Then there exists an* $L_0 \in \mathbb{N}$ *and a* B-M *bimodule* $\mathcal{K}$ *such that* $\mathcal{F}_l$ *is weakly contained in* $L_2(M)\otimes_B \mathcal{K}$ *for all* $l \geq L_0$.

Proof. Let us recall that

$$\langle \zeta\otimes(\zeta')^{\mathrm{op}}(a\otimes_B b), \alpha\otimes_B \beta\rangle = \tau(\beta^* E_B(\alpha^*\zeta a) b\zeta') = \tau(E_B(\alpha^*\zeta a)E_B(b\zeta'\beta^*)).$$

Now we may assume that $(\xi_i)_{i\in I_k}$ is a basis of dimension d_k over B so that

$$P_k(\zeta) = \sum_{i\in I_k} \xi_i E_B(\xi_i^*\zeta) \quad \text{and} \quad E_B(\xi_i^*\xi) \leq 1.$$

This implies

$$\begin{aligned} \tau(\zeta'\xi\zeta\eta) &= \sum_{i\in I_k} \tau(\zeta'\xi\xi_i E_B(\xi_i^*\zeta)\eta) = \sum_{i\in I_k} \tau(\xi_i E_B(\xi_i^*\zeta)\eta\zeta') \\ &= \sum_{i\in I_k} \langle \zeta\otimes(\zeta')^{\mathrm{op}}(1\otimes_B \eta), \xi_i\otimes_B (\xi\xi_i)^*\rangle. \end{aligned}$$

Let $q_0 < 1$ so that $q/q_0 < 1$. Then we define the B-M bimodule

$$\mathcal{K} = \bigoplus_{q/q_0 \le t < 1} L^2(\Gamma_{Q(t)}(B, S \otimes H))$$

with the natural left and right actions. For fixed ξ, η we choose $a = \pm q$ and $|q|/q_0 \le t_0 < \cdots < t_D < b$ for some $b < 1$. This allows us to define $W_{Q(t_i)}(\vec{h}, \vec{a})$ and $W_{Q(t_i)}(\vec{h}', \vec{a}')$ in $\mathcal{K}$. With the help of Corollary 6.8 we deduce that the map $\Phi^+(\zeta) = \sum_{k \ge 2(m+m')} E_M(\xi P_k(\zeta)\eta)$ satisfies

$$\|\Phi^+\|_{L_2(M)\otimes_B \mathcal{K}} \le \sum_\sigma \sum_l |\gamma_l| \sum_{k \ge 2(m+m')} q_0^{(k-m-m')L} \sum_{i \in I_k} \|1 \otimes_B W_{Q(t_l)}(\vec{H}', \vec{a}')\| \, \|\xi_i \otimes_B (W_{Q(t_l)}(\vec{H}, \vec{a})\xi_i)^*\|.$$

Now we note that

$$\|1 \otimes_B W_{Q(t_l)}(\vec{H}', \vec{a}')\| = \|W_{Q(t_l)}(\vec{H}', \vec{a}')\|_{L_2(\Gamma_{Q(t_l)})} \le c(t_l)$$

and

$$\begin{aligned}\|\xi_i \otimes_B (W_{Q(t_l)}(\vec{H}, \vec{a})\xi_i)^*\| &= \tau(W_{Q(t_l)}(\vec{H}, \vec{a})\xi_i E_B(\xi_i^*\xi_i)(W_{Q(t_l)}(\vec{H}, \vec{a})\xi_i)^*) \\ &\le \tau(\xi_i\xi_i^* W_{Q(t_l)}(\vec{H}, \vec{a})^* W_{Q(t_l)}(\vec{H}, \vec{a})) \\ &\le \|W_{Q(t_l)}(\vec{H}, \vec{a})\|^2_{\Gamma_{Q(t_l)}} \le c(t_l).\end{aligned}$$

Thus it suffices to know that $\sum_k q_0^{(k-m-m')L} C d^k$ is finite. Note here that m and m' depend on the Wick word and that we may assume $l \ge L_0$. Thus $q_0^{L_0} d < 1$ and $b < 1$ is enough to achieve summability. Using the second part of Corollary 6.8 we also have summability for $k \le 2(m+m')$. Lemma 6.1 then yields the assertion. □

Corollary 6.10. *Let $M = \Gamma_q(B, S \otimes H)$ and assume that H is finite-dimensional and* $\dim_B(D_k(S)) \le C d^k$ *for some constants $C, d > 0$. Then there exists an B-M bimodule $\mathcal{K}$ such that for $m \ge 1$ large enough we have $\mathcal{F}_m \prec L^2(M) \otimes_B \mathcal{K}$. In particular, for m large enough, $\mathcal{F}_m$ is weakly contained into $L^2(M) \otimes_B L^2(M)$.*

7. The proof of the main theorem and its applications

We first need some preliminaries. Throughout this section we use the notation $M = M(H) = \Gamma_q(B, S \otimes H)$ and $\widetilde{M} = \Gamma_q(B, S \otimes (H \oplus H)) = M(H \oplus H)$. Let

$$\mathcal{M} = (D \mathbin{\bar{\otimes}} \Gamma_q(\ell^2 \otimes H)) \vee M \subset (D \mathbin{\bar{\otimes}} \Gamma_q(\ell^2 \otimes H))^\omega.$$

As in Section 5, let

$$\mathcal{H} \subset \big((L^2(\mathcal{M}) \otimes_{\mathcal{A}} L^2(P)) \otimes \mathcal{F}_q(\ell^2 \otimes H)\big)^\omega$$

be the norm-closed linear span of the sequences

$$\Big(n^{-\frac{m}{2}} \sum_{(j_1,\ldots,j_m)=\sigma} (\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m) y \otimes_{\mathcal{A}} z) \otimes s_{j_1}(h_1)\cdots s_{j_m}(h_m)\Big)$$

for $m \geq 1,\ \sigma \in P_{1,2}(m),\ x_i \in BSB,\ h_i \in H,\ y \in M,\ z \in P$. Take the representations

$$\pi : M \to B(\mathcal{H}), \quad \theta : P^{\mathrm{op}} \to B(\mathcal{H})$$

introduced in Section 5 and define $\mathcal{N} = \pi(M) \vee \theta(P^{\mathrm{op}}) \subset B(\mathcal{H})$. As seen in Section 5, we choose a normal faithful state ϕ on $B_P = \pi(B) \vee \theta(P^{\mathrm{op}}) \subset B(\mathcal{H})$ and then define a normal faithful state ψ on $\mathcal{N}$ by $\psi = \phi \circ E_{B_P}$, where $E_{B_P} : \mathcal{N} \to B_P$ is the normal faithful conditional expectation. Let $d_\psi \in L^1(\mathcal{N})$ be the density of ψ and $\xi_0 = d_\psi^{\frac{1}{2}}$. Then $L^2(\mathcal{N})$ is the norm-closed span of the elements $\pi(x_\sigma)\theta(y^{\mathrm{op}})\xi_0$ for $x_\sigma \in M$ a Wick word and $y \in P$. Let $W_k(M)$ be the linear span of the Wick words of degree k in M and $L^2(B_P) = L^2(B_P, \phi)$ be the standard form for $B_P \subset B(\mathcal{H})$. Then $\mathcal{N}$ is standardly represented on

$$L^2(\mathcal{N}) = \overline{\bigoplus_{k \geq 0} W_k(M) L^2(B_P)}$$

by the formulas

$$\pi(x_\sigma)\theta(y^{\mathrm{op}})(\pi(x_\nu)\theta(z^{\mathrm{op}})\xi_0) = \pi(x_\sigma x_\nu)\theta((zy)^{\mathrm{op}})\xi_0, \quad x_\sigma, x_\nu \in M,\ y, z \in P.$$

The conjugation $\mathcal{J} : L^2(\mathcal{N}) \to L^2(\mathcal{N})$ associated to the standard representation of $\mathcal{N}$ is given by

$$\mathcal{J}(\pi(x_\sigma)\theta(y^{\mathrm{op}})\xi_0) = \sigma^{\psi}_{-\frac{i}{2}}(\pi(x_\sigma^*)\theta(\bar{y}))\xi_0, \quad x_\sigma \in M,\ y \in P,$$

where σ_t^ψ is the modular group on $\mathcal{N}$ associated to ψ. We will also consider $\widetilde{\mathcal{N}} = \mathcal{N}(\widetilde{H})$ constructed in the same way as $\mathcal{N}$ by using $\widetilde{H} = H \oplus H$ instead of H. Thus take

$$\widetilde{\mathcal{H}} \subset \big((L^2(\mathcal{M}) \otimes_{\mathcal{A}} L^2(P)) \otimes \mathcal{F}_q(\ell^2 \otimes \widetilde{H})\big)^\omega$$

to be the norm-closed linear span of the sequences

$$\left(n^{-\frac{m}{2}} \sum_{(j_1,\dots,j_m)=\sigma} (\pi_{j_1}(x_1)\cdots\pi_{j_m}(x_m) y \otimes_{\mathcal{A}} z) \otimes s_{j_1}(\tilde{h}_1)\cdots s_{j_m}(\tilde{h}_m)\right)$$

for $m \geq 1,\ \sigma \in P_{1,2}(m),\ x_i \in BSB,\ \tilde{h}_i \in \widetilde{H}$. Exactly as in Section 5, define the $*$-representations

$$\pi : \widetilde{M} \to B(\widetilde{\mathcal{H}}), \quad \theta : P^{\mathrm{op}} \to B(\widetilde{\mathcal{H}})$$

and then define $\widetilde{\mathcal{N}} = \pi(\widetilde{M}) \vee \theta(P^{\mathrm{op}})$. Then $\widetilde{\mathcal{N}}$ is standardly represented on

$$L^2(\widetilde{\mathcal{N}}) = \overline{\bigoplus_{k \geq 0} W_k(\widetilde{M}) L^2(B_P)},$$

and the associated conjugation $\widetilde{\mathcal{J}} : L^2(\widetilde{\mathcal{N}}) \to L^2(\widetilde{\mathcal{N}})$ is given by the formula

$$\widetilde{\mathcal{J}}(\pi(x_\sigma)\theta(y^{\mathrm{op}})\xi_0) = \sigma^{\psi}_{-\frac{i}{2}}(\pi(x_\sigma^*)\theta(\bar{y}))\xi_0, \quad x_\sigma \in \widetilde{M},\ y \in P,$$

where σ_t^ψ is the modular automorphism group on $\widetilde{\mathcal{N}}$ associated to ψ. For every angle t define the unitary V_t on $L^2(\widetilde{\mathcal{N}})$ by

$$\pi(x_\sigma(x_1, \tilde{h}_1, \dots, x_m, \tilde{h}_m))\theta(y^{\mathrm{op}})\xi_0 \mapsto \pi(x_\sigma(x_1, o_t(\tilde{h}_1), \dots, x_m, o_t(\tilde{h}_m)))\theta(y^{\mathrm{op}})\xi_0.$$

Then the one-parameter group of $*$-automorphisms $\mathrm{Ad}(V_t)$ of $B(L^2(\widetilde{\mathcal{N}}))$ restricts to a group α_t of $*$-automorphisms of $\widetilde{\mathcal{N}}$, acting according to the formula

$$\alpha_t(\pi(x_\sigma(x_1,\tilde{h}_1,\dots,x_m,\tilde{h}_m))\theta(y^{\mathrm{op}})) = \pi(x_\sigma(x_1,o_t(\tilde{h}_1),\dots,x_m,o_t(\tilde{h}_m)))\theta(y^{\mathrm{op}}).$$

When further restricted to $\widetilde{M} = \Gamma_q(B, S \otimes \widetilde{H})$ this group of $*$-automorphisms coincides with the one introduced in Theorem 3.16 and we have the identity

$$T_t(x) = E_M(\alpha_s(x)), \quad x \in M,\ 0 \le s < \tfrac{\pi}{2},$$

where T_t is the heat semigroup introduced in Theorem 3.16 and $t = -\ln(\cos(s))$. We finally introduce the bimodules needed in the deformation argument. To do this, recall that $\widetilde{M} = M(H \oplus H)$ is the generalized q-gaussian algebra generated by B, $s_q(a,h,0)$ and $s_q(a,0,h)$, where $a \in S$ runs through the generating set and $h \in H$ are unit vectors. Let $F \subset H$ be an orthonormal basis. Then we define an M-M bimodule $\mathcal{F}_{=m} \subset L^2(\widetilde{M})$ by

$$\mathcal{F}_{=m} = \overline{\mathrm{span}}^{\|\cdot\|_2}\big\{W_\sigma(k_1,\dots,k_N,a_1,\dots,a_{N'}) : k_i \in F\times\{0\}\cup\{0\}\times F,\ \#\{i \mid k_i \in \{0\}\times F\} = m\big\}.$$

Note that we use reduced Wick words. This means $N = |\sigma_s|$ and the vectors $(k_1,\dots,k_N)$ are the ones obtained after contracting the pairs. Here $\sigma \in P_{1,2}(N')$ and $a_1,\dots,a_{N'}$ are the original coefficients from S. One can see that $\mathcal{F}_{=m}$ is exactly the eigenspace of vectors $\xi \in L^2(\widetilde{M})$ such that $E_{M(0\oplus H)}(\alpha_t(\xi)) = e^{-tm}\xi$ for all (some) $t > 0$. Likewise we define the M-M bimodule $\mathcal{F}^P_{=m} \subset L^2(\widetilde{\mathcal{N}})$ as the $\|\cdot\|_2$-closed span of the elements

$$\pi(W_\sigma(x_1,h_1,\dots,x_m,h_m))\theta(y^{\mathrm{op}})\xi_0, \quad x_i \in BSB,\ h_i \in F\times\{0\}\cup\{0\}\times F,$$

such that exactly m of the vectors h_i belong to $\{0\}\times F$. It's easy to see that $\mathcal{F}^P_{=m}$ can be described by

$$\mathcal{F}^P_{=m} = \{\xi \in L_2(\widetilde{\mathcal{N}}) : E_{\mathcal{N}(0\oplus H)}(\alpha_t(\xi)) = e^{-tm}\xi \text{ for all } t > 0\}.$$

Finally, we set

$$\mathcal{F}_m = \bigoplus_{m' \ge m} \mathcal{F}_{=m'} \subset L^2(\widetilde{M}), \quad \mathcal{F}^P_m = \bigoplus_{m' \ge m} \mathcal{F}^P_{=m'} \subset L^2(\widetilde{\mathcal{N}}).$$

Let's remark that we have the following transversality property, whose proof is virtually the same as that of Proposition 5.1 in [Avsec 2012].

Lemma 7.1. *There exists a constant $C = C(m) > 0$ such that for $0 < t < 2^{-m-1}$ we have*

$$\|V_{t^{m+1}}(\xi) - \xi\| \le C\|e^{\perp}V_t(\xi)\| \quad \textit{for all } \xi \in \bigoplus_{k \ge m+1} L^2_k(\mathcal{N}) \subset L^2(\widetilde{\mathcal{N}}).$$

Theorem 7.2. *Let $M = \Gamma_q(B, S \otimes H)$ associated to a sequence of symmetric independent copies (π_j, B, A, D) and assume that the dimension of $D_k(S)$ over B is finite for every k and that H is finite-dimensional. Let $\mathcal{A} \subset M$ be a von Neumann subalgebra which is amenable relative to B and define $P = \mathcal{N}_M(\mathcal{A})''$. Let $m \ge 1$ be fixed. Then at least one of the following statements holds*:

(1) *The M-M bimodule $\mathcal{F}_m$ is left P-amenable.*

(2) *There exist $t, \delta > 0$ such that* $\inf_{a \in \mathcal{U}(\mathcal{A})} \|T_t(a)\|_2 \ge \delta$.

Proof. The approximately invariant states $\omega_n \in \mathcal{N}_*$ constructed in Theorem 5.1 are implemented by unit vectors $\xi_n \in L^2(\mathcal{N}) \subset L^2(\widetilde{\mathcal{N}})$. Using the Powers–Stormer inequalities we see that the vectors ξ_n have the following properties:

(1) $\langle \pi(x)\xi_n, \xi_n\rangle \to \tau(x),\ x \in M$.

(2) $\|\pi(a)\theta(\bar{a})\xi_n - \xi_n\| \to 0,\ a \in \mathcal{U}(\mathcal{A})$.

(3) $\|\pi(u)\theta(\bar{u})\mathcal{J}\pi(u)\theta(\bar{u})\mathcal{J}\xi_n - \xi_n\| \to 0,\ u \in \mathcal{N}_M(\mathcal{A})$.

Let $e^{\perp} : L^2(\widetilde{\mathcal{N}}) \to \mathcal{F}_m^P$ be the orthogonal projection. We have the following alternative:

Case 1. For every nonzero projection $p \in \mathcal{Z}(P)$ and for every $t > 0$ we have

$$\limsup_n \|e^{\perp} V_t \pi(p)\xi_n\| > \frac{\|p\|_2}{8C}.$$

We will prove that in this case the M-M bimodule $\mathcal{F}_m$ is left P-amenable.

Lemma 7.3. *Let X be the strong operator topology completion of $\mathcal{F}_m$ as a right M-module with respect to the M-valued inner product $\langle x, y\rangle = E_M(x^*y),\ x, y \in \mathcal{F}_m$. Let $\mathcal{L}(X)$ be the von Neumann algebra of adjointable operators on X. Then there exists a normal $*$-homomorphism $\Psi : \mathcal{L}(X) \to B(L^2(\mathcal{F}_m^P))$ such that $\Psi(\mathcal{L}(X)) \subset B(L^2(\mathcal{F}_m^P)) \cap (\mathcal{N}^{\mathrm{op}})' \cap (\theta(P^{\mathrm{op}}))'$.*

Proof. The condition 5.13(iii) implies that $\mathcal{F}_m^P = X \otimes_M L^2(\mathcal{N})$, where the left action on $\mathcal{N}$ is that of $\pi(M)$. Therefore the map $\Psi : \mathcal{L}(X) \to B(\mathcal{F}_m^P)$ given by

$$T \mapsto T \otimes_M \mathrm{id}$$

is a well-defined normal $*$-homomorphism. Let us consider a rank-1 operator $\xi \otimes \bar{\eta} \in \mathcal{L}(X)$ with ξ, η Wick words in $\widetilde{M}$. Then we calculate

$$\Psi(\xi \otimes \bar{\eta})(\pi(x_\sigma)\theta(y^{\mathrm{op}})\xi_0) = \pi(\xi)\pi(E_M(\eta^* x_\sigma))\theta(y^{\mathrm{op}})\xi_0.$$

Let $e_{\mathcal{N}}$ be the orthogonal projection of $\widetilde{\mathcal{N}}$ onto the closure of $\mathcal{N}\xi_0$, which exists thanks to the fact that $E^{\widetilde{\mathcal{N}}}_{B_P}$ is faithful, see Remark 5.14. Then we note that for $\tilde{x}_{\tilde{\sigma}} \in M$ we have

$$\begin{aligned}\langle \pi(E_M(\eta^* x_\sigma))\theta(y^{\mathrm{op}})\xi_0, \tilde{x}_{\tilde{\sigma}}\tilde{y}^{\mathrm{op}}\xi_0\rangle &= \psi(\theta(\tilde{y}^{\mathrm{op}})^*\theta(y^{\mathrm{op}})E^{\widetilde{\mathcal{N}}}_{B_P}(\tilde{x}^*_{\tilde{\sigma}}x_\sigma)) \\ &= \psi(\theta(\tilde{y}^{\mathrm{op}})^*\theta(y^{\mathrm{op}})(E^{\pi(M)}_{\pi(B)} \circ E^{\widetilde{\mathcal{N}}}_{\pi(M)})(\tilde{x}^*_{\tilde{\sigma}}x_\sigma)) \\ &= \langle \pi(E_M(\eta^* x_\sigma))\theta(y^{\mathrm{op}})\xi_0, \pi(\tilde{x}_{\tilde{\sigma}})\theta(\tilde{y}^{\mathrm{op}})\xi_0\rangle.\end{aligned}$$

This shows that

$$\pi(E_M(\eta^* x_\sigma))\theta(y^{\mathrm{op}})\xi_0 = e_{\mathcal{N}}(\pi(\eta^* x_\sigma)\theta(y^{\mathrm{op}})\xi_0).$$

Thus we deduce that for all the rank-1 operators $\xi \otimes \bar{\eta} \in \mathcal{L}(X)$

$$\Psi(\xi \otimes \bar{\eta}) = L_{\pi(\xi)}e_{\mathcal{N}}L_{\pi(\eta^*)}$$

is a right $\mathcal{N}$-module map, and hence belongs to $B(L^2(\mathcal{F}_m^P)) \cap (\mathcal{N}^{\mathrm{op}})'$. It's also trivial to check that $\Psi(\xi \otimes \bar{\eta})$ commutes with the operators $L_{\theta(y^{\mathrm{op}})}$ for all $y \in P$. Since Ψ is normal and the linear span of the rank-1 operators is so-dense in $\mathcal{L}(X)$, we have $\Psi(\mathcal{L}(X)) \subset B(L^2(\mathcal{F}_m^P)) \cap (\mathcal{N}^{\mathrm{op}})' \cap (\theta(P^{\mathrm{op}}))'$, as desired. □

The lemma provides a normal $*$-homomorphism

$$\Psi : B(\mathcal{F}_m) \cap (M^{\mathrm{op}})' \to B(\mathcal{F}_m^P) \cap (\theta(P^{\mathrm{op}}) \vee \widetilde{\mathcal{J}}\pi(M)\widetilde{\mathcal{J}} \vee \widetilde{\mathcal{J}}\theta(P^{\mathrm{op}})\widetilde{\mathcal{J}})'$$

such that $\Psi(\lambda(x)) = \pi(x)$ for $x \in M$, where λ is the natural left action of M on $L^2(\widetilde{M})$. From this point on, the proof proceeds verbatim as in [Popa and Vaes 2014a, proof of Case 1 in Theorem 3.1].

Case 2. There exist a nonzero central projection $p \in \mathcal{Z}(P)$ and $t > 0$ such that

$$\limsup_n \|e^{\perp} V_t \pi(p)\xi_n\| \leq \frac{\|p\|_2}{8C}.$$

In this case we prove that there exist $s, \delta > 0$ such $\|T_s(a)\|_2 \geq \delta$ for all $a \in \mathcal{U}(\mathcal{A})$. Write $\pi(p)\xi_n = \zeta_n + \eta_n$, where $\zeta_n \in \bigoplus_{k \leq m} L_k^2(\mathcal{N})$, $\eta_n \in \bigoplus_{k \geq m+1} L_k^2(\mathcal{N})$. Note that $\|\zeta_n\| \leq 1$, $\|\eta_n\| \leq 1$. Since V_t converges uniformly on $\big(\bigoplus_{k \leq m} L_k^2(\mathcal{N})\big)_1$, there exists a $t_0 > 0$ such that for $0 < s < t_0$ we have

$$\|V_s\xi - \xi\| \leq \min\left\{\frac{\|p\|_2}{8}, \frac{\|p\|_2}{8C}\right\} \quad \text{for } \xi \in \left(\bigoplus_{k \leq m} L_k^2(\mathcal{N})\right)_1.$$

Fix $0 < s < \min\{t^{m+1}, t_0, t_0^{m+1}, 2^{-(m+1)^2}\}$. For every $n \geq 1$ we have the estimate

$$\begin{aligned}
\|V_s\pi(p)\xi_n - \pi(p)\xi_n\| &\leq \|V_s\zeta_n - \zeta_n\| + \|V_s\eta_n - \eta_n\| \\
&\leq \frac{\|p\|_2}{8} + \|V_s\eta_n - \eta_n\| \leq \frac{\|p\|_2}{8} + C\|e^{\perp}V_{{}^{m+1}\!\sqrt{s}}\eta_n\| \\
&\leq \frac{\|p\|_2}{8} + C\|e^{\perp}V_{{}^{m+1}\!\sqrt{s}}\pi(p)\xi_n\| + C\|e^{\perp}V_{{}^{m+1}\!\sqrt{s}}\zeta_n\| \\
&\leq \frac{\|p\|_2}{8} + C\|e^{\perp}V_{{}^{m+1}\!\sqrt{s}}\pi(p)\xi_n\| + C\|e^{\perp}(V_{{}^{m+1}\!\sqrt{s}}\zeta_n - \zeta_n)\| + C\|e^{\perp}\zeta_n\| \\
&\leq \frac{\|p\|_2}{8} + C\|e^{\perp}V_{{}^{m+1}\!\sqrt{s}}\pi(p)\xi_n\| + C\|V_{{}^{m+1}\!\sqrt{s}}\zeta_n - \zeta_n\| \\
&\leq \frac{\|p\|_2}{4} + C\|e^{\perp}V_t\pi(p)\xi_n\|.
\end{aligned}$$

Taking the limsup with respect to n we obtain

$$\limsup_n \|V_s\pi(p)\xi_n - \pi(p)\xi_n\| \leq \frac{3\|p\|_2}{8}.$$

From this point on, the proof proceeds verbatim as in [Popa and Vaes 2014a, proof of Case 2 in Theorem 3.1]. □

Proof of Theorem A. For the first alternative, we use the second item in Theorem 7.2, Proposition 3.20 and Proposition 2.3. For the second alternative, we use the first item in Theorem 7.2, Corollary 6.10 and Remark 2.7. □

Proof of Corollaries B, C, D. These follow immediately from Theorem A. □

References

[Avsec 2012] S. T. Avsec, *Gaussian-like von Neumann algebras and noncommutative Brownian motion*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 2012, available at https://search.proquest.com/docview/1448260344.

[Biane 1997] P. Biane, "Free hypercontractivity", *Comm. Math. Phys.* **184**:2 (1997), 457–474. MR Zbl

[Boutonnet, Houdayer and Vaes 2018] R. Boutonnet, C. Houdayer, and S. Vaes, "Strong solidity of free Araki–Woods factors", *Amer. J. Math.* **140**:5 (2018), 1231–1252. MR Zbl

[Bożejko 1991] M. Bożejko, "A q-deformed probability, Nelson's inequality and central limit theorems", pp. 312–335 in *Nonlinear fields: classical, random, semiclassical* (Karpacz, Poland, 1991), edited by P. Garbaczewski and Z. Popowicz, World Sci., River Edge, NJ, 1991. MR

[Bożejko and Speicher 1991] M. Bożejko and R. Speicher, "An example of a generalized Brownian motion", *Comm. Math. Phys.* **137**:3 (1991), 519–531. MR Zbl

[Bożejko and Speicher 1992] M. Bożejko and R. Speicher, "An example of a generalized Brownian motion, II", pp. 67–77 in *Quantum probability and related topics*, edited by L. Accardi, QP-PQ **7**, World Sci., River Edge, NJ, 1992. MR Zbl

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

[Bożejko and Speicher 1996] M. Bożejko and R. Speicher, "Interpolations between bosonic and fermionic relations given by generalized Brownian motions", *Math. Z.* **222**:1 (1996), 135–159. MR Zbl

[Bożejko, Kümmerer and Speicher 1997] M. Bożejko, B. Kümmerer, and R. Speicher, "q-Gaussian processes: non-commutative and classical aspects", *Comm. Math. Phys.* **185**:1 (1997), 129–154. MR Zbl

[Caspers 2018] M. Caspers, "Gradient forms and strong solidity of free quantum groups", preprint, 2018. arXiv

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

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

[Connes and Jones 1982] A. Connes and V. Jones, "A II_1 factor with two nonconjugate Cartan subalgebras", *Bull. Amer. Math. Soc.* (*N.S.*) **6**:2 (1982), 211–212. MR Zbl

[Connes, Feldman and Weiss 1981] A. Connes, J. Feldman, and B. Weiss, "An amenable equivalence relation is generated by a single transformation", *Ergodic Theory Dynam. Systems* **1**:4 (1981), 431–450. MR Zbl

[Dabrowski 2014] Y. Dabrowski, "A free stochastic partial differential equation", *Ann. Inst. Henri Poincaré Probab. Stat.* **50**:4 (2014), 1404–1455. MR Zbl

[Dykema and Nica 1993] K. Dykema and A. Nica, "On the Fock representation of the q-commutation relations", *J. Reine Angew. Math.* **440** (1993), 201–212. MR Zbl

[Haagerup 1975] U. Haagerup, "The standard form of von Neumann algebras", *Math. Scand.* **37**:2 (1975), 271–283. MR Zbl

[Haagerup 1979] U. Haagerup, "L^p-spaces associated with an arbitrary von Neumann algebra", pp. 175–184 in *Algèbres d'opérateurs et leurs applications en physique mathématique* (Marseille, 1977), edited by M. A. Connes et al., Colloq. Internat. CNRS **274**, CNRS, Paris, 1979. MR Zbl

[Haagerup, Junge and Xu 2010] U. Haagerup, M. Junge, and Q. Xu, "A reduction method for noncommutative L_p-spaces and applications", *Trans. Amer. Math. Soc.* **362**:4 (2010), 2125–2165. MR Zbl

[Houdayer and Vaes 2013] C. Houdayer and S. Vaes, "Type III factors with unique Cartan decomposition", *J. Math. Pures Appl.* (9) **100**:4 (2013), 564–590. MR Zbl

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

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

[Isono 2015b] Y. Isono, "On bi-exactness of discrete quantum groups", *Int. Math. Res. Not.* **2015**:11 (2015), 3619–3650. MR Zbl

[Jones and Popa 1982] V. Jones and S. Popa, "Some properties of MASA's [sic] in factors", pp. 89–102 in *Invariant subspaces and other topics* (Timişoara/Herculane, Romania, 1981), edited by C. Apostol et al., Operator Theory: Adv. Appl. **6**, Birkhäuser, Basel, 1982. MR Zbl

[Junge 2006] M. Junge, "Operator spaces and Araki–Woods factors: a quantum probabilistic approach", *Int. Math. Res. Pap.* **2006** (2006), art. id. 76978. MR Zbl

[Junge and Sherman 2005] M. Junge and D. Sherman, "Noncommutative L^p modules", *J. Operator Theory* **53**:1 (2005), 3–34. MR Zbl

[Junge and Zeng 2015] M. Junge and Q. Zeng, "Ultraproduct methods for mixed q-Gaussian algebras", preprint, 2015. arXiv

[Junge, Longfield and Udrea 2014] M. Junge, S. Longfield, and B. Udrea, "Some classification results for generalized q-Gaussian algebras", 2014. To appear in *Doc. Math.* arXiv

[Junge, Palazuelos, Parcet, Perrin and Ricard 2015] M. Junge, C. Palazuelos, J. Parcet, M. Perrin, and É. Ricard, "Hypercontractivity for free products", *Ann. Sci. Éc. Norm. Sup.* (4) **48**:4 (2015), 861–889. MR Zbl

[Królak 2000] I. Królak, "Wick product for commutation relations connected with Yang–Baxter operators and new constructions of factors", *Comm. Math. Phys.* **210**:3 (2000), 685–701. MR Zbl

[Królak 2006] I. Królak, "Factoriality of von Neumann algebras connected with general commutation relations: finite dimensional case", pp. 277–284 in *Quantum probability* (Będlewo, Poland, 2004), edited by M. Bożejko et al., Banach Center Publ. **73**, Polish Acad. Sci. Inst. Math., Warsaw, 2006. MR Zbl

[Nou 2004] A. Nou, "Non injectivity of the q-deformed von Neumann algebra", *Math. Ann.* **330**:1 (2004), 17–38. MR Zbl

[Nou 2006] A. Nou, "Asymptotic matricial models and QWEP property for q-Araki–Woods algebras", *J. Funct. Anal.* **232**:2 (2006), 295–327. MR Zbl

[Ozawa 2004] N. Ozawa, "Solid von Neumann algebras", *Acta Math.* **192**:1 (2004), 111–117. MR Zbl

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

[Ozawa and Popa 2010b] N. Ozawa and S. Popa, "On a class of II_1 factors with at most one Cartan subalgebra, II", *Amer. J. Math.* **132**:3 (2010), 841–866. MR Zbl

[Paschke 1973] W. L. Paschke, "Inner product modules over B^*-algebras", *Trans. Amer. Math. Soc.* **182** (1973), 443–468. MR Zbl

[Paschke 1974] W. L. Paschke, "The double B-dual of an inner product module over a C^*-algebra B", *Canad. J. Math.* **26**:5 (1974), 1272–1280. MR Zbl

[Peterson and Sinclair 2012] J. Peterson and T. Sinclair, "On cocycle superrigidity for Gaussian actions", *Ergodic Theory Dynam. Systems* **32**:1 (2012), 249–272. MR Zbl

[Pisier 2000] G. Pisier, "An inequality for p-orthogonal sums in non-commutative L_p", *Illinois J. Math.* **44**:4 (2000), 901–923. MR Zbl

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

[Popa 2006a] S. Popa, "On a class of type II_1 factors with Betti numbers invariants", *Ann. of Math.* (2) **163**:3 (2006), 809–899. MR Zbl

[Popa 2006b] S. Popa, "Strong rigidity of II_1 factors arising from malleable actions of w-rigid groups, I", *Invent. Math.* **165**:2 (2006), 369–408. MR Zbl

[Popa and Vaes 2014a] S. Popa and S. Vaes, "Unique Cartan decomposition for II_1 factors arising from arbitrary actions of free groups", *Acta Math.* **212**:1 (2014), 141–198. MR Zbl

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

[Raynaud 2002] Y. Raynaud, "On ultrapowers of non commutative L_p spaces", *J. Operator Theory* **48**:1 (2002), 41–68. MR Zbl

[Ricard 2005] É. Ricard, "Factoriality of q-Gaussian von Neumann algebras", *Comm. Math. Phys.* **257**:3 (2005), 659–665. MR Zbl

[Shlyakhtenko 1999] D. Shlyakhtenko, "A-valued semicircular systems", *J. Funct. Anal.* **166**:1 (1999), 1–47. MR Zbl

[Shlyakhtenko 2004] D. Shlyakhtenko, "Some estimates for non-microstates free entropy dimension with applications to q-semicircular families", *Int. Math. Res. Not.* **2004**:51 (2004), 2757–2772. MR Zbl

[Shlyakhtenko 2009] D. Shlyakhtenko, "Lower estimates on microstates free entropy dimension", *Anal. PDE* **2**:2 (2009), 119–146. MR Zbl

[Sinclair 2011] T. Sinclair, "Strong solidity of group factors from lattices in SO$(n, 1)$ and SU$(n, 1)$", *J. Funct. Anal.* **260**:11 (2011), 3209–3221. MR Zbl

[Śniady 2001] P. Śniady, "Gaussian random matrix models for q-deformed Gaussian variables", *Comm. Math. Phys.* **216**:3 (2001), 515–537. MR Zbl

[Śniady 2004] P. Śniady, "Factoriality of Bożejko–Speicher von Neumann algebras", *Comm. Math. Phys.* **246**:3 (2004), 561–567. MR Zbl

[Speicher 1992] R. Speicher, "A noncommutative central limit theorem", *Math. Z.* **209**:1 (1992), 55–66. MR Zbl

[Speicher 1993] R. Speicher, "Generalized statistics of macroscopic fields", *Lett. Math. Phys.* **27**:2 (1993), 97–104. MR Zbl

[Terp 1982] M. Terp, "Interpolation spaces between a von Neumann algebra and its predual", *J. Operator Theory* **8**:2 (1982), 327–360. MR Zbl

[Voiculescu 1985] D. Voiculescu, "Symmetries of some reduced free product C^*-algebras", pp. 556–588 in *Operator algebras and their connections with topology and ergodic theory* (Buşteni, Romania, 1983), edited by H. Araki et al., Lecture Notes in Math. **1132**, Springer, 1985. MR Zbl

[Voiculescu 1991] D. Voiculescu, "Limit laws for random matrices and free products", *Invent. Math.* **104**:1 (1991), 201–220. MR Zbl

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

[Voiculescu, Dykema and Nica 1992] D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free random variables*, CRM Monograph Series **1**, Amer. Math. Soc., Providence, RI, 1992. MR Zbl

Received 10 Oct 2017. Revised 20 Jul 2018. Accepted 16 Sep 2018.

MARIUS JUNGE: junge@math.uiuc.edu
Department of Mathematics, University of Illinois, Urbana, IL, United States

BOGDAN UDREA: bogdanudrea75@yahoo.com
Institute of Mathematics of the Romanian Academy, Bucharest, Romania

ANALYSIS AND PDE
Vol. 12, No. 7, 2019
dx.doi.org/10.2140/apde.2019.12.1711

COMPLEX INTERPOLATION AND CALDERÓN–MITYAGIN COUPLES OF MORREY SPACES

MIECZYSŁAW MASTYŁO AND YOSHIHIRO SAWANO

We study interpolation spaces between global Morrey spaces and between local Morrey spaces. We prove that for a wide class of couples of these spaces the upper complex Calderón spaces are not described by the K-method of interpolation. A by-product of our results is that couples of Morrey spaces belonging to this class are not Calderón–Mityagin couples. A Banach couple (X_0, X_1) is said to have the universal K-property if all relative interpolation spaces from any Banach couple to (X_0, X_1) are relatively K-monotone. A couple of local Morrey spaces is proved to have the universal K-property once it is a Calderón–Mityagin couple.

1. Introduction

The theory of Calderón–Mityagin couples is a central topic in abstract interpolation theory, since the interpolation spaces relative to such couples are isomorphic to generalized real interpolation spaces. We are interested in unifying this collection of results on Calderón–Mityagin couples. This attempt at unification forms an important component in the general program of describing all interpolation spaces with respect to a given compatible couple of Banach spaces. There is a simple characterization of Calderón–Mityagin couples in terms of the so-called submajorization of the K-functional and orbits. Besides the fundamental example of the couple (L^1, L^∞), which was independently discovered by Calderón [1966] and Mityagin [1965], many other examples were found out later by many mathematicians in interpolation theory, like couples of weighted L^p or of certain types of rearrangement invariant spaces. Unfortunately, it is still difficult to prove or disprove that a given couple of Banach spaces is a Calderón–Mityagin couple. Nevertheless, many Calderón–Mityagin couples have been discovered; we refer, e.g., to [Cwikel and Nilsson 2003; Kalton 1993; Mastyło and Sinnamon 2017] for more about this topic. In this paper we handle couples of Morrey spaces and local Morrey spaces as examples and counterexamples of Calderón–Mityagin couples. Based on the results, we consider the interpolation of Morrey spaces.

Cwikel [1981] conjectured that all interpolation spaces with respect to a given Banach couple are described by K-method whenever all complex interpolation spaces have this property. However, in [Mastyło and Ovchinnikov 1997] the authors disproved this conjecture. This motivates us to study classes of Banach couples for which Cwikel's conjecture is true.

Mastyło was supported by the National Science Centre, Poland, project no. 2015/17/B/ST1/00064. Sawano was supported by Grant-in-Aid for Scientific Research (C) (16K05209), the Japan Society for the Promotion of Science.
MSC2010: primary 46B70; secondary 42A45, 42B30.
Keywords: interpolation spaces, complex method, K-method, Morrey spaces.

The main purpose of the present paper is to study Cwikel's conjecture for the class of Morrey spaces, which play an important role in nonlinear potential analysis and harmonic analysis; see [Adams and Xiao 2004; 2012]. These Banach spaces were used for the first time by Morrey [1938] to prove that certain systems of partial differential equations have Hölder continuous solutions. Morrey spaces are widely used to investigate the local behavior of solutions of partial differential equations, including the Navier–Stokes equations; see, e.g., [Lemarié-Rieusset 2012; Mazzucato 2003; Taylor 1997].

Before we state the main results of the present paper, we introduce some fundamental definitions. For $1 \le q \le p \le \infty$ the (global) Morrey space $\mathcal{M}_q^p$ over $\mathbb{R}^n$ is defined to be the space of all q-locally integrable functions f on $\mathbb{R}^n$ ($f \in L^q_{\mathrm{loc}}$ for short) such that

$$\|f\|_{\mathcal{M}_q^p} := \sup_{(x,r)\in\mathbb{R}^n\times\mathbb{R}_+} |B(x,r)|^{1/p-1/q} \left(\int_{B(x,r)} |f(y)|^q \, dy \right)^{1/q} < \infty.$$

Here and below we write $\mathbb{R}_+ = (0,\infty)$. The symbol $|A|$ stands for the Lebesgue measure of any Lebesgue measurable set A in $\mathbb{R}^n$, and $B(x,r)$ is the open ball in $\mathbb{R}^n$ centered at x of radius $r > 0$. In particular, by the Lebesgue differentiation theorem $\mathcal{M}_q^\infty = L^\infty$ with identical norms. For simplicity of notation, we abbreviate $B(0,r)$ to $B(r)$.

Note that for these spaces sometimes other symbols, such as $\mathcal{L}^{q,\lambda}$ [Peetre 1969] and $L^{q,\lambda}$ [Nakai 2008], are used. Apart from the choice of a different letter $\mathcal{L}$, the second parameter λ is also introduced into the norm in a way different from the above; namely for a measurable function f we define

$$\|f\|_{\mathcal{L}^{q,\lambda}} = \|f\|_{L^{q,\lambda}} = \sup_{x\in\mathbb{R}^n\times\mathbb{R}_+} \left(\frac{1}{r^\lambda} \int_{B(x,r)} |f(y)|^q \, dy \right)^{1/q},$$

where $1 \le q < \infty,\ 0 \le \lambda < n$. Among various function spaces above, we note the following relation:

$$\mathcal{M}_q^p = L^{q,\lambda} = \mathcal{L}^{q,\lambda}, \quad \lambda = n - \frac{nq}{p}.$$

We point out that for technical reasons it is convenient to use a norm equivalent to the original norm of the Morrey space $\mathcal{M}_q^p$ given by

$$\|f\|^*_{\mathcal{M}_q^p} = \sup_Q |Q|^{1/p-1/q} \left(\int_Q |f(y)|^q \, dy \right)^{1/q}, \quad f \in \mathcal{M}_q^p,$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$ with sides parallel to coordinate axes.

It seems worth investigating its local counterpart, which is related to the Beurling algebra B^q and Wiener spaces. The local version of Morrey spaces, where only balls centered at the origin are taken into account, has a connection with studies of N. Wiener [1930; 1932], who considered the spaces of all measurable functions f such that for given $q \in \{1,2\}$

$$\frac{1}{T} \int_0^T |f(y)|^q \, dy$$

is bounded in T or tends to 0 as $T \to \infty$.

In the multidimensional case, a variant of these spaces defined by the norm

$$\|f\|_{B^q} = \sup_{r>0}\left(\frac{1}{|B(r)|}\int_{B(r)} |f(y)|^q\,dy\right)^{1/q}$$

appeared in [Beurling 1964] as the dual of the so-called Beurling algebra.

A local variant of Morrey spaces appeared in [García-Cuerva and Herrero 1994]. The local Morrey space $L\mathcal{M}^p_q$ is defined to be the set of all $f \in L^q_{\mathrm{loc}}$ such that

$$\|f\|_{L\mathcal{M}^p_q} := \sup_{r>0} |B(r)|^{1/p-1/q}\left(\int_{B(r)} |f(y)|^q\,dy\right)^{1/q} < \infty.$$

We note that

$$\|f\|^*_{L\mathcal{M}^p_q} := \sup_{r>0} |Q(r)|^{1/p-1/q}\left(\int_{Q(r)} |f(y)|^q\,dy\right)^{1/q}$$

is an equivalent norm in $L\mathcal{M}^p_q$, where $Q(r) := [-r, r]^n$.

Interpolation properties of classical Morrey spaces were obtained in [Campanato and Murthy 1965; Peetre 1969; Stampacchia 1964]. More and more attention is now being paid to the interpolation of Morrey spaces due to certain properties of Morrey spaces that have become clear recently. For example, as the function $|x|^{-n/p}$ shows, $\mathcal{M}^p_q$ does not contain L^∞_{c} densely. Complex interpolation of Morrey spaces has been studied in [Lemarié-Rieusset 2012; 2013; Yuan et al. 2015]. We mention that Lemarié–Rieusset [2013, case (a), p. 750] proved that if

$$1 \le q_j < p_j < \infty,\quad j \in \{0,1\},\qquad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1},\qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}, \tag{1-1}$$

then for every $\theta \in (0,1)$,

$$[\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}]_\theta \ne \mathcal{M}^p_q$$

whenever $q_0/p_0 \ne q_1/p_1$. For the case when $q_0/p_0 = q_1/p_1$, Lemarié–Rieusset [2013] proved that

$$[\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}]^\theta = \mathcal{M}^p_q.$$

Here and below $[\,\cdot\,]_\theta$ and $[\,\cdot\,]^\theta$ denote the lower and the upper (Calderón) complex methods of interpolation defined in [Calderón 1964], respectively. Lemarié–Rieusset also studied real interpolation of Morrey spaces. In particular under conditions (1-1) we have

$$(L^{q_0}, L^{q_1})_{\theta,q} = L^q,$$

and hence

$$(\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1})_{\theta,q} \hookrightarrow \mathcal{M}^p_q$$

in the sense of continuous embedding. Meanwhile,

$$\mathcal{M}^p_q \hookrightarrow (\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1})_{\theta,\infty}$$

if and only if $q_0/p_0 = q_1/p_1$. We refer to [Lemarié-Rieusset 2013; Yuan et al. 2015] for more details.

In addition to Morrey spaces and local Morrey spaces, we will also consider in our paper the "weak" Morrey space $\mathrm{w}\mathcal{M}^p_q$ and the "weak" local Morrey space $\mathrm{w}L\mathcal{M}^p_q$. For given $1 \le q \le p < \infty$ the weak Morrey space $\mathrm{w}\mathcal{M}^p_q$ is defined to be the quasi-Banach space of all Lebesgue measurable functions f endowed with the quasinorm

$$\|f\|_{\mathrm{w}\mathcal{M}^p_q} = \sup_{\lambda>0} \lambda \|\chi_{\{|f|>\lambda\}}\|_{\mathcal{M}^p_q};$$

Nakai [2008] used the norm

$$\|f\|_{\mathrm{w}L^{q,\lambda}} = \sup_{\lambda>0} \lambda \|\chi_{\{|f|>\lambda\}}\|_{L^{q,\lambda}}$$

to define weak Morrey spaces, while the weak local Morrey space $\mathrm{w}L\mathcal{M}^p_q$ is defined to be the quasi-Banach space of all Lebesgue measurable functions f for which

$$\|f\|_{\mathrm{w}L\mathcal{M}^p_q} = \sup_{\lambda>0} \lambda \|\chi_{\{|f|>\lambda\}}\|_{L\mathcal{M}^p_q} < \infty.$$

When $\Phi(r) = r^q$ and $\phi(r) = r^{-1+\lambda/n}$ and $\lambda/n = 1 - q/p$, this space $\mathrm{w}\mathcal{M}^p_q$ is the same as Nakai's space $L^{(\Phi,\varphi)}_{\mathrm{weak}}$ [Nakai 2008, Definition 6.1, p. 207]. In particular, for every $s \in [1, \infty)$, $L^{(1,n-n/s)}_{\mathrm{weak}}$ coincides with $\mathrm{w}\mathcal{M}^s_1$.

We add some comments on difficulties related to interpolation of Morrey spaces. First we notice that until now there is no complete description of all complex or real interpolation spaces of all couples $(\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1})$ of Morrey spaces with $1 \le q_0 \le p_0 < \infty$ and $1 \le q_1 \le p_1 < \infty$. Regarding real interpolation, it should be pointed out that a general formula remains unknown, even within equivalence for the K-functional of these couples. It is apparently very difficult to find such a formula, and this indeed is one of the nontrivial difficulties of dealing with Morrey couples, especially in the description of interpolation spaces with respect these couples, and in particular real interpolation spaces which just involve the K-functional. Interestingly the situation is completely different in the setting of local Morrey spaces (see Section 5).

In this paper we provide a solution to Cwikel's conjecture in the Morrey space setting. Our new results show that Cwikel's conjecture is still valid for a wide class of global and local Morrey spaces. In particular, because of the fact that a wide class of these couples are not Calderón–Mityagin couples, we have to declare that the problem related to the description of all interpolation spaces for all couples of Morrey spaces is extremely difficult in general.

Let us now describe more precisely the contents of the present paper. In Section 2 we introduce some fundamental definitions and notation used in the paper. In Section 3 we study the upper complex method of interpolation $[\,\cdot\,]^\theta$ for any $\theta \in (0, 1)$. We prove that for any couple (X_0, X_1) of complex Banach lattices on an arbitrary measure space (Ω, Σ, μ) the Gagliardo completion of $[X_0, X_1]^\theta$ with respect to $X_0 + X_1$ coincides isometrically with the Gagliardo completion of the Calderón product $X_0^{1-\theta} X_1^\theta$ with respect to $X_0 + X_1$. In particular we show that if (Ω, Σ, μ) is a σ-finite measure space, then $[X_0, X_1]^\theta = X_0^{1-\theta} X_1^\theta$ with equality of norms whenever each of X_0 and X_1 has the Fatou property. Applying this result to Morrey spaces, we recover the results above due to [Lemarié-Rieusset 2013].

In Sections 4 and 5 we provide a general sufficient condition on a Banach couple (A_0, A_1), which guarantees that (A_0, A_1) is not a Calderón–Mityagin couple. As a by-product, we prove in Section 4 that both couples $(\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1})$ and $(\mathcal{M}^{p_0}_{q_0}, L^\infty)$ are not Calderón–Mityagin couples provided that $q_0 \neq q_1$, $p_0 \neq q_0$ and $p_1 \neq q_1$.

Finally, in Section 5, we describe real interpolation of local Morrey spaces by the upper complex method $[\,\cdot\,]^\theta$ and the classical real method $(\,\cdot\,)_{\theta,\infty}$ for all $\theta \in (0, 1)$. These results are applied to prove that $(L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{q_1}_{p_1})$ is a Calderón–Mityagin couple if and only if $q_0 = q_1$, and in this case this couple has the universal K-property, i.e., (A_0, A_1) and $(L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{q_1}_{p_1})$ are relative Calderón–Mityagin couples for any Banach couple (A_0, A_1). We stress that the key point here is that fact the inclusion $L\mathcal{M}^p_q \hookrightarrow \mathrm{w}L\mathcal{M}^p_q$ is proper for every $1 \leq q \leq p < \infty$; see Lemma 4.7(ii) for the proof, where we use the maximal Hardy–Littlewood operator.

We will use standard notation; in particular given two nonnegative functions f and g defined on the same set A, we write $f \prec g$ or $g \succ f$ if there is a constant $c > 0$ such that $f(x) \leq cg(x)$ for all $x \in A$, while $f \asymp g$ means that both $f \prec g$ and $g \prec f$ hold. If X and Y are topological linear spaces, then $X \hookrightarrow Y$ means that $X \subset Y$ and the inclusion map is continuous. In the case when X and Y are Banach spaces, we write $X \cong Y$ whenever $X = Y$, with *equality* of norms. Throughout the entire paper, C will denote a constant which may have a different value in different appearances.

2. Preliminaries

We will use the standard notation in the theory of Banach spaces and the theory of integration. If X is a Banach space, we denote its (closed) unit ball by B_X. For any measure space (Ω, Σ, μ), the space of all μ-equivalence classes of real-valued Σ-measurable μ-almost everywhere finite functions will be denoted by $L^0(\mu) := L^0(\Omega, \Sigma, \mu)$. This space is a vector lattice under the natural order: $f \leq g$ if and only if $f(s) \leq g(s)$ for μ-almost everywhere $s \in \Omega$.

A linear subspace E of $L^0(\Omega, \Sigma, \mu)$ is called an *ideal* if it is solid, i.e., $|f| \leq |g|$ for some $g \in E$ implies $f \in E$. We will consider Banach lattices on an arbitrary measure space (in general we do not need to assume that the measure is σ-finite, which is usually found in the literature). We recall that a Banach space $X \subset L^0(\mu)$, which is an ideal with a monotone norm (meaning $\|f\|_X \leq \|g\|_X$ for all $f, g \in X$ satisfying $|f| \leq |g|$) is said to be a *Banach lattice* on (Ω, Σ, μ). It is well known that in the theory of Banach lattices on measure spaces we may assume without loss of generality that the measure spaces are complete. A Banach lattice X is called *σ-order continuous* if $x_n \downarrow 0$ implies $\|x_n\|_X \to 0$.

We note that for all choices of two Banach lattices X_0 and X_1 on an arbitrary measure space, $\vec{X} := (X_0, X_1)$ forms a Banach couple in the sense of interpolation theory; see, e.g., [Krein et al. 1982, Corollary 1, p. 42] in the case of σ-finite measures, and for an arbitrary measure space, [Cwikel and Nilsson 2003, Remark 1.41, pp. 34–35].

Let (Ω, Σ, μ) be a measure space, and let $1 \leq p < \infty$. We recall that the *weak Lebesgue* or the *Marcinkiewicz space* $L^{p,\infty}(\mu)$ is made up of all functions $f \in L^0(\mu)$ such that

$$\|f\|_{p,\infty} := \sup_{\lambda>0} \lambda\, \mu(\{x \in \Omega : |f(x)| > \lambda\})^{1/p} < \infty.$$

If $p > 1$, then the quasinorm $\|\cdot\|_{p,\infty}$ is equivalent to the norm

$$\|f\|^*_{p,\infty} := \sup_{A\subset\Omega} \mu(A)^{1/p-1} \int_A |f|\, d\mu,$$

where the supremum is taken over all measurable subsets A of Ω with $\mu(A) > 0$.

If $X \subset L^0(\mu)$ is a Banach lattice and $p \in (1, \infty)$, then its *p-convexification* $X^{(p)}$ is the Banach lattice of all elements $f \in L^0(\mu)$ such that $|f|^p \in X$ with a norm $\|f\|_{X^{(p)}} = \||f|^p\|_X^{1/p}$.

Most of our notation and terminology from interpolation theory is standard and can be found in [Bergh and Löfström 1976; Brudnyi and Krugljak 1991]. For the reader's convenience, we recall some of them.

Let $\vec{X} = (X_0, X_1)$ and $\vec{Y} = (Y_0, Y_1)$ be Banach couples and let $\mathcal{L}(\vec{X}, \vec{Y})$ be the Banach space of all linear operators $T : \vec{X} \to \vec{Y}$ (meaning, as usual, that $T : X_0 + X_1 \to Y_0 + Y_1$ is linear and $T : X_j \to Y_j$ boundedly for $j = 0, 1$), where the norm is given by $\|T\|_{\vec{X}\to\vec{Y}} = \max_{j=0,1} \|T\|_{X_j\to Y_j}$.

Let X be an intermediate space with respect to a Banach couple $\vec{X} = (X_0, X_1)$. The *Gagliardo completion* or *relative completion* of X with respect to $\vec{X}$ is the Banach space X^{c} of all limits in $X_0 + X_1$ of sequences that are bounded in X and endowed with the norm $\|x\|_{X^{\mathrm{c}}} = \inf\{\sup_{k\geq 1} \|x_k\|_X\}$, where the infimum is taken over all bounded sequences $\{x_k\}_{k=1}^\infty$ in X whose limit in $X_0 + X_1$ equals x.

For every Banach couple $\vec{X} = (X_0, X_1)$ and $\theta \in (0, 1)$ the Peetre K-functional of $x \in X_0 + X_1$ is defined by

$$K(t, x; \vec{X}) = K(t, x; X_0, X_1) := \inf\{\|x_0\|_{X_0} + t\|x_1\|_{X_1} : x = x_0 + x_1\}, \quad t > 0,$$

and the real interpolation space $\vec{X}_{\theta,\infty}$ is defined to be the collection of all $x \in X_0 + X_1$ such that

$$\|x\|_{\theta,\infty} := \sup_{t>0} t^{-\theta} K(t, x; \vec{X}) < \infty.$$

Let X and Y be intermediate spaces with respect to $\vec{X}$ and $\vec{Y}$ respectively. We say that they are *relative interpolation spaces* with respect to $\vec{X}$ and $\vec{Y}$ if every $T \in \mathcal{L}(\vec{X}, \vec{Y})$ maps X into Y. The relative interpolation spaces X and Y are said to be *relative K-monotone spaces* with respect to $\vec{X}$ and $\vec{Y}$ if, whenever $x \in X$ and $y \in Y_0 + Y_1$ satisfy $K(t, y; \vec{Y}) \leq K(t, x; \vec{X})$ for all $t > 0$, it follows that $y \in Y$. If $\vec{X} = \vec{Y}$ and $X = Y$, then X is said to be a K-monotone space with respect to $\vec{X}$. Note that $K(t, Tx; \vec{Y}) \leq \|T\|_{\vec{X}\to\vec{Y}} K(t, x; \vec{X})$ for $x \in \vec{X}$. So, if X and Y are relative K-monotone, then they are relative interpolation spaces. If all relative interpolation spaces for $\vec{X}$ and $\vec{Y}$ are relative K-monotone, then we say that $\vec{X}$ and $\vec{Y}$ are relative Calderón–Mityagin couples. In particular, if also $\vec{X} = \vec{Y}$, then $\vec{X}$ is said to be a Calderón–Mityagin couple. We remark here that in a number of papers, various alternative terminologies, such as C-couple or K-adequate couple, are used for the notion of Calderón–Mityagin couples. It is well known and easy to prove, see, e.g., [Cwikel and Nilsson 2003, Remark 1.31], that $\vec{X}$ and $\vec{Y}$ are relative Calderón–Mityagin couples if and only if, for every $x \in X_0 + X_1$ and $y \in Y_0 + Y_1$, the inequality

$$K(t, y; \vec{Y}) \leq K(t, x; \vec{X}), \quad t > 0,$$

implies that there exists an operator $T : \vec{X} \to \vec{Y}$ such that $Tx = y$.

Let $\lambda \geq 1$ be a fixed constant. If we can arrange that $\|T\|_{\vec{X}\to\vec{Y}} \leq \lambda$ for all x and y above, then we say that $\vec{X}$ and $\vec{Y}$ are λ-relatively uniform Calderón–Mityagin couples. An interpolation couple $\vec{X}$ is said to be a uniform Calderón–Mityagin couple if $\vec{X}$ and $\vec{X}$ are λ-relatively uniform Calderón–Mityagin couples for some λ.

If Φ is a Banach lattice of Lebesgue measurable functions on $\mathbb{R}_+$ that contain the function $\min\{1, t\}$, then we can define the Banach space $(X_0, X_1)_\Phi$ of all $x \in X_0 + X_1$ such that $K(\cdot, x; \vec{X}) \in \Phi$ with the norm

$$\|x\| = \|K(\cdot, x; \vec{X})\|_\Phi.$$

The space $(X_0, X_1)_\Phi$ is called the K-space generated by Φ. It is a fundamental result of Brudnyi and Krugljak [1991, Theorem 3.3.20, p. 355] that if (X_0, X_1) is a uniform Calderón–Mityagin couple, then every interpolation space X with respect to (X_0, X_1) is (up to equivalence of norm) a K-space for some Φ. The key ingredient of this result is the K-divisibility theorem first proved by Brudnyi and Krugljak [1981; 1991] and later refined in [Cwikel 1984].

Let (Ω, Σ, μ) be a σ-finite measure space. It is well known that $(L^1(\mu), L^\infty(\mu))$ is a 1-uniform C-couple; see [Calderón 1966]. Several years later, Sedaev and Semenov [1971] proved that every weighted couple $(L^1(w_0), L^1(w_1))$ is a uniform C-couple. For more examples of uniform Calderón–Mityagin couples of Banach lattices we refer to [Cwikel 1981; Cwikel and Nilsson 2003; Kalton 1993; Mastyło and Sinnamon 2017].

3. Upper complex interpolation between Banach lattices

We will use complex methods of interpolation introduced in the fundamental paper [Calderón 1964] to prove isometric relationships between Banach lattices generated by standard operations applied to the Calderón product in the setting of couples of complex Banach lattices on an arbitrary measure space. We will apply these results to couples of Banach lattices enjoying the Fatou property and so in particular to Morrey spaces later.

Let $S := \{z \in \mathbb{C} : 0 < \operatorname{Re} z < 1\}$ be an open strip on the plane. For a given $\theta \in (0, 1)$ and any couple $\vec{X} = (X_0, X_1)$ we denote by $\mathcal{F}(\vec{X})$ the Banach space of all bounded continuous functions $f : \overline{S} \to X_0 + X_1$ on the closure $\overline{S}$ that are analytic on S, and $\mathbb{R} \ni t \mapsto f(j + it) \in X_j$ is a bounded continuous function, for each $j = 0, 1$, and endowed with the norm

$$\|f\|_{\mathcal{F}(\vec{X})} = \max_{j=0,1} \sup_{t\in\mathbb{R}} \|f(j + it)\|_{X_j}.$$

The lower/first complex interpolation space is defined by

$$[X_0, X_1]_\theta := \{f(\theta) : f \in \mathcal{F}(\vec{X})\}$$

and is endowed with the quotient norm. This definition is slightly different from those in [Bergh and Löfström 1976; Calderón 1964], however it gives the same interpolation spaces; see, e.g., [Krein et al. 1982]. We recall that in the original definition it is required in addition that $f \in \mathcal{F}(\vec{X})$ satisfies

$$\lim_{|t|\to\infty} \|f(j + it)\|_{X_j} = 0, \quad j \in \{0, 1\}.$$

Calderón defined a different interpolation method as follows. Let $\mathcal{G}(\vec{X})$ be the Banach space of all continuous functions $g : \overline{S} \to X_0 + X_1$ that are analytic on S, for which there exists $C = C(g) > 0$ such that $\|g(z)\|_{X_0+X_1} \le C(1+|z|)$ for all $z \in S$, and that are endowed with the norm

$$\|g\|_{\mathcal{G}(\vec{X})} := \max_{j=0,1} \left\{ \sup_{s \neq t} \frac{\|g(j+is) - g(j+it)\|_{X_j}}{|s-t|} \right\}.$$

The upper/second complex interpolation space is defined by

$$[X_0, X_1]^\theta := \{g'(\theta) : g \in \mathcal{G}(\vec{X})\}$$

and is endowed with the quotient norm.

Throughout the paper, when the complex methods are applied to a couple (X_0, X_1) of Banach lattices, we mean that $X_j := X_j(\mathcal{C})$ is a complexification of X_j for each $j = 0, 1$.

We need the following lemma:

Lemma 3.1. *Let $\vec{X} = (X_0, X_1)$ be a complex Banach couple, and let $\theta \in (0, 1)$:*

(i) $([X_0, X_1]^\theta)^c \cong ([X_0, X_1]_\theta)^{\mathrm{c}}$.

(ii) $[X_0, X_1]^\theta \cong ([X_0, X_1]_\theta)^{\mathrm{c}}$ *if and only if the unit ball of* $[X_0, X_1]^\theta$ *is closed in* $X_0 + X_1$.

Proof. We claim that $[X_0, X_1]^\theta \hookrightarrow ([X_0, X_1]_\theta)^{\mathrm{c}}$ with norm of continuous inclusion less than or equal to 1. Fix $x \in [X_0, X_1]^\theta$. For $\varepsilon > 0$ there exists $g \in \mathcal{G}(\vec{X})$ such that $x = g'(\theta)$ and

$$\|g\|_{\mathcal{G}(\vec{X})} \le \|x\|_{[X_0,X_1]^\theta} + \varepsilon. \tag{3-1}$$

Consider the sequence $\{g_m\}_{m=1}^\infty$ given by

$$g_m(z) = \frac{g(z + i/m) - g(z)}{i/m}, \quad z \in \overline{S}.$$

Observe that for each $m \in \mathbb{N}$ we have

$$\max_{j=0,1} \sup_{t \in \mathbb{R}} \|g_m(j+it)\|_{X_0+X_1} \le \max_{j=0,1} \sup_{t \in \mathbb{R}} \|g_m(j+it)\|_{X_j} \le \|g\|_{\mathcal{G}(\vec{X})}.$$

Thus it follows by the Phragmén–Lindelöf principle for Banach spaces that

$$\|g_m(z)\|_{X_0+X_1} \le \|g\|_{\mathcal{G}(\vec{X})}, \quad z \in S.$$

We clearly have that each function $g_m : \overline{S} \to X_0 + X_1$ is continuous and analytic on the strip S. Thus we conclude that $g_m \in \mathcal{F}(\vec{X})$ with $\|g_m\|_{\mathcal{F}(\vec{X})} \le \|g\|_{\mathcal{G}(\vec{X})}$. Hence $g_m(\theta) \in [X_0, X_1]_\theta$ and

$$\|g_m(\theta)\|_{[X_0,X_1]_\theta} \le \|g\|_{\mathcal{G}(\vec{X})}, \quad m \in \mathbb{N}.$$

Since

$$\lim_{m \to \infty} g_m(\theta) = g'(\theta) = x \quad \text{(convergence in } X_0 + X_1\text{)},$$

$x \in [X_0, X_1]_\theta^{\mathrm{c}}$ and so we deduce by (3-1) that

$$\|x\|_{[X_0,X_1]_\theta^{\mathrm{c}}} \le \|x\|_{[X_0,X_1]^\theta} + \varepsilon.$$

Letting ε tend to 0, this estimate completes the proof of claim. Applying the well-known continuous inclusion map with norm less than or equal to 1, see [Bergh and Löfström 1976, Theorem 4.3.1], we learn

$$[X_0, X_1]_\theta \hookrightarrow [X_0, X_1]^\theta$$

which completes the proof of (i).

The proof of (ii) is obvious by (i) and the fact that the unit ball of $([X_0, X_1]_\theta)^{\mathrm{c}}$ is closed in $X_0 + X_1$. □

We will also need results on relationships between the upper complex space $[X_0, X_1]^\theta$ and the *Calderón product* $X_0^{1-\theta} X_1^\theta$ defined for any couple (X_0, X_1) of Banach lattices over a measure space (Ω, Σ, μ), which consists of all $f \in L^0(\mu)$ such that $|f| \le \lambda\, |f_0|^{1-\theta} |f_1|^\theta$ μ-a.e. for some $\lambda > 0$ and $f_j \in B_{X_j}$, $j \in \{0, 1\}$. It is well known, see [Calderón 1964, 13.5, p. 123], that $X_0^{1-\theta} X_1^\theta$ is a Banach lattice endowed with the norm

$$\|f\|_{X_0^{1-\theta} X_1^\theta} = \inf\{\lambda > 0 : |f| \le \lambda\, |f_0|^{1-\theta} |f_1|^\theta,\ f_0 \in B_{X_0},\ f_1 \in B_{X_1}\}$$

and that $X_0^{1-\theta} X_1^\theta$ is continuously embedded into $X_0 + X_1$, which is also a Banach lattice.

We come now to the first theorem of this section.

Theorem 3.2. *For any couple of Banach lattices $\vec{X} = (X_0, X_1)$ over a measure space (Ω, Σ, μ), the following continuous inclusion relation holds with norm less than or equal to* 1:

$$X_0^{1-\theta} X_1^\theta \hookrightarrow [X_0, X_1]^\theta.$$

Before we turn to the proof of this theorem, some comments seem called for. First, in the case of σ-finite measure, this result follows from the vector-valued inclusion proved in [Calderón 1964, p. 125] by taking $B = B_0 = B_1 = \mathbb{C}$. For the reader's convenience, we include a different and transparent proof without assuming the σ-finiteness of the underlying measure space.

Proof of Theorem 3.2. Let $f \in X_0^{1-\theta} X_1^\theta$, so that the estimate $|f| \le |f_0|^{1-\theta} |f_1|^\theta$ holds for some $f_0 \in X_0$ and $f_1 \in X_1$. Note that the support of f is contained in the intersection of the supports of f_0 and f_1. Hence without loss of generality we may suppose that f_0 and f_1 are not equal to zero on Ω.

We define functions

$$F(z) := |f_0|^{1-z} |f_1|^z, \quad G(z) := \int_\theta^z F(w)\, dw, \quad z \in \overline{S}.$$

We claim that $G \in \mathcal{G}(\vec{X})$. To show this we fix $s, t \in \mathbb{R}$ and $j \in \{0, 1\}$. We observe that

$$G(j + is) - G(j + it) = \int_{j+it}^{j+is} |f_0|^{1-w} |f_1|^w\, dw$$

yields $|G(j + is) - G(j + it)| \le |s - t|\, |f_j|$. This implies

$$\max_{j=0,1} \sup_{-\infty < s < t < \infty} \frac{\|G(j + is) - G(j + it)\|_{X_j}}{|s - t|} \le \max_{j=0,1} \|f_j\|_{X_j}. \tag{3-2}$$

We will show that $G : S \to X_0 + X_1$ is analytic. To see this, consider the functions F_0 and F_1 defined by

$$F_0(z) = \chi_{\{|f_0|\ge|f_1|\}}|f_0|^{1-z}|f_1|^z, \quad F_1(z) = F(z) - F_0(z), \quad z \in \overline{S}.$$

We estimate

$$|F_0(z)| = \chi_{\{|f_0|\ge|f_1|\}}|f_0|^{1-\operatorname{Re} z}|f_1|^{\operatorname{Re} z} \le |f_0|.$$

Likewise we have

$$|F_1(z)| \le |f_1|.$$

We define functions

$$G_0(z) = \int_\theta^z F_0(w)\,dw, \quad G_1(z) = \int_\theta^z F_1(w)\,dw \quad z \in \overline{S}.$$

Since $|G_0(z+h) - G_0(z)| \le |h||f_0|$ for all $z, h \in \mathbb{C}$ such that $z+h, z \in \overline{S}$, it follows that $G_0 : \overline{S} \to X_0$ is a continuous function. Similarly, we can establish that $G_1 : \overline{S} \to X_1$ is also continuous.

We now show that the mapping $G_0 : S \to X_0$ is analytic. To this end we fix $0 < \varepsilon < \frac{1}{2}$ and consider the open strip $S_\varepsilon = \{z \in S : \varepsilon < \operatorname{Re} z < 1-\varepsilon\}$. We note that $F_j(z) \in X_j$ for $z \in S_\varepsilon$.

Since

$$\chi_{\{|f_0|\ge|f_1|\}}\left|\left(\frac{|f_1|}{|f_0|}\right)^\varepsilon \exp\left(h\log\left(\frac{|f_1|}{|f_0|}\right)\right) - \left(\frac{|f_1|}{|f_0|}\right)^\varepsilon\left(1 + h\log\left(\frac{|f_1|}{|f_0|}\right)\right)\right| \prec |h|^2,$$

we conclude that

$$G_0(z+h) - G_0(z) - hF_0(z) = \mathrm{O}(|h|^2) \quad \text{as } h \to 0,$$

in X_0 uniformly over $z \in S_\varepsilon$. Similarly, we can show that

$$G_1(z+h) - G_1(z) - hF_1(z) = \mathrm{O}(|h|^2) \quad \text{as } h \to 0,$$

in X_1 uniformly over $z \in S_\varepsilon$. Combining these calculations, we see that $G|_S = G_0|_S + G_1|_S : S \to X_0 + X_1$ is analytic.

To finish the proof of the claim, we need only to observe that

$$\|G_j(z)\|_{X_j} \le \|G_j(z) - G_j(\operatorname{Re}(z))\|_{X_j} + \|G_j(\operatorname{Re}(z)) - G_j(\theta)\|_{X_j} = \mathrm{O}(|z|+1)$$

for $j \in \{0, 1\}$ keeping in mind that $G_j(\theta) = 0$.

Now observe that $G'(\theta) = |f_0|^{1-\theta}|f_1|^\theta$ and by (3-2) $\|G\|_{\mathcal{G}(\vec{X})} \le \max_{j=0,1} \|f_j\|_{X_j}$. Thus we deduce that $|f| \in [X_0, X_1]^\theta$. Since $[X_0, X_1]^\theta$ is a Banach lattice and f_0 and f_1 are arbitrary, we conclude that

$$X_0^{1-\theta}X_1^\theta \hookrightarrow [X_0, X_1]^\theta,$$

with norm of the continuous inclusion map less than or equal to 1. □

Remark 3.3. The inclusion in the above theorem is proper in general. To see this we recall that Lozanovskii [1972] constructed a closed Banach sublattice Y_0 of a weighted Banach lattice $L^\infty(w)$ on $((0, 1), m)$ with Lebesgue measure m, where

$$L^\infty(w) = \{f \in L^\infty(0, 1) : wf \in L^\infty(0, 1)\}$$

with $w(t) = t$ for all $t \in (0,1)$ and endowed with the norm $\|f\|_{L^\infty(w)} = \|wf\|_{L^\infty}$, such that $(L^\infty(w))^{1-\theta}(L^\infty)^\theta$ and $Y_0^{1-\theta}(L^\infty)^\theta$ are not relative interpolation spaces with respect to $(L^\infty(w), L^\infty)$ and (Y_0, L^∞) for any $\theta \in (0,1)$. We complexify these spaces. Since $[\,\cdot\,]^\theta$ is an interpolation functor in the class of complex Banach spaces, it follows that for the couple $(X_0, X_1) = (Y_0(\mathbb{C}), L^\infty(\mathbb{C}))$ the inclusion

$$X_0^{1-\theta} X_1^\theta \subset [X_0, X_1]^\theta$$

is proper for an arbitrary $\theta \in (0,1)$.

The following result shows isometric relationships between Banach lattices generated by standard operations and Calderón's constructions in the setting of couples of complex Banach lattices on an arbitrary measure space:

Theorem 3.4. *Let $\vec{X} = (X_0, X_1)$ be a couple of complex Banach lattices on an arbitrary measure space (Ω, Σ, μ). Then the following statements are true for all $\theta \in (0,1)$:*

(i) $[X_0, X_1]_\theta \cong (X_0^{1-\theta} X_1^\theta)^\circ$.

(ii) $([X_0, X_1]^\theta)^{\mathrm{c}} \cong (X_0^{1-\theta} X_1^\theta)^{\mathrm{c}}$.

(iii) $[X_0, X_1]^\theta \cong X_0^{1-\theta} X_1^\theta$ *whenever the unit ball of* $X_0^{1-\theta} X_1^\theta$ *is closed in* $X_0 + X_1$.

Proof. We begin with (i). Since $[\vec{X}]_\theta \cong ([X_0, X_1]^\theta)^\circ$ is a closed subspace of $[\vec{X}]^\theta$ and the norm in $[\vec{X}]_\theta$ is the restriction of the norm in $[\vec{X}]^\theta$, see [Bergh 1979], if follows from Theorem 3.2 that

$$(X_0^{1-\theta} X_1^\theta)^\circ \hookrightarrow ([X_0, X_1]^\theta)^\circ \cong [X_0, X_1]_\theta,$$

with norm of the continuous inclusion map less or equal to 1.

To obtain the reverse inclusion, we recall that Calderón proved

$$[X_0, X_1]_\theta \hookrightarrow (X_0^{1-\theta} X_1^\theta)^\circ$$

for any σ-finite measure space. The proof of the following continuous inclusion map with norm less than or equal to 1 given in [Calderón 1964, (i), p. 125], see also [Krein et al. 1982, pp. 240–241], works for any measure space: combining two the above continuous inclusions, we obtain the statement (i).

To finish the proofs of (ii) and (iii) at the same time, we observe that the above inclusion, combined with Lemma 3.1 and Theorem 3.2, yields continuous inclusion maps with norm less than or equal to 1,

$$X_0^{1-\theta} X_1^\theta \hookrightarrow [X_0, X_1]^\theta \hookrightarrow ([X_0, X_1]_\theta)^{\mathrm{c}} \hookrightarrow ((X_0^{1-\theta} X_1^\theta)^\circ)^{\mathrm{c}} \hookrightarrow (X_0^{1-\theta} X_1^\theta)^{\mathrm{c}}.$$

Clearly these inclusions complete the proofs of statements (ii) and (iii). □

Remark 3.5. We note that in the case of σ-finite measure spaces the above statement (i) was proved by Shestakov [1974], who extended Calderón's result [1964] proved under the assumption that $X_0^{1-\theta} X_1^\theta$ is σ-order continuous.

We will apply Theorem 3.4 to some class of Banach lattices. In what follows we assume that a measure space (Ω, Σ, μ) is such that $L^0(\Omega, \Sigma, \mu)$ is a Dedekind complete vector lattice (i.e., every subset of $L^0(\mu)$ order bounded from above has a supremum). We refer to [Fremlin 1974, Theorem 64 B, p. 170] for a description and general examples of such measure spaces. Let us just notice here that such measure spaces are semifinite (i.e., for any $A \in \Sigma$ with $\mu(A) > 0$, there is $B \in \Sigma$ such that $B \subset A$ and $0 < \mu(B) < \infty$). Since $L^0(\Omega, \Sigma, \mu)$ is a Dedekind complete vector lattice, it follows that for any subset $E \subset L^0(\Omega, \Sigma, \mu)$, the set $\{\chi_{\operatorname{supp} x} : x \in E\}$ is order bounded in $L^0(\mu)$. If we put $x_0 := \sup\{\chi_{\operatorname{supp} x} : x \in E\}$, then $\operatorname{supp}(E)$ exists and is given by $\operatorname{supp}(E) = \operatorname{supp} x_0$.

We shall need to use some results on Köthe duality. We recall that the *Köthe dual* space X' of any Banach lattice X on (Ω, Σ, μ) is defined to be the space of all $x' \in L^0(\mu)$ with $\operatorname{supp} x \subset \operatorname{supp}(X)$ such that $xx' \in L^1(\mu)$ for all $x \in X$. It is well known that X' is a Banach lattice on (Ω, Σ, μ) equipped with the norm

$$\|x'\|_{X'} := \sup\left\{ \int_\Omega |xx'|\, d\mu : \|x\|_X \le 1 \right\}, \quad x' \in X'.$$

As usual, the Köthe dual space of X' is denoted by X''. If a Banach lattice X is such that $X \cong X''$, then X is said to be *maximal*.

We note that if X is a Banach lattice on a σ-finite measure space (Ω, Σ, μ), then $X \cong X''$ if and only if X has the Fatou property; see, [Kantorovich and Akilov 1982, Theorem 7, p. 191]. We recall that X is said to have the *Fatou property* if for any sequence $\{f_m\}_{m=1}^\infty$ of nonnegative elements from X such that $f_m \uparrow f$ for $f \in L^0(\Omega)$ and $\sup_{m\ge 1} \|f_m\|_X < \infty$, one has $f \in X$ and $\|f_m\|_X \uparrow \|f\|_X$.

We will need the following lemma.

Lemma 3.6. *Let X_0, X_1 and X be Banach lattices on a measure space (Ω, Σ, μ). If X is an intermediate space with respect to (X_0, X_1), then $X^{\mathrm{c}} \hookrightarrow X''$ with the norm inclusion less than or equal to* 1.

Proof. Fix $x \in X^{\mathrm{c}}$. Then in a similar fashion to the proof of [Cwikel and Nilsson 2003, Lemma 1.16], we claim that for a given $\varepsilon > 0$, there exists a sequence $\{y_m\}$ in X such that $0 \le y_m \uparrow |x|$ μ-a.e., $\|y_m\|_X \le (1+\varepsilon)\|x\|_{X^{\mathrm{c}}}$ for each $m \in \mathbb{N}$ and $y_m \to y$ in $X_0 + X_1$. In fact, it follows by the definition of X^{c} that we can find a sequence $\{z_m\}$ in X such that $\|z_m\|_X \le (1+\varepsilon)\|x\|_{X^{\mathrm{c}}}$ for each $m \in \mathbb{N}$ and $z_m \to z$ in $X_0 + X_1$. If we set $y_m = \min\{\max\{0, z_m\}, |x|\}$ for each $m \in \mathbb{N}$, then we obtain the desired sequence. We conclude by Lebesgue's monotone convergence theorem that for any $x' \in X'$,

$$\int_\Omega |xx'|\, d\mu = \lim_{m\to\infty} \int_\Omega y_m |x'|\, d\mu \le (1+\varepsilon)\|x\|_{X^{\mathrm{c}}} \|x'\|_{X'}.$$

Since $\varepsilon > 0$ is arbitrary, the desired continuous inclusion follows. □

We are now ready to state the following result.

Theorem 3.7. *Assume that a measure space (Ω, Σ, μ) is such that $L^0(\Omega, \Sigma, \mu)$ is a Dedekind complete vector lattice and it satisfies the following condition: if $A \subset \Omega$ is such that $A \cap B \in \Sigma$ for every set $B \in \Sigma$, $\mu(B) < \infty$, then $A \in \Sigma$. Let $\vec{X} = (X_0, X_1)$ be an arbitrary couple of complex Banach lattices on (Ω, Σ, μ). If X_0 and X_1 are both maximal and* $\operatorname{supp}(X_0) = \operatorname{supp}(X_1) = \Omega$, *then*

$$[X_0, X_1]^\theta \cong X_0^{1-\theta} X_1^\theta.$$

Proof. Since both X_0 and X_1 are maximal, it follows by the original Lozanovskii duality formula [1978]

$$(X_0^{1-\theta}X_1^{\theta})' \cong (X_0')^{1-\theta}(X_1')^{\theta}$$

that $X_0^{1-\theta}X_1^{\theta}$ is also a maximal Banach lattice. Thus, we deduce from Lemma 3.6 that

$$X_0^{1-\theta}X_1^{\theta} \hookrightarrow (X_0^{1-\theta}X_1^{\theta})^c \hookrightarrow (X_0^{1-\theta}X_1^{\theta})'' \cong X_0^{1-\theta}X_1^{\theta}$$

with the norm of the inclusion maps less than or equal to 1. □

Corollary 3.8. *Let $\vec{X} = (X_0, X_1)$ be an arbitrary couple of complex Banach lattices with the Fatou property on a σ-finite measure space (Ω, Σ, μ). If* $\operatorname{supp}(X_0) = \operatorname{supp}(X_1) = \Omega$, *then*

$$[X_0, X_1]^{\theta} \cong X_0^{1-\theta}X_1^{\theta}.$$

Proof. Clearly any σ-finite measure space (Ω, Σ, μ) satisfies the desired condition from Theorem 3.7. Moreover, it is well known that $L^0(\Omega, \Sigma, \mu)$ is a Dedekind complete vector lattice, see [Kantorovich and Akilov 1982], and so the desired statement follows from Theorem 3.7. □

Remark 3.9. Lozanovskii proved that for all $\theta \in (0, 1)$ we have

$$(X_0^{1-\theta}X_1^{\theta})' = (X_0')^{1-\theta}(X_1')^{\theta}$$

with equality of norms; see [Lozanovskii 1969, Theorem 2]. Using this result for $\theta = \frac{1}{2}$, Lozanovskii [1969] showed $X^{1/2}(X')^{1/2} \simeq L^2$ for any Banach lattice on a given σ-finite measure space. Thus taking X which does not enjoy the Fatou property, we conclude that the Fatou property of $X_0^{1-\theta}X_1^{\theta}$ does not always imply that the Fatou property holds for X_0 and X_1. For further examples refer to [Reisner 1993], where among others it is shown (see Example 2) that there exist σ-order continuous Banach sequence lattices X and Y that do not enjoy the Fatou property such that $X^{1-\theta}Y^{\theta}$ is σ-order continuous and has the Fatou property.

Simple calculation shows that for any Banach lattice X and every $1 < r < \infty$ we have $X^{1/r}(L^\infty)^{1-1/r} \cong X^{(r)}$; thus by Theorem 3.7 we obtain the following useful formula:

Corollary 3.10. *Let X be a Banach lattice with the Fatou property on a σ-finite measure space. Then, for any $\theta \in (0, 1)$,*

$$[X, L^\infty]^{\theta} \cong X^{1-\theta}(L^\infty)^{\theta} \cong X^{(r)},$$

where $r = (1-\theta)^{-1}$.

An immediate application of our results is the following variant of the Riesz–Thorin interpolation theorem:

Theorem 3.11. *Let (X_0, X_1) and (Y_0, Y_1) be couples of complex Banach lattices on measure spaces. Then for every linear operator $T : (X_0, X_1) \to (Y_0, Y_1)$ and all $\theta \in (0, 1)$, T is bounded from $X_0^{1-\theta}X_1^{\theta}$ into $(Y_0^{1-\theta}Y_1^{\theta})^{\mathrm{c}}$ and satisfies*

$$\|T\|_{X_0^{1-\theta}X_1^{\theta} \to (Y_0^{1-\theta}Y_1^{\theta})^{\mathrm{c}}} \le \|T\|_{X_0 \to Y_0}^{1-\theta}\|T\|_{X_1 \to Y_1}^{\theta}.$$

In particular if Y_0 *and* Y_1 *are Banach lattices on a σ-finite measure space* (Ω, Σ, μ) *with* $\operatorname{supp}(Y_0) = \operatorname{supp}(Y_1) = \Omega$ *and both enjoy the Fatou property, then* T *is bounded from* $X_0^{1-\theta} X_1^{\theta}$ *into* $Y_0^{1-\theta} Y_1^{\theta}$ *with the norm estimate*

$$\|T\|_{X_0^{1-\theta} X_1^{\theta} \to Y_0^{1-\theta} Y_1^{\theta}} \le \|T\|_{X_0 \to Y_0}^{1-\theta} \|T\|_{X_1 \to Y_1}^{\theta}.$$

Proof. We have $\|T\|_{[X_0,X_1]^\theta \to [X_0,X_1]^\theta} \le \|T\|_{X_0 \to Y_0}^{1-\theta} \|T\|_{X_1 \to Y_1}^{\theta}$ according to [Bergh and Löfström 1976, Theorem 4.1.4]. Since $X_0^{1-\theta} X_1^{\theta} \hookrightarrow [X_0, X_1]^\theta$ by Theorem 3.2 with the inclusion constant less than or equal to 1 and $[Y_0, Y_1]^\theta \hookrightarrow (Y_0^{1-\theta} Y_1^{\theta})^{\mathrm{c}}$ by Lemma 3.1 again with the inclusion constant less than or equal to 1, the first required estimate follows. This estimate combined with the continuous inclusions shown in the proof of Theorem 3.7 yields the second estimate. □

The following lemma is surely well known to specialists, but we include a proof for the sake of completeness.

Lemma 3.12. *If* $0 < \theta < 1$ *then for any couple* $\vec{X} = (X_0, X_1)$ *of complex Banach spaces, we have* $[X_0, X_1]^\theta \hookrightarrow (X_0, X_1)_{\theta,\infty}$ *with the norm of the inclusion map less or equal to* 1.

Proof. It is well known that for any Banach couples $\vec{X} = (X_0, X_1)$ and $\vec{Y} = (Y_0, Y_1)$ and any operator $T : \vec{X} \to \vec{Y}$, we have the following estimate for restrictions of T (see [Bergh and Löfström 1976, Theorem 4.1.4]):

$$\|T\|_{[\vec{X}]^\theta \to [\vec{Y}]^\theta} \le \|T\|_{X_0 \to Y_0}^{1-\theta} \|T\|_{X_1 \to Y_1}^{\theta}. \tag{3-3}$$

We fix $t > 0$ and $x \in [X_0, X_1]^\theta$. By the Hahn–Banach theorem, there exists a continuous linear functional $x^* \in (X_0 + X_1)^*$ such that $x^*(x) = K(t, x; \vec{X})$ and

$$|x^*(y)| \le K(t, y; \vec{X}), \quad y \in X_0 + X_1.$$

This implies that the linear operator $T : X_0 + X_1 \to \mathbb{C}$ defined by $T(y) = x^*(y)$ for all $y \in X_0 + X_1$ satisfies $T : (X_0, X_1) \to (\mathbb{C}, \mathbb{C})$ with $\|T\|_{X_0 \to \mathbb{C}} \le 1$ and $\|T\|_{X_1 \to \mathbb{C}} \le t$.

Now we apply the estimate (3-3) to $(Y_0, Y_1) = (\mathbb{C}, \mathbb{C})$ and the obvious equality $[\mathbb{C}, \mathbb{C}]^\theta \cong \mathbb{C}$ to get that

$$K(t, x; \vec{X}) = x^*(x) = T(x) \le t^\theta \|x\|_{[\vec{X}]^\theta}.$$

Since $t > 0$ and $x \in [X_0, X_1]^\theta$ are arbitrary, the proof is complete. □

We conclude with the remark that if (X_0, X_1) is a couple of complex Banach lattices which enjoy the Fatou property, then the formula $[X_0, X_1]^\theta = X_0^{1-\theta} X_1^{\theta}$ (up to equivalence of norms) is a consequence of abstract interpolation results, combined with relationships between the Köthe duality results and the orbital descriptions of the upper complex method and of other interpolation constructions; see [Ovchinnikov 1984, pp. 474–492] for more details.

4. On Calderón–Mityagin couples of Morrey spaces

As is mentioned in the Introduction, one of the fundamental problems in the theory of interpolation spaces is the description of all interpolation spaces X with respect to a given compatible couple of Banach spaces $\vec{X} = (X_0, X_1)$. Cwikel [1981] posed the question of whether in fact *all* Calderón–Mityagin couples can be identified by checking whether their complex interpolation spaces are K-spaces. In [Mastyło and Ovchinnikov 1997] the authors provided counterexamples, which give a negative answer to this question.

Also as is mentioned in the Introduction there is no complete description of the complex interpolation spaces between Morrey spaces in the general case. We show in this section that the complex spaces with respect to any couple of Morrey spaces which are not L^p-spaces cannot be described by the K-method of interpolation. In particular this implies that these couples are not Calderón–Mityagin couples.

Before we state and prove the main results of this section we need some technical observations. Suppose that we have a nonnegative measurable function on $\mathbb{R}^n$ which is rotationally symmetric, so that it can be expressed as $f(|x|)$, where $f:[0,\infty)\to\mathbb{R}_+$ is a measurable function. Recall that $|B(x,r)|=v_n r^n$ for all $x\in\mathbb{R}^n$, where $v_n:=\pi^{n/2}/\Gamma(1+n/2)$ is the measure of the unit ball $B(1)$.

Then a standard calculation via applying spherical coordinates gives

$$\int_{\mathbb{R}^n} f(|x|)\,dx=\frac{2\pi^{n/2}}{\Gamma(n/2)}\int_{\mathbb{R}_+} f(t)t^{n-1}\,dt.$$

Since the ball $B(x,r)$ has the same measure as $B(0,r)$ for every $r>0$, by choosing $f=\chi_{(0,r)}$, we obtain

$$|B(x,r)|=|B(0,r)|=\int_{\mathbb{R}^n}\chi_{(0,r)}(|y|)\,dy=\frac{2\pi^{n/2}}{\Gamma(n/2)}\int_0^r t^{n-1}\,dt=v_n r^n.$$

Combining these formulas we conclude that if $1<s<\infty$ and $f(t)=t^{-n/s}\chi_{(0,r)}$ for all $t\geq 0$, then

$$\int_{B(0,r)}|x|^{-n/s}\,dx=v_n\frac{s}{s-1}r^{n-n/s}.$$

This shows that the function $x\mapsto|x|^{-n/s}$ belongs to $L\mathcal{M}_1^s$ and its norm is equal to ${v_n}^{1/s}s/(s-1)$. In what follows we will use that this function belongs to the Morrey space $\mathcal{M}_1^s$. For the sake of completeness we include a proof of this fact.

Proposition 4.1. *If $1<s<n$, then the function $x\mapsto|x|^{-n/s}$ belongs to $\mathcal{M}_1^s$ and its norm is equal to ${v_n}^{1/s}s/(s-1)$.*

Proof. Let $g(x)=|x|^{-n/s}$ for all $x\in\mathbb{R}^n$ (we put $g(0):=0$). We observe that

$$|\{y\in\mathbb{R}^n:|g(y)|>\tau\}|=|B(0,\tau^{-s/n})|=v_n\tau^{-s},\quad \tau>0.$$

Hence, for any $x\in\mathbb{R}^n$ and any $r>0$, we have

$$\begin{aligned}\int_{B(x,r)}|g(y)|\,dy&=\int_0^\infty|\{y\in\mathbb{R}^n:\chi_{B(x,r)}(y)g(y)>\tau\}|\,d\tau\\&\leq\int_0^\infty\min\big\{|B(x,r)|,\,|\{y\in\mathbb{R}^n:g(y)>\tau\}|\big\}\,d\tau=v_n\int_0^\infty\min\{r^n,\tau^{-s}\}\,d\tau\\&=v_n r^{n(1-1/s)}+v_n\int_{r^{-n/s}}^\infty\tau^{-s}\,d\tau=v_n\frac{s}{s-1}\,r^{n(1-1/s)}\\&={v_n}^{1/s}\frac{s}{s-1}|B(x,r)|^{1-1/s}.\end{aligned}$$

The above estimate combined with previous calculation completes the proof. ☐

According to Lemma 3.12, for a Banach couple (A_0, A_1), $[A_0, A_1]^\theta$ is a subspace of $(A_0, A_1)_{\theta,\infty}$ for any $\theta \in (0, 1)$. More precisely, $[A_0, A_1]^\theta \hookrightarrow (A_0, A_1)_{\theta,\infty}$ with the norm of the inclusion map less than or equal to 1. Concerning this inclusion, we will also need the following useful Proposition 4.2:

Here and below in this paper we let $1 \le q_0 \le p_0 \le \infty$ and $1 \le q_1 \le p_1 \le \infty$ satisfy $\min(p_0, p_1) < \infty$.

We will apply Proposition 4.2 below to Morrey spaces. Again according to Lemma 3.12,

$$[\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}]^\theta \hookrightarrow (\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1})_{\theta,\infty}.$$

We have the following result.

Proposition 4.2. *Let $\vec{A} = (A_0, A_1)$ be a Banach couple, and let A be any interpolation space with respect to $\vec{A}$ that is a proper subspace of $\vec{A}_{\theta,\infty}$ for some $\theta \in (0, 1)$. If there exists $a_\infty \in A$ such that one of the following conditions is satisfied, then $\vec{A}$ is not a Calderón–Mityagin couple*:

(i) $K(t, a_\infty; \vec{A}) \succ t^\theta$ *for all* $t \in (0, \infty)$.

(ii) $A_1 \hookrightarrow A_0$ *and there exists* $t_0 > 0$ *such that* $K(t, a_\infty; \vec{A}) \succ t^\theta$ *for all* $t \in (0, t_0]$.

Proof. (i) For every $a \in \vec{A}_{\theta,\infty}$ we set $x_a = \|a\|_{\vec{A}_{\theta,\infty}} a_\infty \in A$. Our hypothesis gives that

$$K(t, x_a; \vec{A}) \succ t^\theta \|a\|_{\vec{A}_{\theta,\infty}} \ge K(t, a; \vec{A}), \quad t > 0.$$

Suppose that A is a proper subset of $\vec{A}_{\theta,\infty}$. If $\vec{A}$ were a Calderón–Mityagin couple, then we would get $a = T(x_a)$ for some operator $T : \vec{A} \to \vec{A}$, and therefore, by interpolation, $a \in A$. Since $a \in \vec{A}_{\theta,\infty}$ is arbitrary, this would imply that $\vec{A}_{\theta,\infty} \subset A$ and

$$A = \vec{A}_{\theta,\infty},$$

which is a contradiction with our hypothesis.

(ii) There is no loss of generality in assuming that t_0 is equal to $\|\mathrm{id}\|_{A_1 \to A_0}$, the operator norm of the embedding of A_1 into A_0. It is clear that $K(t, x; \vec{A}) = \|a\|_{A_0}$ for every $a \in A_0$ and $t \ge t_0$. In this case our hypothesis implies that for each $a \in \vec{A}_{\theta,\infty}$ there exists an element $x_a \in A$ which is a suitable scalar multiple of a_∞ which satisfies

$$K(t, a; \vec{A}) \prec K(t, x_a; \vec{A}), \quad t \in (0, t_0],$$

and also $\|a\|_{A_0} \le \|x_a\|_{A_0}$. Therefore, $K(t, a; \vec{A}) \prec K(t, x_a; \vec{A})$ for all $t > 0$ and this enables us to complete the proof via the same reasoning as in part (i). □

Theorem 4.3. *Let $1 \le q_0 < p_0 < \infty$ and $1 \le q_1 \le p_1 \le \infty$ with $q_0 \ne q_1$. Then for any $\theta \in (0, 1)$ the inclusion $[\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}]^\theta \subset (\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1})_{\theta,\infty}$ is proper.*

Proof. Define p and q by

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

We distinguish two cases: (1) $p_1 < \infty$ and (2) $p_1 = \infty$.

Case (1): $p_1 < \infty$. Since $p_0 < \infty$ and $q_0 \neq q_1$, the proof of [Lemarié-Rieusset 2013, case (b), p. 751] shows that $(\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1})_{\theta,\infty}$ is not embedded into $\mathcal{M}^{p}_{q}$. Meanwhile, if we go through an argument similar to the one to prove $(\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1})_{\theta,q} \hookrightarrow \mathcal{M}^{p}_{q}$ using the result $(L^{q_0}, L^{q_1})_{\theta,q} = L^q$ by Calderón, we have $[\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}]^\theta \hookrightarrow \mathcal{M}^{p}_{q}$. Combining these observations, we conclude that $[\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}]^\theta$ is a proper subspace of $(\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1})_{\theta,\infty}$.

Case (2): $p_1 = \infty$. We have $\mathcal{M}^{p_1}_{q_1} = L^\infty$ isometrically for all choices of q_1 in the permitted range $[1, \infty]$.

We apply [Cwikel and Gulisashvili 2000, Theorem 4] again to see that $(X, L^\infty)_{\theta,\infty}$ is the space of all $f \in L^0$ in a quasi-Banach lattice X_θ endowed with the quasinorm

$$\|f\|_{X_\theta} := \sup_{\lambda>0} \lambda \, \|\chi_{\{|f|>\lambda\}}\|_X^{1-\theta} < \infty,$$

whenever X is a Banach function lattice on a σ-finite complete nonatomic measure space (Ω, Σ, μ). By Corollary 3.10, we have $[\mathcal{M}^{p_0}_{q_0}, L^\infty]^\theta = (\mathcal{M}^{p_0}_{q_0})^{(r)} = \mathcal{M}^{rp_0}_{rq_0}$. Furthermore, the inclusion $\mathcal{M}^{p}_{q} \subset \mathrm{w}\mathcal{M}^{p}_{q}$ is proper for every $1 \le q < p < \infty$ (see Lemma 4.7(ii)), we conclude that the inclusion

$$[\mathcal{M}^{p_0}_{q_0}, L^\infty]^\theta = \mathcal{M}^{rp_0}_{rq_0} \hookrightarrow (\mathcal{M}^{p_0}_{q_0}, L^\infty)_{\theta,\infty} = \mathrm{w}\mathcal{M}^{rp_0}_{rq_0}$$

is also proper, and so the proof is complete. □

We also will need the following results: the first one is motivated by [Brudnyi and Krugljak 1991, Theorem 4.5.5].

Theorem 4.4. *Let $\vec{X} = (X_0, X_1)$ be a Banach couple, and let $r \in (0,1)$, $\theta \in (0,1)$ and $\gamma \in (1,\infty)$ be fixed. Assume that for each $j \in \mathbb{J} = \mathbb{Z}$ (resp. $j \in \mathbb{J} = \mathbb{Z}_+$) there exists $v_j \in X_0 + X_1$ such that*

$$\min\{1, r^{-j}t\} \le K(t, v_j; \vec{X}) \le \gamma \min\{1, r^{-j}t\}, \quad t > 0 \ \ (\textit{resp. } t \in (0,1]).$$

Then, for a certain positive integer N which depends on r, θ and γ, the element $x_\theta \in X_0 + X_1$ defined by

$$x_\theta = \sum_{j \in \mathbb{J} \cap N\mathbb{Z}} r^{j\theta} v_j$$

satisfies

$$K(t, x_\theta, \vec{X}) \asymp t^\theta \quad \textit{for all } t > 0 \ \ (\textit{resp. for all } t \in (0,1]).$$

Proof. Fix $\theta \in (0,1)$ and $r \in (0,1)$. It is easy to check that there is a constant $C = C(r,\theta) > 1$ such that for each positive integer N we have

$$\sum_{j \in \mathbb{J} \cap N\mathbb{Z}} r^{j\theta} \min\{1, r^{-j}t\} \le C t^\theta, \quad t > 0. \tag{4-1}$$

Thus, we conclude by our hypothesis that for $\mathbb{J} = \mathbb{Z}$ (resp. $\mathbb{J} = \mathbb{Z}_+$) the series

$$x_\theta := \sum_{j \in \mathbb{J} \cap N\mathbb{Z}} r^{j\theta} v_j$$

converges in $X_0 + X_1$.

Combining (4-1) with the right-hand inequality of our hypothesis yields that for x_θ we have, for all $t \in (0,\infty)$ (resp. $t \in (0,1]$),

$$K(t, x_\theta; \vec{X}) \le \sum_{j\in\mathbb{J}} r^{j\theta} K(t, v_j; \vec{X}) \le \gamma \sum_{j\in\mathbb{J}} r^{j\theta} \min\{1, r^{-j}t\} \le \gamma C t^\theta.$$

The sums $\sum_{m=1}^{\infty} r^{N\theta m}$ and $\sum_{m=-\infty}^{-1} r^{-N(1-\theta)m}$ can be made arbitrarily small by choosing $N \in \mathbb{N}$ large enough. So it is clear that we can choose a positive integer N which depends only on r, θ and γ and which is large enough to satisfy

$$\gamma \sum_{m\in\mathbb{Z}\setminus\{0\}} \min\{r^{N\theta m}, r^{-N(1-\theta)m}\} \le \frac{1}{2}.$$

For this N and for each $j \in \mathbb{J} \cap N\mathbb{Z}$, setting $j_0 = j/N$, we consequently have

$$\begin{aligned}\gamma \sum_{k\in\mathbb{J}\cap N\mathbb{Z}\setminus\{j\}} r^{k\theta} \min\{1, r^{j-k}\} &\le r^{j\theta}\gamma \sum_{k\in N\mathbb{Z}\setminus\{j\}} \min\{r^{(k-j)\theta}, r^{(j-k)(1-\theta)}\} \\ &\le r^{j\theta}\gamma \sum_{m\in\mathbb{Z}\setminus\{j_0\}} \min\{r^{N(m-j_0)\theta}, r^{N(j_0-m)(1-\theta)}\} \\ &= r^{j\theta}\gamma \sum_{m\in\mathbb{Z}\setminus\{0\}} \min\{r^{Nm\theta}, r^{-Nm(1-\theta)}\} \le \frac{1}{2} r^{j\theta}.\end{aligned}$$

So, we get that for each $j \in \mathbb{J} \cap N\mathbb{Z}$,

$$\begin{aligned}K(r^j, x_\theta; \vec{X}) &\ge K(r^j, r^{j\theta} v_j; \vec{X}) - K\Bigl(r^j, \sum_{k\in\mathbb{J}\cap N\mathbb{Z}\setminus\{j\}} r^{k\theta} v_k; \vec{X}\Bigr) \\ &\ge r^{j\theta} - \sum_{k\in\mathbb{J}\cap N\mathbb{Z}\setminus\{j\}} r^{k\theta} K(r^j, v_k; \vec{X}) \\ &\ge r^{j\theta} - \gamma \sum_{k\in\mathbb{J}\cap N\mathbb{Z}\setminus\{j\}} r^{k\theta} \min\{1, r^{j-k}\} \ge \frac{1}{2} r^{j\theta}.\end{aligned}$$

To conclude the proof, observe that for a given $t > 0$ (resp. $t \in (0,1]$) there is an integer $j \in N\mathbb{Z}$ (resp. $j \in \mathbb{Z}_+ \cap N\mathbb{Z}$) such that $r^j \le t \le r^{j-N}$. Then the above estimate yields

$$K(t, x_\theta; \vec{X}) \ge K(r^j, x_\theta; \vec{X}) \ge \tfrac{1}{2} r^{j\theta} \ge \tfrac{1}{2} r^{N\theta} t^\theta.$$

Thus $K(t, x_\theta; \vec{X}) \asymp t^\theta$ for all $t > 0$ (resp. $t \in (0,1]$), as required. □

As an application we obtain the following result:

Lemma 4.5. *Let the set $\mathbb{J}$ be either $\mathbb{Z}$ or $\mathbb{Z}_+$. Let $\mathcal{T}$ be the interval $(0,\infty)$ if $\mathbb{J} = \mathbb{Z}$ or $(0,1]$ if $\mathbb{J} = \mathbb{Z}_+$. Let (X_0, X_1) be a couple of Banach lattices on a complete σ-finite measure space (Ω, Σ, μ), and let $r \in (0,1)$, $\theta \in (0,1)$ be fixed. Assume that there exists a sequence $\{F_j\}_{j\in\mathbb{J}}$ in Σ with positive measures such that $\|\chi_{F_j}\|_{X_1} \asymp r^{-j} \|\chi_{F_j}\|_{X_0}$ and that $K(t, \chi_{F_j}; \vec{X}) \asymp \min\{\|\chi_{F_j}\|_{X_0}, t\|\chi_{F_j}\|_{X_1}\}$ for each $j \in \mathbb{J}$*

and all $t \in \mathcal{T}$. Then there exists a positive integer N for which the function f_θ defined by

$$f_\theta = \sum_{j \in \mathbb{J} \cap N\mathbb{Z}} r^{j\theta} \frac{\chi_{F_j}}{\|\chi_{F_j}\|_{X_0}}$$

has the following two properties:

(i) *$f_\theta \in X_0 + X_1$ and $K(t, f_\theta; \vec{X}) \asymp t^\theta$ for all $t \in \mathcal{T}$.*

(ii) *If furthermore $\{F_j\}_{j\in\mathbb{J}}$ is a nondecreasing sequence and if $\chi_{\bigcup_{j\in\mathbb{J}\cap N\mathbb{Z}} F_j} \in X_0$ and if $f_\theta^{1/\theta} \in X_1$, then $f_\theta \in {X_0}^{1-\theta} {X_1}^{\theta}$.*

Proof. (i) If we let $v_j := \chi_{F_j} / \|\chi_{F_j}\|_{X_0}$ we have, for each $j \in \mathbb{J}$,

$$K(t, v_j; \vec{X}) \asymp \min\{1, r^{-j} t\}, \quad t \in \mathcal{T}.$$

Thus the statement (i) follows from Theorem 4.4.

(ii) This is a direct consequence of the decomposition:

$$f_\theta = (\chi_{\bigcup_{j\in\mathbb{J}\cap N\mathbb{Z}} F_j})^{1-\theta} (f_\theta^{1/\theta})^\theta \in X_0^{1-\theta} X_1^\theta. \qquad \square$$

For $s > 1$, we write $s' = s/(s-1)$. We will need the following useful lemma:

Lemma 4.6. *For a given $s > 1$ we put $\alpha = 2^{-s'}$. For each $\varepsilon \in \{0,1\}^n$, we define an affine map $f_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ by*

$$f_\varepsilon(x) = \alpha x + (1-\alpha)\varepsilon, \quad x \in \mathbb{R}^n.$$

We also define two sequences $\{F_j\}_{j=0}^\infty$ and $\{E_j\}_{j=0}^\infty$ of subsets of $\mathbb{R}^n$ by $F_j = \alpha^{-j} E_j$ for each $j \geq 0$, where $E_0 := [0,1]^n$ and E_j are given by

$$E_j := \bigcup_{\varepsilon \in \{0,1\}^n} f_\varepsilon(E_{j-1}), \quad j \in \mathbb{N}.$$

Then the following statements are true:

(i) *$F_j \subset F_{j+1}$ for each $j \geq 0$.*

(ii) *F_j is made up of 2^{jn} pairwise disjoint cubes of volume 1 for each $j \in \mathbb{N}$.*

(iii) *For all $1 < u < \infty$ and each $j \geq 0$ we have*

$$\|\chi_{F_j}\|_{\mathcal{M}_1^u} \asymp \max\{1, \alpha^{jn/u - jn} 2^{-jn}\} = \max\{1, \alpha^{jn/u - jn/s}\},$$

where the constants of equivalence do not depend on j.

(iv) *Let $1 \leq q_0 \leq p < \infty$ satisfy $s = p/q_0$. Then $\chi_{\bigcup_{j\in\mathbb{Z}_+} F_j} \in \mathcal{M}_{q_0}^p$.*

(v) *For every $x \in \mathbb{R}^n$ and each $j \in \mathbb{N}$, we have*

$$M\chi_{F_j}(x) \succ \sum_{k=1}^{j} \frac{\chi_{F_k}(\alpha^{j-k} x)}{\|\chi_{F_k}(\alpha^{j-k} \cdot)\|_{L\mathcal{M}_1^s}^*},$$

where $\|\cdot\|^*_{L\mathcal{M}^s_1}$ *is the local Morrey norm generated by cubes*:

$$\|f\|^*_{L\mathcal{M}^s_1} := \sup_{r>0} r^{n/p-n} \int_{[-r,r]^n} |f(y)|\,dy < \infty.$$

Proof. The statements (i) and (ii) are obvious. So we concentrate on other parts.

(iii) The statement follows from the equivalence below obtained via standard calculations using an equivalent norm in Morrey spaces generated by cubes, and the fact that α satisfies $2\alpha = \alpha^{1/s}$:

$$\|\chi_{E_j}\|^*_{\mathcal{M}^u_1} \asymp \max\{\alpha^{jn/u}, 2^{jn}\alpha^{jn}\} = 2^{jn}\alpha^{jn} \max\{1,\, \alpha^{jn/u-jn}2^{-jn}\}, \quad j \geq 0.$$

We include the proof of the equivalence $\|\chi_{E_j}\|^*_{\mathcal{M}^u_1} \asymp \max\{\alpha^{jn/u}, 2^{jn}\alpha^{jn}\}$ for the reader's convenience. From the definition of $\|\cdot\|^*_{\mathcal{M}^u_1}$, we have

$$\|\chi_{E_j}\|^*_{\mathcal{M}^u_1} = \sup_Q |Q|^{1/u-1}\left(\int_Q \chi_{E_j}(y)\,dy\right) = \sup_Q |Q|^{1/u-1}|Q \cap E_j|,$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$ with sides parallel to coordinate axes. Since $E_j \subset [0,1]^n$, we get

$$\|\chi_{E_j}\|^*_{\mathcal{M}^u_1} = \sup_Q |Q|^{1/u-1}|Q \cap [0,1]^n \cap E_j|.$$

Thus it follows that we may suppose that Q intersects $[0,1]^n$. Assume that Q is such a cube. Translate Q to have a cube R of the same volume as Q so that $Q \cap [0,1]^n \subset R$ and $R \cap [0,1]^n$ is a cube. Then

$$\begin{aligned} |Q|^{1/u-1}|Q \cap [0,1]^n \cap E_j| &\leq |R|^{1/u-1}|R \cap [0,1]^n \cap E_j| \\ &\leq |R \cap [0,1]^n|^{1/u-1}|R \cap [0,1]^n \cap E_j|. \end{aligned}$$

So, we arrive at

$$\|\chi_{E_j}\|^*_{\mathcal{M}^u_1} = \sup_Q |Q|^{1/u-1}|Q \cap E_j|,$$

where the supremum is taken over all cubes Q in $[0,1]^n$ with sides parallel to coordinate axes.

Using the cubes $[0,1]^n$ and $[0,\alpha^j]^n$, we have

$$\|\chi_{E_j}\|^*_{\mathcal{M}^u_1} \geq \max\{\alpha^{jn/u}, 2^{jn}\alpha^{jn}\}.$$

To show the opposite estimate, we notice that if $|Q| \leq \alpha^{jn}$, then we have

$$|Q|^{1/u-1}|Q \cap E_j| \leq |Q|^{1/u} \leq \alpha^{jn/u}.$$

Assume that $\alpha^{kn} \leq |Q| \leq \alpha^{(k-1)n}$ for some $k \in \{1, \ldots, j\}$. We first observe that it follows by $E_{k-1} \supset E_j$ that

$$|Q \cap E_j| = |Q \cap E_{k-1} \cap E_j|.$$

Now note that E_{k-1} is made up of $2^{n(k-1)}$ disjoint compact cubes of volume $\alpha^{(k-1)n}$. In view of the size of Q, we know Q can intersect at most 2^n of them, say, $Q^1, \dots, Q^L$ with $L \le 2^n$. Then we have

$$|Q \cap E_j| = |Q \cap E_{k-1} \cap E_j| \le \sum_{l=1}^{L} |Q^l \cap E_j| = L 2^{n(j-k+1)} \alpha^{jn} \le 2^{n+n(j-k+1)} \alpha^{jn},$$

since $Q^l \cap E_j$ is made up of $2^{n(j-l)}$ disjoint cubes of volume α^{jn}. As a consequence,

$$|Q|^{1/u-1} |Q \cap E_j| \le \alpha^{kn/u-kn} 2^{n+n(j-k+1)} \alpha^{jn} \le 4^n \max\{\alpha^{jn/u}, 2^{jn}\alpha^{jn}\},$$

as required.

(iv) Since $\mathcal{M}^u_w = (\mathcal{M}^{u/w}_1)^{(w)}$ for every $1 \le u < w < \infty$, (iii) yields that for all integers $j \ge 0$, we have the equivalence

$$\|\chi_{F_j}\|_{\mathcal{M}^p_{q_0}} \asymp \max\{1, \alpha^{-jn/s+jn/s}\} = 1.$$

(v) Fix $j \in \mathbb{N}$. Let

$$G_k = \{x \in \mathbb{R}^n : \alpha^{j-k} x \in F_k\}, \quad k \in \{0, \dots, j\}.$$

Then $G_0 = [0, \alpha^{-j}]^n \supset G_1 \supset \cdots \supset G_j = F_j$. If $x \in G_j$, then the conclusion is clear since $F_j = G_j$. If $x \in \mathbb{R}^n \setminus G_0$, then the conclusion is again clear since the right-hand side is zero. Assume otherwise; $x \in G_k \setminus G_{k+1}$ for some $k \in \{0, \dots, j-1\}$. Let $H_k(x)$ be the connected component of G_k containing x. By translation we may assume that $x \in [0, \alpha^{-j+k}]^n = H_k(x)$. Then

$$\begin{aligned} M\chi_{F_j}(x) &\ge \frac{|F_j \cap [0, \alpha^{-j+k}]^n|}{|H_k(x)|} = 2^{(j-k)n} \alpha^{(j-k)n} \\ &\succ \frac{\chi_{F_k}(\alpha^{j-k} x)}{\|\chi_{F_k}(\alpha^{j-k}\,\cdot)\|^*_{L\mathcal{M}^s_1}} \succ \sum_{m=1}^{j} \frac{\chi_{F_m}(\alpha^{j-m} x)}{\|\chi_{F_m}(\alpha^{j-m}\,\cdot)\|^*_{L\mathcal{M}^s_1}}. \end{aligned}$$ □

Lemma 4.7 below is somewhat known. However we include a proof for completeness since we will use it later. In the proof we will use the Hardy–Littlewood maximal operator $M : L^1_{\mathrm{loc}} \to L^0$, which is defined by

$$Mf(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\, dy, \quad x \in \mathbb{R}^n.$$

Lemma 4.7. *The following statements are true*:

(i) *The Hardy–Littlewood maximal operator M is unbounded both on $\mathcal{M}^s_1$ and on $L\mathcal{M}^s_1$ for every $s \in [1, \infty)$.*

(ii) *The inclusions $L\mathcal{M}^p_q \hookrightarrow \mathrm{w}L\mathcal{M}^p_q$ and $\mathcal{M}^p_q \hookrightarrow \mathrm{w}\mathcal{M}^p_q$ are proper for all $1 \le q \le p < \infty$.*

Proof. (i) Let $s = 1$. It is a classical fact that if Mf is in L^1 for $f \in L^1_{\mathrm{loc}}$, then $f = 0$ a.e. and M cannot be bounded on L^1. Let $1 < s < \infty$. The statement that M is not bounded in $\mathcal{M}^s_1$ is an immediate consequence of Nakai's result [2008, Corollary 2.5, p. 205] on necessary and sufficient conditions for the boundedness of M on generalized Orlicz–Morrey spaces $L^{(\Phi,\phi)}$. In fact the Morrey space $\mathcal{M}^s_1$ is the

Orlicz–Morrey space generated by the Young function $\Phi(t) = t$ and the function $\phi(t) = t^{-1/s}$ for all $t \ge 0$. Recall that $\mathcal{M}_1^s = L^{(1,\lambda)}$ and $\mathrm{w}\mathcal{M}_1^s = L^{(1,\lambda)}_{\mathrm{weak}}$ if $\lambda = n - n/s$.

The necessary condition $\Phi \in \nabla_2$ (i.e., that there exists $k \ge 1$ such that $\Phi(t) \le \frac{1}{2k}\Phi(kt)$ for all $t > 0$) is not satisfied. A careful analysis of the proof of the mentioned result of Nakai based on a key observation in [Nakai 2008, Lemma 4.10] also gives that the operator M is not bounded in $L\mathcal{M}_1^s$. We point out here that by using the sequence $\{F_j\}_{j=1}^\infty$ defined in the proof of Theorem 4.8 below, we can also disprove that M is bounded on $L\mathcal{M}_1^s$ for all $s > 1$. We include a short and transparent proof of our own, for the reader's convenience. Let $\{F_j\}_{j=0}^\infty$ be the sequence constructed in the proof of Lemma 4.6. Let $j \in \mathbb{N}$ be arbitrary. Then from Lemma 4.6(v), we get

$$|[-\alpha^{-j}, \alpha^{-j}]^n\|^{1/s-1} \int_{[-\alpha^{-j},\alpha^{-j}]^n} M\chi_{F_j}(x)\,dx \succ \sum_{k=1}^{j} \frac{\|\chi_{F_k}(\alpha^{j-k}\cdot)\|^*_{\mathcal{M}_1^s}}{\|\chi_{F_k}(\alpha^{j-k}\cdot)\|^*_{\mathcal{M}_1^s}} = j.$$

This proves that M is not bounded on $L\mathcal{M}_1^s$.

(ii) We apply the fact that M is bounded from $\mathcal{M}_1^s$ to $\mathrm{w}\mathcal{M}_1^s$, see, e.g,. [Nakai 2008, Corollary 6.3, p. 207], and also that M is bounded from $L\mathcal{M}_1^s$ to $\mathrm{w}L\mathcal{M}_1^s$. The second fact easily follows from [Burenkov and Guliyev 2004, Lemma 10], which yields that there exists a positive constant C such that for all $f \in L^1_{\mathrm{loc}}$ and $r > 0$,

$$\sup_{\lambda>0} \lambda|\{x \in B(r) : Mf(x) > \lambda\}| \le Cr^n \int_r^\infty \left(\frac{1}{t^{n+1}} \int_{B(t)} |f(x)|\,dx\right) dt.$$

If $f \in L\mathcal{M}_1^s$, then simple calculus yields that

$$\sup_{\lambda>0} \lambda|\{x \in B(r) : Mf(x) > \lambda\}| \le Cr^n \int_r^\infty t^{-n/s-1} \|f\|_{L\mathcal{M}_1^s}\,dt = Cr^{n-n/s}\|f\|_{L\mathcal{M}_1^s}.$$

As a result we get

$$\|Mf\|_{\mathrm{w}\mathcal{M}_1^s} = \sup_{\lambda,r>0} |B(r)|^{1/s-1}|\{x \in B(r) : Mf(x) > \lambda\}| \le C\|f\|_{L\mathcal{M}_1^s},$$

as required. Let $f_0 = \lim_{j\to\infty} \chi_{F_j} \in L\mathcal{M}_1^s$, where each F_j is as in Lemma 4.6. From the proof of (i), Lemma 4.6(iv) and (v), $Mf_0 \in \mathrm{w}L\mathcal{M}_1^s \setminus L\mathcal{M}_1^s$. If $1 < q < p < \infty$, then we let $s := p/q$ and define $g_0 = (Mf_0)^{1/q}$. Combining these observations, we get $g_0 \in (\mathrm{w}L\mathcal{M}_1^s)^{(q)} = \mathrm{w}L\mathcal{M}_q^p$ and $g_0 \notin (L\mathcal{M}_1^s)^{(q)} = L\mathcal{M}_q^p$ and so $L\mathcal{M}_q^p \ne \mathrm{w}L\mathcal{M}_q^p$. Similarly we can show that $\mathcal{M}_q^p \ne \mathrm{w}\mathcal{M}_q^p$.

Since the case $p = q$ reduces to the well-known fact that $L^p \ne L^{p,\infty}$, the proof is complete. □

With all these preliminary results, we are now ready to state our main result of this section, which shows that Cwikel's conjecture is valid in a wide class of Morrey spaces.

Theorem 4.8. *Let $1 \le q_0 < p_0 < \infty$ and $1 \le q_1 < p_1 \le \infty$ with $q_0 \ne q_1$. Then for any $\theta \in (0, 1)$ the upper complex interpolation space $[\mathcal{M}_{q_0}^{p_0}, \mathcal{M}_{q_1}^{p_1}]^\theta$ is not a K-monotone couple with respect to $(\mathcal{M}_{q_0}^{p_0}, \mathcal{M}_{q_1}^{p_1})$, and so $(\mathcal{M}_{q_0}^{p_0}, \mathcal{M}_{q_1}^{p_1})$ is not a Calderón–Mityagin couple.*

Proof. It follows from Proposition 4.2 and Theorem 4.3 that it is suffices to prove that for every $\theta \in (0, 1)$ there exists $f \in [\mathcal{M}_{q_0}^{p_0}, \mathcal{M}_{q_1}^{p_1}]^\theta$ such that $t^\theta \asymp K(t, f; \mathcal{M}_{q_0}^{p_0}, \mathcal{M}_{q_1}^{p_1})$.

We will distinguish three cases:

Case 1: $p_0, p_1 < \infty$ with $p_0 \neq p_1$. Let $1/p = (1-\theta)/p_0 + \theta/p_1$ and $1/q = (1-\theta)/q_0 + \theta/q_1$. Consider the functions f, g_0 and g_1 given by $f(x) = |x|^{-n/p}$ and $g_j(x) = |x|^{-n/p_j}$ for all $x \in \mathbb{R}^n \setminus \{0\}$, for $j \in \{0, 1\}$. By the preceding discussion in introduction of this section, it follows that $g_0 \in \mathcal{M}^{p_0}_{q_0}$ and $g_1 \in \mathcal{M}^{p_1}_{q_1}$.

Since Morrey spaces enjoy the Fatou property, it follows from Theorem 3.7 that

$$(\mathcal{M}^{p_0}_{q_0})^{1-\theta}(\mathcal{M}^{p_1}_{q_1})^{\theta} \cong [\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}]^{\theta}.$$

Combining the above facts with $f = g_0^{1-\theta} g_1^{\theta}$, we get $f \in [\mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}]^{\theta}$.

We claim that

$$t^{\theta} \prec K(t, f; \mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}), \quad t > 0.$$

First we notice that $\mathcal{M}^{p_j}_{q_j} \hookrightarrow \mathcal{M}^{p_j}_{1}$ with the norm of the inclusion map 1 for $j \in \{0, 1\}$:

$$\begin{aligned} K(t, f; \mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}) &\geq K(t, f; \mathcal{M}^{p_0}_{1}, \mathcal{M}^{p_1}_{1}) = \inf\{\|f_0\|_{\mathcal{M}^{p_0}_1} + t\,\|f_1\|_{\mathcal{M}^{p_1}_1} : f_0 + f_1 = f\} \\ &\geq \inf_{f_0+f_1=f} \sup_{x\in\mathbb{R}^n, r>0} \int_{B(x,r)} \big(|B(x,r)|^{1/p_0-1}|f_0(y)| + t\,|B(x,r)|^{1/p_1-1}|f_1(y)|\big)\,dy. \end{aligned}$$

Thus applying the formula which was explained at the beginning of this section,

$$|B(0,r)|^{1/s-1} \int_{B(0,r)} |x|^{-n/s}\,dx = C(s),$$

we obtain the following estimates for all $t > 0$:

$$\begin{aligned} &K(t, f; \mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}) \\ &\quad\geq \inf_{f_0+f_1=f} \sup_{x\in\mathbb{R}^n, r>0} \min\{|B(x,r)|^{1/p_0-1}, t\,|B(x,r)|^{1/p_1-1}\} \int_{B(x,r)} (|f_0(y)| + |f_1(y)|)\,dy \\ &\quad\geq \sup_{x\in\mathbb{R}^n, r>0} \min\{|B(x,r)|^{1/p_0-1}, t\,|B(x,r)|^{1/p_1-1}\} \int_{B(x,r)} |f(y)|\,dy \\ &\quad\geq \sup_{r>0} \min\{|B(r)|^{1/p_0-1}, t\,|B(r)|^{1/p_1-1}\} \int_{B(r)} |f(y)|\,dy \\ &\quad\asymp \sup_{r>0} \min\{|B(r)|^{1/p_0-1/p}, t\,|B(r)|^{1/p_1-1/p}\}. \end{aligned}$$

If we calculate the last expression, we obtain

$$K(t, f; \mathcal{M}^{p_0}_{q_0}, \mathcal{M}^{p_1}_{q_1}) \asymp \sup_{r>0} \min\left\{1, \frac{t}{|B(r)|^{1/p_0-1/p_1}}\right\} |B(r)|^{\theta/p_0-\theta/p_1} = \sup_{s>0} \min\left\{1, \frac{t}{s}\right\} s^{\theta} = t^{\theta}.$$

Case 2: $p_1 = \infty$. We will use [Cwikel and Gulisashvili 2000, Lemma 6] from which it follows that if X is a Banach function lattice on a σ-finite complete measure space (Ω, Σ, μ) and f is a nonnegative function in $(X, L^{\infty})_{\theta,\infty}$ for some $\theta \in (0, 1)$, then for each $y > 0$ we have $\chi_{E(y,f)} \in X$ and

$$\|f\chi_{E(y,f)}\|_X \leq K(\|\chi_{E(y,f)}\|_X, f; X, L^{\infty}), \tag{4-2}$$

where $\chi_{E(y,f)} = \{x \in \Omega : f(x) > y\}$.

At this stage we observe that it follows from Corollary 3.10 and Proposition 4.1 that for the function f given by $f(x) = |x|^{-n(1-\theta)/p_0}$ for all $x \in \mathbb{R}^n \setminus \{0\}$ with $f(0) := 0$ we have

$$f \in (\mathcal{M}^{p_0}_{q_0})^{1/(1-\theta)} = [\mathcal{M}^{p_0}_{q_0}, L^\infty]^\theta \hookrightarrow (\mathcal{M}^{p_0}_{q_0}, L^\infty)_{\theta,\infty}.$$

Now we apply the above estimate (4-2) of the K-functional for the couple $(\mathcal{M}^{p_0}_{q_0}, L^\infty)$. First notice that it is easy to check that

$$\|\chi_{B(r)}\|_{\mathcal{M}^{p_0}_{q_0}} = |B(r)|^{1/p_0} = (v_n r)^{n/p_0}, \quad r > 0,$$

and that, for function f shown above, we have

$$E(y, f) = \{x \in \mathbb{R}^n : f(x) > y\} = B(1/y^{p_0/n(1-\theta)}).$$

For $t > 0$ let us take $y := y(t) = (v_n^{n/p_0} t^{-1})^{1-\theta}$. Then we get $\|\chi_{E(y,f)}\|_{\mathcal{M}^{p_0}_{q_0}} = t$. Hence, we obtain

$$K(t, f; \mathcal{M}^{p_0}_{q_0}, L^\infty) \ge \|f\chi_{E(y,f)}\|_{\mathcal{M}^{p_0}_{q_0}} \ge y\|\chi_{E(y,f)}\|_{\mathcal{M}^{p_0}_{q_0}} = (v_n^{n/p_0} t^{-1})^{1-\theta} t = v_n^{n(1-\theta)/p_0} t^\theta.$$

Case 3: $1 \le q_0 < q_1 < p := p_0 = p_1 < \infty$. First observe that if $\vec{X} = (X_0, X_1)$ is a Banach couple such that the norm of the inclusion map $X_1 \hookrightarrow X_0$ is less than or equal to 1, then $K(t, x; \vec{X}) = \|x\|_{X_0}$ for every $t \ge 1$. We learn from Hölder's inequality that the couple $(X_0, X_1) := (\mathcal{M}^p_{q_0}, \mathcal{M}^p_{q_1})$ enjoys this property.

To finish we will apply Lemmas 4.5(ii) and 4.6. To do this we will use a sequence $\{F_j\}_{j\ge 0}$ of Lebesgue measurable subsets in $\mathbb{R}^n$, constructed in the proof of Lemma 4.6, which satisfies the conditions of the Lemma 4.5(ii). As a result,

$$g_\theta := \sum_{j=0}^\infty r^{j\theta} \chi_{F_j} \in (\mathcal{M}^p_{q_0})^{1-\theta} (\mathcal{M}^p_{q_1})^\theta$$

and

$$K(t, g_\theta; \mathcal{M}^p_{q_0}, \mathcal{M}^p_{q_1}) \asymp t^\theta, \quad t \in (0, 1]. \qquad \square$$

We conclude this section with the following result:

Proposition 4.9. *Assume there exists $\theta \in (0, 1)$ such that the inclusion $[\mathcal{M}^{s_0}_1, \mathcal{M}^{s_1}_1]^\theta \hookrightarrow (\mathcal{M}^{s_0}_1, \mathcal{M}^{s_1}_1)_{\theta,\infty}$ is proper for every $s_0, s_1 \in (1, \infty)$ with $s_0 \ne s_1$. Then $(\mathcal{M}^{p_0}_q, \mathcal{M}^{p_1}_q)$ is not a Calderón–Mityagin couple for all $p_0, p_1 \in (1, \infty)$ with $p_0 \ne p_1$ and all $1 \le q < \min\{p_0, p_1\}$.*

Proof. It is easy to verify that for any couple (X_0, X_1) of Banach lattices and every $1 < q < \infty$, we have

$$(X_0^{(q)})^{1-\theta} (X_1^{(q)})^\theta \cong (X_0^{1-\theta} X_1^\theta)^{(q)}.$$

Thus thanks to the well-known equivalence

$$K(t, f; X_0^{(q)}, X_1^{(q)}) \asymp K(t^q, |f|^q; X_0, X_1)^{1/q}, \quad f \in X_0^{(q)} + X_1^{(q)},$$

up to equivalence of norms, we get

$$(X_0^{(q)}, X_1^{(q)})_{\theta,\infty} = (X_0, X_1)^{(q)}_{\theta,\infty}.$$

These formulas combined with the Fatou property of $\mathcal{M}_q^{p_j} = (\mathcal{M}^{p_j/q})^{(q)}$ imply for $(X_0, X_1) := (\mathcal{M}_1^{p_0/q}, \mathcal{M}_1^{p_1/q})$ that the inclusion

$$[\mathcal{M}_q^{p_0}, \mathcal{M}_q^{p_1}]^\theta \hookrightarrow (\mathcal{M}_q^{p_0}, \mathcal{M}_q^{p_1})_{\theta,\infty}$$

is proper. Since $p_0 \neq p_1$, it follows from the proof of Theorem 4.8 that there exists $f \in [\mathcal{M}_q^{p_0}, \mathcal{M}_q^{p_1}]^\theta$ such that

$$K(t, f; \mathcal{M}_q^{p_0}, \mathcal{M}_q^{p_1}) \asymp t^\theta.$$

The required statement now follows from Proposition 4.2. □

To conclude this section, we note that it is natural to ask whether $[\mathcal{M}_1^{p_0}, \mathcal{M}_1^{p_1}]^\theta \neq (\mathcal{M}_1^{p_0}, \mathcal{M}_1^{p_1})_{\theta,\infty}$ in the set-theoretical sense for all $\theta \in (0, 1)$ and $p_0, p_1 \in (1, \infty)$ with $p_0 \neq p_1$.

5. On Calderón–Mityagin couples of local Morrey spaces

We now study Calderón–Mityagin couples of local Morrey spaces. The following interpolation results will play a key role. We proceed in a couple of simple steps, which seem independently interesting themselves.

Lemma 5.1. *If* $1 \le q_0 < p_0 < \infty$ *and* $1/p = (1-\theta)/p_0$ *and* $1/q = (1-\theta)/q_0$ *for* $\theta \in (0, 1)$, *then the following formulas are true*:

(i) $[L\mathcal{M}_{q_0}^{p_0}, L^\infty]^\theta \cong (L\mathcal{M}_{q_0}^{p_0})^{1-\theta}(L^\infty)^\theta \cong L\mathcal{M}_q^p$.

(ii) $(L\mathcal{M}_{q_0}^{p_0}, L^\infty)_{\theta,\infty} = \mathrm{w}L\mathcal{M}_q^p$.

If we reexamine the proof of Theorem 5.2, we obtain the proof of Lemma 5.1 as a special case of Theorem 5.2. However, we give a proof using what we have shown.

Proof. Since $L\mathcal{M}_{q_0}^{p_0}$ has the Fatou property, statement (i) follows from Corollary 3.10. In a similar fashion to the proof of Theorem 4.3, we explain that (ii) follows by [Cwikel and Gulisashvili 2000, Theorem 4]. □

We now handle the case $q_1 < p_1 < \infty$ by a different method.

Theorem 5.2. *Let* $1 \le q_j < p_j < \infty$ *for* $j \in \{0, 1\}$ *and* $\theta \in (0, 1)$. *If* $1/p = (1-\theta)/p_0 + \theta/p_1$ *and* $1/q = (1-\theta)/q_0 + \theta/q_1$, *then we have the following properties*:

(i) $[L\mathcal{M}_{q_0}^{p_0}, L\mathcal{M}_{q_1}^{p_1}]^\theta = L\mathcal{M}_q^p$.

(ii) *If* $q = q_0 = q_1$, *then* $(L\mathcal{M}_{q_0}^{p_0}, L\mathcal{M}_{q_1}^{p_1})_{\theta,\infty} = L\mathcal{M}_q^p$.

(iii) *If* $q_0 \neq q_1$, *then* $(L\mathcal{M}_{q_0}^{p_0}, L\mathcal{M}_{q_1}^{p_1})_{\theta,\infty} = \mathrm{w}L\mathcal{M}_q^p$.

Proof. We will use an equivalent norm on local Morrey spaces. It is obvious that for any $1 \le q \le p < \infty$ the functional $\|\cdot\|'$ defined on $L\mathcal{M}_q^p$ by

$$\|f\|' = \sup_{k\in\mathbb{Z}} |B(2^k)|^{1/p-1/q} \left(\int_{B(2^k)} |f(y)|^q \, dy \right)^{1/q}, \quad f \in L\mathcal{M}_q^p,$$

is a norm equivalent to the original norm on $L\mathcal{M}_q^p$.

We claim that if in addition $q \neq p$, then the formula

$$\|f\|^* := \sup_{k\in\mathbb{Z}} |B(2^k)|^{1/p-1/q}\left(\int_{B(2^k)\setminus B(2^{k-1})} |f(y)|^q\,dy\right)^{1/q}$$

also gives an equivalent norm on $L\mathcal{M}_q^p$.

From the definition it is clear that $\|\cdot\|^* \geq \|\cdot\|'$. To prove the converse inequality, we observe that

$$\left(\int_{B(2^j)\setminus B(2^{j-1})} |f(y)|^q\,dy\right)^{1/q} \leq |B(2^j)|^{-1/p+1/q}\|f\|^*, \quad j \in \mathbb{Z}.$$

If we use the triangle inequality, we get that for each $k \in \mathbb{Z}$

$$\begin{aligned}\left(\int_{B(2^k)} |f(y)|^q\,dy\right)^{1/q} &= \left\|\sum_{j=-\infty}^{k} f\chi_{B(2^j)\setminus B(2^{j-1})}\right\|_q \\ &\leq \left(\sum_{j=0}^{\infty} 2^{jn/p-jn/q}\right)|B(2^k)|^{-1/p+1/q}\|f\|^* \\ &= c_{p,q}|B(2^k)|^{-1/p+1/q}\|f\|^*.\end{aligned}$$

This implies that $\|f\|' \leq c_{p,q}\|f\|^*$ for all $f \in L\mathcal{M}_q^p$.

In what follows we will need some general interpolation formulas. In order to state them we introduce some additional notation. For a given sequence $\{\vec{A}^k\}_{k\in\mathbb{Z}}$ of Banach couples, where $\vec{A}^k = (A_0^k, A_1^k)$, we define a Banach couple

$$\ell_\infty(\{\vec{A}_k\}) := (\ell_\infty(\{A_0^k\}_{k\in\mathbb{Z}}), \ell_\infty(\{A_1^k\}_{k\in\mathbb{Z}})),$$

where $\ell_\infty(\{A_j^k\}_{k\in\mathbb{Z}})$ for each $j \in \{0,1\}$ is a Banach space of all bounded sequences $\{a_j^k\}_{k\in\mathbb{Z}} \in \prod_{k\in\mathbb{Z}} A_j^k$ endowed with the uniform norm.

We omit the standard proofs of the following formulas:

$$[\ell_\infty(\{A_0^k\}_{k\in\mathbb{Z}}), \ell_\infty(\{A_1^k\}_{k\in\mathbb{Z}})]^\theta = \ell_\infty(\{[A_0^k, A_1^k]^\theta\}_{k\in\mathbb{Z}}),$$
$$(\ell_\infty(\{A_0^k\}_{k\in\mathbb{Z}}), \ell_\infty(\{A_1^k\}_{k\in\mathbb{Z}}))_{\theta,\infty} = \ell_\infty(\{(A_0^k, A_1^k)_{\theta,\infty}\}_{k\in\mathbb{Z}}).$$

We consider a sequence $\{(A_0^k, A_1^k)\}_{k\in\mathbb{Z}}$ given by $A_j^k := w_j^k L^{q_j}$, with $w_j^k = |B(2^k)|^{1/p_j-1/q_j}$ for each $k \in \mathbb{Z}$ and $j \in \{0,1\}$, and endowed with norms $\|f\|_{A_j^k} = w_j^k\|f\|_{L^{q_j}}$.

Since $[L^{q_0}, L^{q_1}]^\theta = L^q$ if $q_0 \neq q_1$, $(L^{q_0}, L^{q_1})_{\theta,\infty} = L^{q,\infty}$ and $(L^q, L^q)_{\theta,\infty} = L^q$, then the above vector-valued formulas easily yield

$$[\ell_\infty(\{A_0^k\}_{k\in\mathbb{Z}}), \ell_\infty(\{A_1^k\}_{k\in\mathbb{Z}})]^\theta = \ell_\infty(\{[w_0^k L^{q_0}, w_1^k L^{q_1}]^\theta\}_{k\in\mathbb{Z}}) = \ell_\infty(\{w^k L^q\}_{k\in\mathbb{Z}}),$$
$$(\ell_\infty(\{A_0^k\}_{k\in\mathbb{Z}}), \ell_\infty(\{A_1^k\}_{k\in\mathbb{Z}}))_{\theta,\infty} = \ell_\infty(\{(w_0^k L^q, w_1^k L^q)_{\theta,\infty}\}_{k\in\mathbb{Z}}) = \ell_\infty(\{w^k L^{q,\infty}\}_{k\in\mathbb{Z}}),$$
$$(\ell_\infty(\{A_0^k\}_{k\in\mathbb{Z}}), \ell_\infty(\{A_1^k\}_{k\in\mathbb{Z}}))_{\theta,\infty} = \ell_\infty(\{(w_0^k L^{q_0}, w_1^k L^{q_1})_{\theta,\infty}\}_{k\in\mathbb{Z}}) = \ell_\infty(\{w^k L^{q,\infty}\}_{k\in\mathbb{Z}}),$$

where $\{w^k\}_{k\in\mathbb{Z}} = \{|B(2^k)|^{1/p-1/q}\}$ for each $k \in \mathbb{Z}$ and

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

It is easy to check that $\mathrm{w}L\mathcal{M}^p_q = \ell_\infty(\{w^k L^{q,\infty}\})$ with equality of quasinorms and so the last formula can take the form

$$(\ell_\infty(\{A_0^k\}_{k\in\mathbb{Z}}), \ell_\infty(\{A_1^k\}_{k\in\mathbb{Z}}))_{\theta,\infty} = \mathrm{w}L\mathcal{M}^p_q, \quad q_0 \neq q_1.$$

Now by the discussion before about equivalent norms on Morrey spaces, it follows that

$$\|f\|_{L\mathcal{M}^p_q} \asymp \sup_{k\in\mathbb{Z}} |B(2^k)|^{1/p-1/q} \left(\int_{B(2^k)\setminus B(2^{k-1})} |f(y)|^q \, dy \right)^{1/q}, \quad f \in L\mathcal{M}^p_q.$$

This equivalence implies that operators U and V given by

$$Uf = \{f\chi_{B(2^k)\setminus B(2^{k-1})}\}_{k\in\mathbb{Z}}, \quad f \in L\mathcal{M}^{p_0}_{q_0} + L\mathcal{M}^{p_1}_{q_1},$$
$$V(\{f_k\}_{k\in\mathbb{Z}}) = \sum_{k\in\mathbb{Z}} f_k \chi_{B(2^k)\setminus B(2^{k-1})}, \quad \{f_k\}_{k\in\mathbb{Z}} \in \ell_\infty(\{A_0^k\}_{k\in\mathbb{Z}}) + \ell_\infty(\{A_1^k\}_{k\in\mathbb{Z}})$$

are bounded between the Banach couples

$$U : (L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}) \to (\ell_\infty(\{A_0^k\}_{k\in\mathbb{Z}}), \ell_\infty(\{A_1^k\}_{k\in\mathbb{Z}})),$$
$$V : (\ell_\infty(\{A_0^k\}_{k\in\mathbb{Z}}), \ell_\infty(\{A_1^k\}_{k\in\mathbb{Z}})) \to (L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}).$$

We conclude by the vector-valued interpolation formulas shown above that

$$U : [L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}]^\theta \to \ell_\infty(\{|B(2^k)|^{1/p-1/q} L^q\})$$

is bounded. In particular this yields the continuous inclusion

$$[L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}]^\theta \hookrightarrow L\mathcal{M}^p_q.$$

The boundedness of an operator V from $\ell_\infty(\{|B(2^k)|^{1/p-1/q} L^q\}_{k\in\mathbb{Z}})$ into $[L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}]^\theta$ yields the reverse continuous inclusion

$$L\mathcal{M}^p_q \hookrightarrow [L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}]^\theta$$

and so $[L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}]^\theta = L\mathcal{M}^p_q$, as required.

Similarly we obtain the remaining formulas and this completes the proof. □

Remark 5.3. We notice that using maps U and V defined in the proof of Theorem 5.2 we easily conclude the following equivalence for the K-functional of local Morrey couples: if $1 \le q_j < p_j < \infty$ for $j \in \{0,1\}$, then for all $f \in L\mathcal{M}^{p_0}_{q_0} + L\mathcal{M}^{p_1}_{q_1}$ and $t > 0$,

$$K(t, f; L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}) \asymp K\big(t, \{f\chi_{B(2^k)\setminus B(2^{k-1})}\}_{k\in\mathbb{Z}}; \ell_\infty(\{w_0^k L^{q_0}\}_{k\in\mathbb{Z}}), \ell_\infty(\{w_1^k L^{q_1}\}_{k\in\mathbb{Z}})\big)$$
$$\asymp \sup_{k\in\mathbb{Z}} K(t, f\chi_{B(2^k)\setminus B(2^{k-1})}; w_0^k L^{q_0}, w_1^k L^{q_1}),$$

where $w_j^k = |B(2^k)|^{1/p_j-1/q_j}$ for each $k \in \mathbb{Z}$ and $j \in \{0,1\}$.

We apply the above results to study Calderón–Mityagin couples of local Morrey spaces. Our main result of this section shows that Cwikel's conjecture is valid in the class of local Morrey spaces.

Theorem 5.4. *Let* $1 \le q_0 < p_0 < \infty$ *and* $1 \le q_1 < p_1 < \infty$. *The following are equivalent*:

(i) $q_0 = q_1$.

(ii) $(L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1})$ *has the universal* K*-property.*

(iii) $(L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1})$ *is a Calderón–Mityagin couple.*

Proof. (i) $\Rightarrow$ (ii). We invoke the following result from [Cwikel and Peetre 1981]: for any Banach couple (A_0, A_1) and any numbers $\theta_0, \theta_1 \in (0,1)$ the couple of interpolation spaces $((A_0, A_1)_{\theta_0,\infty}, (A_0, A_1)_{\theta_1,\infty})$ has the universal K-property. To establish that the couple $(L\mathcal{M}^{p_0}_{q}, L\mathcal{M}^{p_1}_{q})$ falls under this scope, we fix $0 < \varepsilon < \min\{p_0 - q, p_1 - q\}$ and set $u_0 = p_0 - \varepsilon$, $u_1 = p_0 + \varepsilon$, $v_0 = p_1 - \varepsilon$ and $v_1 = p_1 + \varepsilon$. Then for $\theta_0, \theta_1 \in (0,1)$ given by $\theta_0 = \frac{1}{2}(1 + \varepsilon/p_0)$ and $\theta_1 = \frac{1}{2}(1 + \varepsilon/p_1)$ we have

$$\frac{1}{p_0} = \frac{1-\theta_0}{u_0} + \frac{\theta_0}{u_1}, \qquad \frac{1}{p_1} = \frac{1-\theta_1}{v_0} + \frac{\theta_1}{v_1}.$$

If $q = q_0 = q_1$, then it follows from Theorem 5.2 that

$$(L\mathcal{M}^{u_0}_{q}, L\mathcal{M}^{u_1}_{q})_{\theta_0,\infty} = L\mathcal{M}^{p_0}_{q}, \quad (L\mathcal{M}^{v_0}_{q}, L\mathcal{M}^{v_1}_{q})_{\theta_1,\infty} = L\mathcal{M}^{p_1}_{q}.$$

(ii) $\Rightarrow$ (iii). This is obvious.

(iii) $\Rightarrow$ (i). We consider the contrapositive. Suppose that $q_0 \ne q_1$. We have two cases:

Case 1: $p_0 \ne p_1$. Then as before the function f defined by $f(x) = |x|^{-n/p}$ for almost all $x \in \mathbb{R}^n$ is in $L\mathcal{M}^{p}_{q}$, and

$$K(t, f; L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}) \asymp t^{\theta}, \quad t > 0.$$

Meanwhile, we conclude from Theorem 5.2 that

$$f \in [L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}]^{\theta} = L\mathcal{M}^{p}_{q}$$

and so it follows from Lemma 4.7 and Theorem 5.2 that the inclusion

$$[L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}]^{\theta} \hookrightarrow (L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1})_{\theta,\infty}$$

is proper. Applying Proposition 4.2(i), we deduce that $(L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1})$ is not a Calderón–Mityagin couple and so we get a contradiction.

Case 2: $p := p_0 = p_1$. Since the cube $[0,1]^n$ appears in the definition of the sequence $\{E_j\}_{j=0}^{\infty}$ constructed in the proof of Lemma 4.6, the same conclusion as in the case of Morrey spaces yields that for $\{F_j\}_{j\ge 0} := \{\alpha^{-j} E_j\}_{j \ge 0}$ with the same $\alpha \in \left(0, \frac{1}{2}\right)$, we get

$$\|\chi_{F_j}\|_{L\mathcal{M}^{u}_{1}} \asymp \max\{1, \alpha^{-jn/u + jn} 2^{-jn}\} = \max\{1, \alpha^{-jn/u + jn/s}\}, \quad j \ge 0.$$

Then as we did in the proof of Theorem 4.8, we apply Lemma 4.5 for the couple $(X_0, X_1) := (L\mathcal{M}^{p}_{q_0}, L\mathcal{M}^{p}_{q_1})$ to find $f \in [L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}]^{\theta} = L\mathcal{M}^{p}_{q}$ such that

$$K(t, f; L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1}) \asymp t^{\theta}, \quad t \in (0,1].$$

Then similar to the above proof of Case 1, applying Proposition 4.2(ii), we deduce that $(L\mathcal{M}^{p_0}_{q_0}, L\mathcal{M}^{p_1}_{q_1})$ is not a Calderón–Mityagin couple. But this is a contradiction. □

Acknowledgement

We thank the referee for the helpful comments that led to improvement of the paper.

References

[Adams and Xiao 2004] D. R. Adams and J. Xiao, "Nonlinear potential analysis on Morrey spaces and their capacities", *Indiana Univ. Math. J.* **53**:6 (2004), 1629–1663. MR Zbl

[Adams and Xiao 2012] D. R. Adams and J. Xiao, "Morrey spaces in harmonic analysis", *Ark. Mat.* **50**:2 (2012), 201–230. MR Zbl

[Bergh 1979] J. Bergh, "On the relation between the two complex methods of interpolation", *Indiana Univ. Math. J.* **28**:5 (1979), 775–778. MR Zbl

[Bergh and Löfström 1976] J. Bergh and J. Löfström, *Interpolation spaces: an introduction*, Grundlehren der Math. Wissenschaften **223**, Springer, 1976. MR Zbl

[Beurling 1964] A. Beurling, "Construction and analysis of some convolution algebras", *Ann. Inst. Fourier* (*Grenoble*) **14**:2 (1964), 1–32. MR Zbl

[Brudnyi and Krugljak 1991] Y. A. Brudnyi and N. Y. Krugljak, *Interpolation functors and interpolation spaces, I*, North-Holland Math. Library **47**, North-Holland, Amsterdam, 1991. MR Zbl

[Brudnyi, Y. A. and Kruglyak, N. Y. 1981] Y. A. Brudnyĭ and N. Y. Kruglyak, "Functors of real interpolation", *Dokl. Akad. Nauk SSSR* **256**:1 (1981), 14–17. In Russian; translated in *Soviet Math. Dokl.* **23**:1 (1981), 5–8. MR

[Burenkov and Guliyev 2004] V. I. Burenkov and H. V. Guliyev, "Necessary and sufficient conditions for boundedness of the maximal operator in local Morrey-type spaces", *Studia Math.* **163**:2 (2004), 157–176. MR Zbl

[Calderón 1964] A. P. Calderón, "Intermediate spaces and interpolation: the complex method", *Studia Math.* **24** (1964), 113–190. MR Zbl

[Calderón 1966] A. P. Calderón, "Spaces between L^1 and L^∞ and the theorem of Marcinkiewicz", *Studia Math.* **26** (1966), 273–299. MR Zbl

[Campanato and Murthy 1965] S. Campanato and M. K. V. Murthy, "Una generalizzazione del teorema di Riesz–Thorin", *Ann. Scuola Norm. Sup. Pisa* (3) **19**:1 (1965), 87–100. MR Zbl

[Cwikel 1981] M. Cwikel, "Monotonicity properties of interpolation spaces, II", *Ark. Mat.* **19**:1 (1981), 123–136. MR Zbl

[Cwikel 1984] M. Cwikel, "K-divisibility of the K-functional and Calderón couples", *Ark. Mat.* **22**:1 (1984), 39–62. MR Zbl

[Cwikel and Gulisashvili 2000] M. Cwikel and A. Gulisashvili, "Interpolation on families of characteristic functions", *Studia Math.* **138**:3 (2000), 209–224. MR Zbl

[Cwikel and Nilsson 2003] M. Cwikel and P. G. Nilsson, "Interpolation of weighted Banach lattices", pp. 1–105 in Mem. Amer. Math. Soc. **787**, Amer. Math. Soc., Providence, RI, 2003. MR Zbl

[Cwikel and Peetre 1981] M. Cwikel and J. Peetre, "Abstract K and J spaces", *J. Math. Pures Appl.* (9) **60**:1 (1981), 1–49. MR Zbl

[Fremlin 1974] D. H. Fremlin, *Topological Riesz spaces and measure theory*, Cambridge Univ. Press, 1974. MR Zbl

[García-Cuerva and Herrero 1994] J. García-Cuerva and M.-J. L. Herrero, "A theory of Hardy spaces associated to the Herz spaces", *Proc. London Math. Soc.* (3) **69**:3 (1994), 605–628. MR Zbl

[Kalton 1993] N. J. Kalton, "Calderón couples of rearrangement invariant spaces", *Studia Math.* **106**:3 (1993), 233–277. MR Zbl

[Kantorovich and Akilov 1982] L. V. Kantorovich and G. P. Akilov, *Functional analysis*, 2nd ed., Pergamon Press, Oxford, 1982. MR Zbl

[Krein et al. 1982] S. G. Kreĭn, Y. I. Petunin, and E. M. Semenov, *Interpolation of linear operators*, Trans. Math. Monographs **54**, Amer. Math. Soc., Providence, R.I., 1982. MR Zbl

[Lemarié-Rieusset 2012] P. G. Lemarié-Rieusset, "The role of Morrey spaces in the study of Navier–Stokes and Euler equations", *Eurasian Math. J.* **3**:3 (2012), 62–93. MR Zbl

[Lemarié-Rieusset 2013] P. G. Lemarié-Rieusset, "Multipliers and Morrey spaces", *Potential Anal.* **38**:3 (2013), 741–752. Correction in **41**:4 (2014), 1359–1362. MR Zbl

[Lozanovskii 1969] G. Y. Lozanovskii, "On some Banach lattices", *Sibirsk. Mat. Ž.* **10**:3 (1969), 584–599. In Russian; translated in *Siberian Math. J.* **10**:3 (1969), 419–431. MR Zbl

[Lozanovskii 1972] G. Y. Lozanovskii, "A remark on a certain interpolation theorem of Calderón", *Funkcional. Anal. i Priložen.* **6**:4 (1972), 89–90. In Russian; translated in *Funct. Anal. Appl.* **6**:4 (1972), 333–334. MR

[Lozanovskii 1978] G. Y. Lozanovskii, "Transformations of ideal Banach spaces by means of concave functions", pp. 122–148 in *Qualitative and approximate methods for the investigation of operator equations, III*, edited by P. P. Zabreiko, Yaroslavl Gos. Univ., 1978. In Russian. MR Zbl

[Mastyło and Ovchinnikov 1997] M. Mastyło and V. I. Ovchinnikov, "On the relation between complex and real methods of interpolation", *Studia Math.* **125**:3 (1997), 201–218. MR Zbl

[Mastyło and Sinnamon 2017] M. Mastyło and G. Sinnamon, "Calderón–Mityagin couples of Banach spaces related to decreasing functions", *J. Funct. Anal.* **272**:11 (2017), 4460–4482. MR Zbl

[Mazzucato 2003] A. L. Mazzucato, "Besov–Morrey spaces: function space theory and applications to non-linear PDE", *Trans. Amer. Math. Soc.* **355**:4 (2003), 1297–1364. MR Zbl

[Mityagin 1965] B. Mityagin, "An interpolation theorem for modular spaces", *Mat. Sb.* **66(108)**:4 (1965), 473–482. In Russian; translated in *Interpolation spaces and allied topics in analysis*, edited by M. Cwikel and J. Peetre, Lecture Notes in Math. **1070** (1984), Springer, 10–23. Zbl

[Morrey 1938] C. B. Morrey, Jr., "On the solutions of quasi-linear elliptic partial differential equations", *Trans. Amer. Math. Soc.* **43**:1 (1938), 126–166. MR Zbl

[Nakai 2008] E. Nakai, "Orlicz–Morrey spaces and the Hardy–Littlewood maximal function", *Studia Math.* **188**:3 (2008), 193–221. MR Zbl

[Ovchinnikov 1984] V. I. Ovchinnikov, "The method of orbits in interpolation theory", *Math. Rep.* **1**:2 (1984), 349–515. MR Zbl

[Peetre 1969] J. Peetre, "On the theory of $\mathscr{L}_{p,\lambda}$ spaces", *J. Funct. Anal.* **4** (1969), 71–87. MR Zbl

[Reisner 1993] S. Reisner, "Some remarks on Lozanovskyi's intermediate normed lattices", *Bull. Polish Acad. Sci. Math.* **41**:3 (1993), 189–196. MR Zbl

[Sedaev and Semenov 1971] A. A. Sedaev and E. M. Semenov, "The possibility of describing interpolation spaces in terms of Peetre's K-method", *Optimizacija* **4(21)** (1971), 98–114. In Russian. MR Zbl

[Shestakov 1974] V. A. Shestakov, "Complex interpolation in Banach spaces of measurable functions", *Vestnik Leningrad Univ.* **19** (1974), 64–68. In Russian. MR Zbl

[Stampacchia 1964] G. Stampacchia, "$\mathscr{L}^{(p,\lambda)}$-spaces and interpolation", *Comm. Pure Appl. Math.* **17** (1964), 293–306. MR Zbl

[Taylor 1997] M. E. Taylor, "Microlocal analysis on Morrey spaces", pp. 97–135 in *Singularities and oscillations* (Minneapolis, 1994/1995), edited by J. Rauch and M. Taylor, IMA Vol. Math. Appl. **91**, Springer, 1997. MR Zbl

[Wiener 1930] N. Wiener, "Generalized harmonic analysis", *Acta Math.* **55**:1 (1930), 117–258. MR Zbl

[Wiener 1932] N. Wiener, "Tauberian theorems", *Ann. of Math.* (2) **33**:1 (1932), 1–100. MR Zbl

[Yuan et al. 2015] W. Yuan, W. Sickel, and D. Yang, "Interpolation of Morrey–Campanato and related smoothness spaces", *Sci. China Math.* **58**:9 (2015), 1835–1908. MR Zbl

Received 12 Oct 2017. Revised 3 Sep 2018. Accepted 19 Dec 2018.

MIECZYSŁAW MASTYŁO: mastylo@amu.edu.pl
Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland

YOSHIHIRO SAWANO: ysawano@tmu.ac.jp
Department of Mathematics and Information Sciences, Tokyo Metropolitan University, Tokyo, Japan

mathematical sciences publishers

msp

ANALYSIS AND PDE
Vol. 12, No. 7, 2019
dx.doi.org/10.2140/apde.2019.12.1741

MONOTONICITY AND LOCAL UNIQUENESS FOR THE HELMHOLTZ EQUATION

BASTIAN HARRACH, VALTER POHJOLA AND MIKKO SALO

This work extends monotonicity-based methods in inverse problems to the case of the Helmholtz (or stationary Schrödinger) equation $(\Delta + k^2 q)u = 0$ in a bounded domain for fixed nonresonance frequency $k > 0$ and real-valued scattering coefficient function q. We show a monotonicity relation between the scattering coefficient q and the local Neumann-to-Dirichlet operator that holds up to finitely many eigenvalues. Combining this with the method of localized potentials, or Runge approximation, adapted to the case where finitely many constraints are present, we derive a constructive monotonicity-based characterization of scatterers from partial boundary data. We also obtain the local uniqueness result that two coefficient functions q_1 and q_2 can be distinguished by partial boundary data if there is a neighborhood of the boundary part where $q_1 \geq q_2$ and $q_1 \not\equiv q_2$.

1. Introduction

Let $\Omega \subseteq \mathbb{R}^n$, $n \geq 2$, be a bounded Lipschitz domain with unit outer normal ν. For a fixed nonresonance frequency $k > 0$, we study the relation between a real-valued scattering coefficient function $q \in L^\infty(\Omega)$ in the Helmholtz equation (or time-independent Schrödinger equation)

$$(\Delta + k^2 q)u = 0 \quad \text{in } \Omega \tag{1}$$

and the local (or partial) Neumann-to-Dirichlet (NtD) operator

$$\Lambda(q) : L^2(\Sigma) \to L^2(\Sigma), \quad g \mapsto u|_\Sigma,$$

where $u \in H^1(\Omega)$ solves (1) with Neumann data

$$\partial_\nu u|_{\partial\Omega} = \begin{cases} g & \text{on } \Sigma, \\ 0 & \text{else.} \end{cases}$$

Here $\Sigma \subseteq \partial\Omega$ is assumed to be an arbitrary nonempty relatively open subset of $\partial\Omega$. Since k is a nonresonance frequency, $\Lambda(q)$ is well-defined and is easily shown to be a self-adjoint compact operator.

We will show that

$$q_1 \leq q_2 \quad \text{implies} \quad \Lambda(q_1) \leq_{\text{fin}} \Lambda(q_2),$$

MSC2010: primary 35R30; secondary 35J05.
Keywords: inverse coefficient problems, Helmholtz equation, stationary Schrödinger equation, monotonicity, localized potentials.

where the inequality on the left-hand side is to be understood pointwise almost everywhere, and the right-hand side denotes that $\Lambda(q_2) - \Lambda(q_1)$ possesses only finitely many negative eigenvalues. Based on a slightly stronger quantitative version of this monotonicity relation, and an extension of the technique of localized potentials [Gebauer 2008] to spaces with finite codimension, we deduce the following local uniqueness result for determining q from $\Lambda(q)$.

Theorem 1.1. *Let $O \subseteq \overline{\Omega}$ be a connected relatively open set with $O \cap \Sigma \neq \varnothing$ and $q_1 \leq q_2$ on O. Then*

$$\Lambda(q_1) = \Lambda(q_2) \qquad \textit{implies} \qquad q_1 = q_2 \quad \textit{in } O.$$

Moreover, if $q_1|_O \not\equiv q_2|_O$, then $\Lambda(q_2) - \Lambda(q_1)$ has infinitely many positive eigenvalues.

Theorem 1.1 will be proven in Section 5. Note that this result removes the assumption $q_1, q_2 \in L^\infty_+(\Omega)$ from the local uniqueness result in [Harrach and Ullrich 2017], and that it implies global uniqueness if $q_1 - q_2$ is piecewise-analytic; see Corollary 5.2. Note also that in dimension $n = 2$, Imanuvilov, Uhlmann and Yamamoto [Imanuvilov et al. 2015] have proven global uniqueness with partial boundary data for potentials $q \in W^{1,p}(\Omega)$, $p > 2$. Compared to the result in [Imanuvilov et al. 2015], Theorem 1.1 is both less restrictive, as it holds for L^∞-potentials and any dimension $n \geq 2$, and more restrictive, as it relies on a local definiteness condition that is not required in [Imanuvilov et al. 2015].

Additionally to Theorem 1.1, we will also derive a constructive monotonicity-based method to detect a scatterer in an otherwise homogeneous domain. Let the scatterer $D \subseteq \Omega$ be an open set such that $\overline{D} \subseteq \Omega$ and the complement $\Omega \setminus \overline{D}$ is connected, and let

$$\begin{aligned} q(x) &= 1 \quad \text{for } x \in \Omega \setminus D \text{ (a.e.), and} \\ 1 < q_{\min} \leq q(x) &\leq q_{\max} \quad \text{for } x \in D \text{ (a.e.),} \end{aligned}$$

with constants $q_{\min}, q_{\max} > 0$. For an open set $B \subseteq \Omega$, we define the self-adjoint compact operator

$$T_B : L^2(\Sigma) \to L^2(\Sigma), \qquad \int_\Sigma g T_B h \, \mathrm{d}s := \int_B k^2 u_1^{(g)} u_1^{(h)} \, \mathrm{d}x,$$

where $u_1^{(g)}, u_1^{(h)} \in H^1(\Omega)$ solve (1) with $q \equiv 1$ and Neumann data g, h respectively.

Theorem 1.2. *For all $0 < \alpha \leq q_{\min} - 1$,*

$$B \subseteq D \quad \textit{if and only if} \quad \alpha T_B \leq_{\text{fin}} \Lambda(q) - \Lambda(1).$$

We will also give a bound on the number of negative eigenvalues in the case $B \subseteq D$, and prove a similar result for scatterers with negative contrast in Section 6.

Let us give some references on related works and comment on the origins and relevance of our result. The inverse problem considered in this work is closely related to the inverse conductivity problem of determining the positive conductivity function γ in the equation $\nabla \cdot (\gamma \nabla u) = 0$ in a bounded domain in $\mathbb{R}^n$ from knowledge of the associated Neumann-to-Dirichlet operator. This is also known as the problem of electrical impedance tomography or the Calderón problem [1980; 2006]. For a short list of seminal contributions for full boundary data let us refer to [Kohn and Vogelius 1984; 1985; Druskin 1998; Sylvester and Uhlmann 1987; Nachman 1996; Astala and Päivärinta 2006; Haberman and Tataru

2013; Caro and Rogers 2016]. For the uniqueness problem with partial boundary data there are rather precise results if $n = 2$ (see [Imanuvilov et al. 2010; 2015] and the survey [Guillarmou and Tzou 2013]), but in dimensions $n \geq 3$ it is an open question whether measurements on an arbitrary open set $\Sigma \subseteq \partial\Omega$ suffice to determine the unknown coefficient. We refer to [Kenig et al. 2007; Isakov 2007; Kenig and Salo 2013; Krupchyk and Uhlmann 2016] and the overview article [Kenig and Salo 2014] for known results, which either impose strong geometric restrictions on the inaccessible part of the boundary or require measurements of Dirichlet and Neumann data on sets that cover a neighborhood of the so-called front face

$$F(x_0) = \{x \in \partial\Omega : (x - x_0) \cdot \nu(x) \leq 0\}$$

for a point x_0 outside the closed convex hull of Ω. Also note that partial boundary data determines full boundary data by unique continuation if there exists a connected neighborhood of the full boundary on which the coefficient is known, so that uniqueness also holds in this case; see [Ammari and Uhlmann 2004].

Theorem 1.1, as well as the previous work [Harrach and Ullrich 2017], give uniqueness results where the measurements are made on an arbitrary open set $\Sigma \subseteq \partial\Omega$. Our result shows that a coefficient change in the positive or negative direction in a neighborhood of Σ (or an open subset of Σ) always leads to a change in the Neumann–Dirichlet-operator irrespectively of what happens outside this neighborhood, or the geometry or topology of the domain. Note however that our uniqueness result requires that there is a neighborhood of the boundary part on which the coefficient change is of definite sign. Our uniqueness result does not cover coefficient changes that are infinitely oscillating between positive and negative values when approaching the boundary.

Our result is based on combining monotonicity estimates (similar to those originally derived in [Kang et al. 1997; Ikehata 1998]) with localized potentials. Other theoretical uniqueness results have been obtained by this approach in [Arnold and Harrach 2013; Gebauer 2008; Harrach 2009; 2012; Harrach and Seo 2010; Harrach and Ullrich 2017]. Also note that monotonicity relations have been used in various ways in the study of inverse problems; see, e.g., [Kohn and Vogelius 1984; 1985; Isakov 1988; Alessandrini 1990; Ikehata 1999], where uniqueness results are established by methods that involve monotonicity conditions and blow-up arguments.

Monotonicity-based methods for detecting regions (or inclusions) where a coefficient function differs from a known background have been introduced by Tamburrino and Rubinacci [2002] for the inverse conductivity problem. In that paper they propose to simulate boundary measurements for a number of test regions and then use the fact that a monotonicity relation between the simulated and the true measurements will hold, if the test region lies inside the true inclusion. The work [Harrach and Ullrich 2013] used the technique of localized potentials [Gebauer 2008] to prove that this is really an if-and-only-if relation for the case of continuous measurements modeled by the NtD operator. Moreover, [Harrach and Ullrich 2013] also showed that this if-and-only-if relation still holds when the simulated measurements are replaced by linearized approximations so that the monotonicity method can be implemented without solving any forward problems other than that for the known background medium. For a list of recent works on monotonicity-based methods, let us refer to [Harrach et al. 2015; 2019; Harrach and Ullrich

2015; Harrach and Minh 2016; 2018; Maffucci et al. 2016; Tamburrino et al. 2016; Barth et al. 2017; Garde 2018; Garde and Staboulis 2017; 2019; Harrach and Lin 2017; Su et al. 2017; Ventre et al. 2017; Brander et al. 2018; Griesmaier and Harrach 2018; Zhou et al. 2018; Harrach and Meftahi 2019; Harrach 2019; Harrach and Lin 2019].

Previous monotonicity-based results often considered second-order equations with positive bilinear forms, such as the conductivity equation. So far, this positivity has been the key to proving monotonicity inequalities between the coefficient and the Neumann-to-Dirichlet operator, and previous results fail to hold in general for equations involving a positive frequency $k > 0$ (or a negative potential for the Schrödinger equation). In this article, we remove this limitation and introduce methods for more general elliptic models. We will focus on the Helmholtz equation in a bounded domain as a model case, but the ideas might be applicable to inverse boundary value and scattering problems for, e.g., Helmholtz, Maxwell, and elasticity equations. The main technical novelty of this work is that we treat compact perturbations of positive bilinear forms by extending the monotonicity relations to only hold true up to finitely many eigenvalues, and extend the localized potentials arguments to hold on spaces of finite codimension.

It should also be noted that the localized potentials arguments in [Gebauer 2008] stem from the ideas of the factorization method that was originally developed for scattering problems involving far-field measurements of the Helmholtz equation by Kirsch [1998], see also [Kirsch and Grinberg 2008], and then extended to the inverse conductivity problem in [Brühl and Hanke 2000; Brühl 2001]; see also the overview article [Harrach 2013]. For the inverse conductivity problem, the monotonicity method has the advantage over the factorization method that it allows a convergent regularized numerical implementation, see [Harrach and Ullrich 2013, Remark 3.5; Garde and Staboulis 2019], and that it can also be used for the indefinite case where anomalies of larger and smaller conductivity are present. The localized potentials approach in [Gebauer 2008] has recently been extended to show the possibility of localizing and concentrating electromagnetic fields in [Harrach et al. 2018].

The paper is structured as follows. In Section 2 we discuss the well-posedness of the Helmholtz equation outside resonance frequencies, introduce the Neumann-to-Dirichlet-operators, and give a unique continuation result from sets of positive measure. Sections 3 and 4 contain the main theoretical tools for this work. In Section 3, we introduce a Loewner order of compact self-adjoint operators that holds up to finitely many negative eigenvalues, and show that increasing the scattering index monotonically increases the Neumann-to-Dirichlet-operator in the sense of this new order. We also characterize the connection between the finite number of negative eigenvalues that have to be excluded in the Loewner ordering and the Neumann eigenvalues for the Laplacian. Section 4 extends the localized potentials result from [Gebauer 2008] to the Helmholtz equation and shows that the energy terms appearing in the monotonicity relation can be controlled in spaces of finite codimension. We give two independent proofs of this result, one using a functional analytic relation between operator norms and the ranges of their adjoints, and an alternative proof that is based on a Runge approximation argument. Sections 5 and 6 then contain the main results of this work on local uniqueness for the bounded Helmholtz equation and the detection of scatterers by monotonicity comparisons; see Theorem 1.1 and 1.2 above.

A preliminary version of these results has been published as the extended abstract [Harrach et al. 2017]. The bound on the number of negative eigenvalues in the monotonicity inequalities derived in this work has recently been improved in [Harrach et al. 2019].

2. The Helmholtz equation in a bounded domain

We start by summarizing some properties of the Neumann-to-Dirichlet-operators, discuss well-posedness and the role of resonance frequencies, and state a unique continuation result for the Helmholtz equation in a bounded domain.

2A. ***Neumann-to-Dirichlet operators.*** Throughout this work, let $\Omega \subseteq \mathbb{R}^n$, $n \geq 2$, denote a bounded domain with Lipschitz boundary and outer unit normal ν, and let $\Sigma \subseteq \partial\Omega$ be an open subset of $\partial\Omega$. For a frequency $k \geq 0$ and a real-valued scattering coefficient function $q \in L^\infty(\Omega)$, we consider the Helmholtz equation with (partial) Neumann boundary data $g \in L^2(\Sigma)$, i.e., finding $u \in H^1(\Omega)$ with

$$(\Delta + k^2 q)u = 0 \quad \text{in } \Omega, \qquad \partial_\nu u|_{\partial\Omega} = \begin{cases} g & \text{on } \Sigma, \\ 0 & \text{else.} \end{cases} \tag{2}$$

We also denote the solution by $u_q^{(g)}$ instead of u if the choice of g and q is not clear from the context.

The Neumann problem (2) is equivalent to the variational formulation of finding $u \in H^1(\Omega)$ such that

$$\int_\Omega (\nabla u \cdot \nabla v - k^2 q u v)\,\mathrm{d}x = \int_{\partial\Omega} g v|_{\partial\Omega}\,\mathrm{d}s \quad \text{for all } v \in H^1(\Omega). \tag{3}$$

We introduce the bounded linear operators

$$I : H^1(\Omega) \to H^1(\Omega),$$
$$j : H^1(\Omega) \to L^2(\Omega),$$
$$M_q : L^2(\Omega) \to L^2(\Omega),$$

where I denotes the identity operator, j is the compact embedding from H^1 to L^2, and M_q is the multiplication operator by q. We furthermore use $\langle \cdot, \cdot \rangle$ to denote the $H^1(\Omega)$ inner product and define the operators

$$K := j^* j \quad \text{and} \quad K_q := j^* M_q j,$$

which are compact self-adjoint linear operators from $H^1(\Omega)$ to $H^1(\Omega)$. By

$$\gamma_\Sigma : H^1(\Omega) \to L^2(\Sigma), \quad v \mapsto v|_\Sigma,$$

we denote the compact trace operator.

With this notation (3) can be written as

$$\langle (I - K - k^2 K_q)u, v \rangle = \int_{\partial\Omega} g(\gamma_\Sigma v)\,\mathrm{d}s \quad \text{for all } v \in H^1(\Omega),$$

so that the Neumann problem for the Helmholtz equation (2) is equivalent to the equation

$$(I - K - k^2 K_q)u = \gamma_\Sigma^* g. \tag{4}$$

Our results on identifying the scattering coefficient q will require that $I - K - k^2 K_q$ is continuously invertible, which is equivalent to the fact that k *is not a resonance frequency*, or, equivalently; that 0 *is not a Neumann eigenvalue*, see Lemmas 2.2 and 3.10. Note that this implies, in particular, that $k > 0$ and $q \not\equiv 0$. We can then define the Neumann-to-Dirichlet operator (with Neumann data prescribed and Dirichlet data measured on the same open subset $\Sigma \subseteq \partial\Omega$)

$$\Lambda(q) : L^2(\Sigma) \to L^2(\Sigma), \quad g \mapsto u|_\Sigma, \quad \text{where } u \in H^1(\Omega) \text{ solves (2).}$$

The Neumann-to-Dirichlet operator satisfies

$$\Lambda(q) = \gamma_\Sigma (I - K - k^2 K_q)^{-1} \gamma_\Sigma^*, \tag{5}$$

which shows that $\Lambda(q)$ is a compact self-adjoint linear operator.

We will show in Section 3 that there is a monotonicity relation between the scattering coefficient q and the Neumann-to-Dirichlet-operator $\Lambda(q)$. Increasing q will increase $\Lambda(q)$ in the sense of operator definiteness up to finitely many eigenvalues. The number of eigenvalues that do not follow the increase will be bounded by the number defined in the following lemma. Note that here, and throughout the paper, we always count the number of eigenvalues of a compact self-adjoint operator with multiplicity according to the dimension of the associated eigenspaces.

Lemma 2.1. *Given $k > 0$ and $q \in L^\infty(\Omega)$, let $d(q)$ be the number of eigenvalues of $K + k^2 K_q$ that are larger than* 1, *and let $V(q)$ be the sum of the associated eigenspaces. Then $d(q) = \dim(V(q)) \in \mathbb{N}_0$ is finite*, *and*

$$\int_\Omega (|\nabla v|^2 - k^2 q |v|^2)\,\mathrm{d}x \geq 0 \quad \text{for all } v \in V(q)^\perp,$$

where $V(q)^\perp$ denotes the orthocomplement of $V(q)$ in $H^1(\Omega)$.

Proof. Since

$$\langle (I - K - k^2 K_q) v, v \rangle = \int_\Omega (|\nabla v|^2 - k^2 q |v|^2)\,\mathrm{d}x,$$

the assertion follows from the spectral theorem for compact self-adjoint operators. □

We will show in Lemma 3.10 that $d(q)$ agrees with the number of positive Neumann eigenvalues of $\Delta + k^2 q$. If $q(x) \leq q_{\max} \in \mathbb{R}$ for all $x \in \Omega$ (a.e.) then $d(q) \leq d(q_{\max})$, and $d(q_{\max})$ is the number of Neumann eigenvalues of the Laplacian Δ that are larger than $-k^2 q_{\max}$; see Corollary 3.11.

2B. ***Resonance frequencies.*** We now summarize some results on the solvability of the Helmholtz equation (2) outside of resonance frequencies.

Lemma 2.2. *Let $q \in L^\infty(\Omega)$.*

(a) *For each $k \geq 0$, the following properties are equivalent*:

(i) *For each $F \in L^2(\Omega)$ and $g \in L^2(\partial\Omega)$, there exists a unique solution $u \in H^1(\Omega)$ of*

$$(\Delta + k^2 q)u = F \quad \text{in } \Omega, \qquad \partial_\nu u|_{\partial\Omega} = g, \tag{6}$$

and the solution depends linearly and continuously on F and g.

(ii) *The homogeneous Neumann problem*

$$(\Delta + k^2 q)u = 0 \quad \textit{in } \Omega, \qquad \partial_\nu u|_{\partial\Omega} = 0, \tag{7}$$

admits only the trivial solution $u \equiv 0$.

(iii) *The operator* $I - K - k^2 K_q : H^1(\Omega) \to H^1(\Omega)$ *is continuously invertible.*

We call k *a* ***resonance frequency*** *if the properties* (i)–(iii) *do not hold.*

(b) *If* $q \not\equiv 0$, *then the set of resonance frequencies is countable and discrete.*

Proof. (a) Clearly, (i) implies (ii), and, using the equivalence of (2) and (4), (ii) implies that $I - K - k^2 K_q$ is injective. Since K and K_q are compact, the operator $I - K - k^2 K_q$ is Fredholm of index 0. Hence, injectivity of $I - K - k^2 K_q$ already implies that $I - K - k^2 K_q$ is continuously invertible, so that (ii) implies (iii). Finally, $u \in H^1(\Omega)$ solves (6) if and only if

$$\int_\Omega (\nabla u \cdot \nabla v - k^2 q u v)\, \mathrm{d}x = -\int_\Omega F v \, \mathrm{d}x + \int_{\partial\Omega} g v|_{\partial\Omega} \, \mathrm{d}S \quad \text{for all } v \in H^1(\Omega).$$

This is equivalent to

$$\langle (I - K - k^2 K_q)u, v\rangle = -\int_\Omega F j(v)\, \mathrm{d}x + \int_{\partial\Omega} g \gamma_{\partial\Omega}(v)\, \mathrm{d}s \quad \text{for all } v \in H^1(\Omega),$$

and thus equivalent to

$$(I - K - k^2 K_q)u = -j^* F + \gamma_{\partial\Omega}^* g,$$

so that (iii) implies (i).

(b) We extend I, K, and K_q to the Sobolev space of complex-valued functions

$$I, K, K_q : H^1(\Omega; \mathbb{C}) \to H^1(\Omega; \mathbb{C}).$$

For $k \in \mathbb{C}$ we then define

$$R(k) := K + k^2 K_q : H^1(\Omega; \mathbb{C}) \to H^1(\Omega; \mathbb{C}).$$

$R(k)$ is a family of compact operators depending analytically on $k \in \mathbb{C}$. The analytic Fredholm theorem, see, e.g., [Reed and Simon 1972, Theorem VI.14], now implies that either $I - R(k)$ is not invertible for all $k \in \mathbb{C}$, or that there is a countable discrete set $Z \subseteq \mathbb{C}$ such that $I - R(k)$ is continuously invertible when $k \in \mathbb{C} \setminus Z$. Hence, to prove (b), it suffices to show that there exists $k \in \mathbb{C}$ for which $I - R(k)$ is invertible.

We will show that this is the case for any $0 \neq k \in \mathbb{C}$ with $\mathrm{Re}(k^2) = 0$. In fact, $(I - R(k))u = 0$ implies

$$0 = \int_\Omega (\nabla u \cdot \nabla v - k^2 q u v)\, \mathrm{d}x \quad \text{for all } v \in H^1(\Omega; \mathbb{C}).$$

Using $v := \bar{u}$ and taking the real part yields that $0 = \int_\Omega |\nabla u|^2 \, \mathrm{d}x$, which shows that u must be constant, and that

$$\int_\Omega k^2 q u v \, \mathrm{d}x = 0 \quad \text{for all } v \in H^1(\Omega; \mathbb{C}).$$

Together with $k^2 \neq 0$, and $q \not\equiv 0$, this shows that $u \equiv 0$. Hence, $I - R(k)$ is injective and thus invertible for all $0 \neq k \in \mathbb{C}$ with $\mathrm{Re}(k^2) = 0$. □

2C. ***Unique continuation.*** We will make use of a unique continuation property for the Helmholtz equation from sets of positive measure. In two dimensions, this follows from a standard reduction to quasiconformal mappings. However, since we could not find a proof in the literature, we will first give the argument following [Alessandrini 2012] and references therein (in fact [Alessandrini 2012] proves strong unique continuation for more general equations). See also [Astala et al. 2009] for basic facts on quasiconformal mappings in the plane.

Lemma 2.3. *Let $\Omega \subset \mathbb{R}^2$ be a connected open set, and suppose that $u \in H^1_{\mathrm{loc}}(\Omega)$ is a weak solution of*

$$-\operatorname{div}(A\nabla u) + du = 0 \quad \text{in } \Omega,$$

where $A \in L^\infty(\Omega, \mathbb{R}^{n\times n})$ is symmetric and satisfies $A(x)\xi \cdot \xi \geq c_0|\xi|^2$ for some $c_0 > 0$, and $d \in L^{q/2}(\Omega)$ for some $q > 2$. If u vanishes in a set E of positive measure, then $u \equiv 0$ in Ω.

Proof. It is enough to show that u vanishes in some ball, since then weak (or strong) unique continuation [Alessandrini 2012] implies that $u \equiv 0$. Let x_0 be a point of density 1 in E and let $U_r := B_r(x_0)$ and $E_r := E \cap U_r$. There is $r_0 > 0$ so that if $r < r_0$, then $U_r \subset \Omega$ and E_r has positive measure.

We will now work in U_r. Observe first that there is $p > 2$ so that $u \in W^{1,p}(U_r)$ [Astala et al. 2009, Theorem 16.1.4]. In particular u is Hölder continuous and we may assume (after removing a set of measure zero from E) that $u(x) = 0$ for all $x \in E_r$. The first step is to show that $\nabla u = 0$ a.e. on E_r. Let N_1 be the set of points in E_r where u is not differentiable, and let N_2 be the set of points of density < 1 in E_r. Then N_1 and N_2 have zero measure. Fix a point $x \in E_r \setminus (N_1 \cup N_2)$ and a unit direction e. There is a sequence (x_j) with $x_j \in B(x, 1/j) \cap E_r$ so that $\left|(x_j - x)/|x_j - x| - e\right| \leq 1/j$ for j large (for if not, then all points in E_r near x would be outside a fixed sector in direction e, which contradicts the fact that x has density 1). Since u is differentiable at x,

$$u(x_j) - u(x) = \nabla u(x) \cdot (x_j - x) + o(|x - x_j|).$$

Dividing by $|x - x_j|$ and using that $u(x_j) = u(x) = 0$ implies that $\nabla u(x) \cdot e = 0$. It follows that ∇u vanishes in $E_r \setminus (N_1 \cup N_2)$, so indeed

$$u = 0 \quad \text{in } E_r, \qquad \nabla u = 0 \quad \text{a.e. in } E_r. \tag{8}$$

The next step is to reduce to the case where $d = 0$. As in [Alessandrini 2012, Proposition 2.4], we choose r small enough so that there is a nonvanishing $w \in W^{1,p}(U_r)$ satisfying

$$-\operatorname{div}(A\nabla w) + dw = 0 \quad \text{in } U_r,$$
$$\tfrac{1}{2} \leq w \leq 2 \quad \text{in } U_r, \qquad \|\nabla w\|_{L^p(U_r)} \leq 1.$$

We write $v = u/w$. It follows that $v \in W^{1,p}(U_r)$ is a weak solution of

$$-\operatorname{div}(\tilde{A}\nabla v) = 0 \quad \text{in } U_r,$$

where $\tilde{A} = w^2 A$ is L^∞ and uniformly elliptic. Moreover, (8) implies that

$$v = 0 \quad \text{in } E_r, \qquad \nabla v = 0 \quad \text{a.e. in } E_r. \tag{9}$$

To prove the lemma, we will show that $v \equiv 0$ in some ball.

Let $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Since $\tilde{A}\nabla v$ is divergence-free, there is a real-valued function $\tilde{v} \in H^1(U_r)$ satisfying

$$\nabla \tilde{v} = J(\tilde{A}\nabla v). \tag{10}$$

Such a function $\tilde{v}$ is unique up to an additive constant. Define

$$f = v + i\tilde{v}.$$

As in [Alessandrini 2012], $f \in H^1(U_r)$ solves an equation of the form

$$\partial_{\bar{z}} f = \mu \partial_z f + \nu \bar{\partial}_z f \quad \text{in } U_r,$$

where $\|\mu\|_{L^\infty(U_r)} + \|\nu\|_{L^\infty(U_r)} < 1$. It follows that f is a quasiregular map and by the Stoilow factorization [Astala et al. 2009, Theorem 5.5.1] it has the representation

$$f(z) = F(\chi(z)), \quad z \in U,$$

where χ is a quasiconformal map $\mathbb{C} \to \mathbb{C}$ and F is a holomorphic function on $\chi(U)$.

Finally, the Jacobian determinant J_f of f is given by

$$J_f(z) = F'(\chi(z)) J_\chi(z).$$

Using (9) and (10), we see that $J_f = 0$ a.e. in E_r. Moreover, since χ is quasiconformal, J_χ can only vanish in a set of measure zero [Astala et al. 2009, Corollary 3.7.6]. It follows that $F'(\chi(z)) = 0$ for a.e. $z \in E_r$. Then the Taylor series of the analytic function F' at $\chi(x_0)$ must vanish (otherwise one would have $F'(\chi(z)) = (\chi(z) - \chi(x_0))^N g(\chi(z))$, where $g(\chi(x_0)) \neq 0$ and the only zero near x_0 would be $z = x_0$). Thus $F' = 0$ near x_0, so F is constant, f is also constant, and $v = 0$ near x_0. □

We can now state the unique continuation property for any dimension $n \geq 2$ in the form that we will utilize in the later sections. As in [Harrach and Ullrich 2013, Definition 2.2] we say that a relatively open subset $O \subseteq \overline{\Omega}$ is *connected to* Σ if O is connected and $\Sigma \cap O \neq \varnothing$.

Theorem 2.4. (a) *Let $u \in H^1(\Omega)$ solve*

$$(\Delta + k^2 q)u = 0 \quad \text{in } \Omega. \tag{11}$$

If $u|_E = 0$ for a subset $E \subseteq \Omega$ with positive measure then $u(x) = 0$ for all $x \in \Omega$ (a.e.)

(b) *Let $u \in H^1(\Omega)$, $\Delta u \in L^2(\Omega)$, and*

$$(\Delta + k^2 q)u = 0 \quad \text{in } \Omega \setminus C$$

for a closed set C for which $\overline{\Omega} \setminus C$ is connected to Σ. If $u|_\Sigma = 0$ and $\partial_\nu u|_\Sigma = 0$, then $u(x) = 0$ for all $x \in \Omega \setminus C$ (a.e.)

Proof. For $n = 2$, (a) follows from Lemma 2.3. For $n \geq 3$, (a) is shown in [Harrach and Ullrich 2017, Theorem 4.2] (see also [Regbaoui 2001, proof of Theorem 2.1]) by combining the following two results:

(i) If $u \in H^1(\Omega)$ solves (11) and vanishes on a measurable set of positive measure then u has a zero of infinite order; see, e.g., [de Figueiredo and Gossez 1992, Proposition 3; Hadi and Tsouli 2001, Theorem 2.1].

(ii) The trivial solution $u \equiv 0$ is the only $H^1(\Omega)$-solution of (11) that has a zero of infinite order; see, e.g, [Hörmander 1985, Theorem 17.2.6].

Part (b) follows from (a) by extending u by zero on $B \setminus \Omega$, where B is a small ball with $B \cap \partial\Omega \subseteq \Sigma$; see the proof of Lemma 4.4(c) in [Harrach and Ullrich 2017]. □

3. Monotonicity and localized potentials for the Helmholtz equation

In this section we show that increasing the scattering coefficient leads to a larger Neumann-to-Dirichlet operator in a certain sense. For this result, the Neumann-to-Dirichlet operators are ordered by an extension of the Loewner order of compact self-adjoint operators that holds up to finitely many negative eigenvalues.

3A. ***A Loewner order up to finitely many eigenvalues.*** We start by giving a rigorous definition and characterization of this ordering.

Definition 3.1. Let $A, B : X \to X$ be two self-adjoint compact linear operators on a Hilbert space X. For a number $d \in \mathbb{N}_0$, we write

$$A \leq_d B \quad \text{or} \quad \langle Ax, x\rangle \leq_d \langle Bx, x\rangle$$

if $B - A$ has at most d negative eigenvalues. We also write $A \leq_{\text{fin}} B$ if $A \leq_d B$ holds for some $d \in \mathbb{N}_0$, and we write $A \leq B$ if $A \leq_d B$ holds for $d = 0$.

Note that for $d = 0$ this is the standard partial ordering of compact self-adjoint operators in the sense of operator definiteness (also called Loewner order). Also note that "$\leq_{\text{fin}}$" and "$\leq_d$" (for $d \neq 0$) are not partial orders since they are clearly not antisymmetric. Obviously, "$\leq_{\text{fin}}$" and "$\leq_d$" are reflexive, and "$\leq_{\text{fin}}$" is also transitive (see Lemma 3.4 below) and thus a so-called preorder.

To characterize this new ordering, we will make use of the following lemma.

Lemma 3.2. *Let $A : X \to X$ be a self-adjoint compact linear operator on a Hilbert space X with inner product $\langle \cdot, \cdot \rangle$ inducing the norm $\|\cdot\|$. Let $d \in \mathbb{N}_0$ and $r \in \mathbb{R}$, $r \geq 0$.*

(a) *The following statements are equivalent*:

(i) *A has at most d eigenvalues larger than r.*

(ii) *There exists a compact self-adjoint operator $F : X \to X$ with*

$$\dim(\mathcal{R}(F)) \leq d \qquad \textit{and} \qquad \langle (A - F)x, x\rangle \leq r\|x\|^2 \quad \textit{for all } x \in X,$$

where $\mathcal{R}(F)$ stands for the range of F.

(iii) *There exists a subspace* $W \subset X$ *with* $\operatorname{codim}(W) \leq d$ *such that*

$$\langle Aw, w\rangle \leq r\|w\|^2 \quad \textit{for all } w \in W.$$

(iv) *There exists a subspace* $V \subset X$ *with* $\dim(V) \leq d$ *such that*

$$\langle Av, v\rangle \leq r\|v\|^2 \quad \textit{for all } v \in V^{\perp}.$$

(b) *The following statements are equivalent*:

(i) A *has* (*at least*) d *eigenvalues larger than* r.

(ii) *There exists a subspace* $V \subset X$ *with* $\dim(V) \geq d$ *such that*

$$\langle Av, v\rangle > r\|v\|^2 \quad \textit{for all } v \in V.$$

Proof. (a) We start by showing that (i) implies (ii). Let A have at most d eigenvalues larger than $r \geq 0$. Let $(\lambda_k)_{k\in\mathbb{N}}$ be the nonzero eigenvalues of A, ordered in such a way that $\lambda_k \leq r$ for $k > d$. Let $\mathcal{N}(A)$ denote the kernel of A and let $(v_k)_{k\in\mathbb{N}} \in X$ be a sequence of corresponding eigenvectors forming an orthonormal basis of $\mathcal{N}(A)^{\perp}$. Then

$$Ax = \sum_{k=1}^{\infty} \lambda_k v_k \langle v_k, x\rangle \quad \text{for all } x \in X,$$

and (ii) follows with $F : X \to X$ defined by

$$F : x \mapsto \sum_{k=1}^{d} \lambda_k v_k \langle v_k, x\rangle \quad \text{for all } x \in X.$$

The implication from (ii) to (iii) follows by setting $W := \mathcal{N}(F)$ since

$$\operatorname{codim}(W) = \dim(W^{\perp}) = \dim(\mathcal{R}(F)) \leq d$$

and

$$\langle Aw, w\rangle = \langle (A - F)w, w\rangle \geq 0.$$

Part (iii) implies (iv) by setting $V := W^{\perp}$.

To show that (iv) implies (i), we assume that (i) is not true, so that A has at least $d+1$ eigenvalues larger than $r \geq 0$. We sort the positive eigenvalues of A in decreasing order to obtain

$$\lambda_1 \geq \cdots \geq \lambda_d \geq \lambda_{d+1} > r.$$

Then, by the Courant–Fischer–Weyl min-max principle, see, e.g., [Lax 2002, p. 318], we have that the minimum over all d-dimensional subspaces $V \subset X$ must satisfy

$$\min_{\substack{V\subset X \\ \dim(V)=d}} \max_{\substack{v\in V^{\perp} \\ \|v\|=1}} \langle Av, v\rangle = \lambda_{d+1} > r,$$

which shows that (iv) cannot be true. Hence, (iv) implies (i).

(b) This can be shown analogously to (a). Part (ii) follows from (i) by choosing V as the sum of eigenspaces for eigenvalues larger than r, and (ii) implies (i) by using the Courant–Fischer–Weyl min-max principle. □

Corollary 3.3. *Let $A, B : X \to X$ be two self-adjoint compact linear operators on a Hilbert space X with inner product $\langle \cdot, \cdot \rangle$. For any number $d \in \mathbb{N}_0$, the following statements are equivalent*:

(a) $A \leq_d B$.

(b) *There exists a compact self-adjoint operator $F : X \to X$ with*

$$\dim(\mathcal{R}(F)) \leq d \qquad \textit{and} \qquad \langle (B - A + F)x, x \rangle \geq 0 \quad \textit{for all } x \in X.$$

(c) *There exists a subspace $W \subset X$ with* $\operatorname{codim}(W) \leq d$ *such that*

$$\langle (B - A)w, w \rangle \geq 0 \quad \textit{for all } w \in W.$$

(d) *There exists a subspace $V \subset X$ with* $\dim(V) \leq d$ *such that*

$$\langle (B - A)v, v \rangle \geq 0 \quad \textit{for all } v \in V^{\perp}.$$

Proof. This follows from Lemma 3.2(a) with $r = 0$ and A replaced by $A - B$. □

Lemma 3.4. *Let $A, B, C : X \to X$ be self-adjoint compact linear operators on a Hilbert space X. For $d_1, d_2 \in \mathbb{N}_0$*

$$A \leq_{d_1} B \quad \textit{and} \quad B \leq_{d_2} C \qquad \textit{implies} \qquad A \leq_{d_1+d_2} C,$$
$$A \leq_{\text{fin}} B \quad \textit{and} \quad B \leq_{\text{fin}} C \qquad \textit{implies} \qquad A \leq_{\text{fin}} C.$$

Proof. This follows from the characterization in Corollary 3.3(b). □

3B. ***A monotonicity relation for the Helmholtz equation.*** With this new ordering, we can show a monotonicity relation between the scattering index and the Neumann-to-Dirichlet-operators. Note that the dimension bound in the last line of the following theorem has recently been improved to $d(q_2) - d(q_1)$ in [Harrach et al. 2019].

Theorem 3.5. *Let $q_1, q_2 \in L^\infty(\Omega) \setminus \{0\}$. Assume that $k > 0$ is not a resonance for q_1 or q_2, and let $d(q_2) \in \mathbb{N}_0$ be defined as in Lemma 2.1.*

Then there exists a subspace $V \subset L^2(\Sigma)$ with $\dim(V) \leq d(q_2)$ *such that*

$$\int_\Sigma g(\Lambda(q_2) - \Lambda(q_1))g \, \mathrm{d}s \geq \int_\Omega k^2 (q_2 - q_1)|u_1^{(g)}|^2 \, \mathrm{d}x \quad \textit{for all } g \in V^{\perp}.$$

In particular

$$q_1 \leq q_2 \quad \textit{implies} \quad \Lambda(q_1) \leq_{d(q_2)} \Lambda(q_2).$$

Remark 3.6. Note that by interchanging q_1 and q_2, Theorem 3.5 also yields that there exists a subspace $V \subset L^2(\Sigma)$ with $\dim(V) \leq d(q_1)$ such that

$$\int_\Sigma g(\Lambda(q_2) - \Lambda(q_1))g \, \mathrm{d}s \leq \int_\Omega k^2 (q_2 - q_1)|u_2^{(g)}|^2 \, \mathrm{d}x \quad \text{for all } g \in V^{\perp}.$$

To prove Theorem 3.5 we will use the following lemmas.

Lemma 3.7. *Let $q_1, q_2 \in L^\infty(\Omega) \setminus \{0\}$. Assume that $k > 0$ is not a resonance for q_1 or q_2. Then, for all $g \in L^2(\Sigma)$,*

$$\int_\Sigma g(\Lambda(q_2) - \Lambda(q_1))g \,\mathrm{d}s + \int_\Omega k^2(q_1 - q_2)|u_1^{(g)}|^2 \,\mathrm{d}x = \int_\Omega (|\nabla(u_2^{(g)} - u_1^{(g)})|^2 - k^2 q_2 |u_2^{(g)} - u_1^{(g)}|^2) \,\mathrm{d}x,$$

where $u_1^{(g)}$, $u_2^{(g)}$ solve the Helmholtz equation (2) *with Neumann boundary data g and $q = q_1$, $q = q_2$ respectively.*

Proof. Define the bilinear form

$$B_q(u, v) = \int_\Omega (\nabla u \cdot \nabla v - k^2 q u v) \,\mathrm{d}x, \quad u, v \in H^1(\Omega).$$

Writing $u_1 = u_1^{(g)}$ and $u_2 = u_2^{(g)}$, from the definition of the NtD map and from (3) we have

$$\int_\Sigma g\Lambda(q_1)g \,\mathrm{d}s = \int_\Sigma (\partial_\nu u_1)u_1 \,\mathrm{d}s = 2\int_\Sigma (\partial_\nu u_2)u_1 \,\mathrm{d}s - \int_\Sigma (\partial_\nu u_1)u_1 \,\mathrm{d}s = 2B_{q_2}(u_2, u_1) - B_{q_1}(u_1, u_1)$$

and

$$\int_\Sigma g\Lambda(q_2)g \,\mathrm{d}s = \int_\Sigma (\partial_\nu u_2)u_2 \,\mathrm{d}s = B_{q_2}(u_2, u_2).$$

We thus obtain that

$$\begin{aligned}\int_\Sigma g(\Lambda(q_2) - \Lambda(q_1))g \,\mathrm{d}s &= B_{q_2}(u_2, u_2) - 2B_{q_2}(u_2, u_1) + B_{q_1}(u_1, u_1) \\ &= B_{q_2}(u_2 - u_1, u_2 - u_1) - B_{q_2}(u_1, u_1) + B_{q_1}(u_1, u_1).\end{aligned}$$ □

We will show that the bilinear forms in the right-hand sides in Lemma 3.7 are positive up to a finite-dimensional subspace.

Lemma 3.8. *Let $q_1, q_2 \in L^\infty(\Omega) \setminus \{0\}$ for which $k > 0$ is not a resonance. There exists a subspace $V \subset L^2(\Sigma)$ with $\dim(V) \le d(q_2)$ such that for all $g \in V^\perp$*

$$\int_\Omega \left(|\nabla(u_2^{(g)} - u_1^{(g)})|^2 - k^2 q_2 |u_2^{(g)} - u_1^{(g)}|^2\right) \mathrm{d}x \ge 0.$$

Proof. Using Lemma 2.1, we have

$$\int_\Omega \left(|\nabla(u_2^{(g)} - u_1^{(g)})|^2 - k^2 q_2 |u_2^{(g)} - u_1^{(g)}|^2\right) \mathrm{d}x \ge 0$$

for all $g \in L^2(\Sigma)$ with $u_2^{(g)} - u_1^{(g)} \in V(q_2)^\perp$. The solution operators

$$S_j : L^2(\Sigma) \to H^1(\Omega), \quad g \mapsto u_j^{(g)}, \quad \text{where } u_j^{(g)} \in H^1(\Omega) \text{ solves (2)}, \quad j \in \{1, 2\},$$

are linear and bounded, and

$$(S_2 - S_1)g = u_2^{(g)} - u_1^{(g)} \in V(q_2)^\perp \quad \text{if and only if} \quad g \in ((S_2 - S_1)^* V(q_2))^\perp.$$

Since $\dim(S_2 - S_1)^* V(q_2) \le \dim V(q_2) = d(q_2)$, the assertion follows with $V := (S_2 - S_1)^* V(q_2)$. □

Proof of Theorem 3.5. This now immediately follows from combining Lemmas 3.7 and 3.8. □

3C. *The number of negative eigenvalues.* We will now further investigate the number $d(q) \in \mathbb{N}_0$ (defined in Lemma 2.1) that bounds the number of negative eigenvalues in the monotonicity relations derived in Section 3B. We will show that $d(q)$ depends monotonously on the scattering index q and show that $d(q)$ is less than or equal to the number of Neumann eigenvalues for the Laplacian which are larger than $-k^2 q_{\max}$, where $q_{\max} \geq q(x)$ for all $x \in \Omega$ (a.e.).

Lemma 3.9. *Let $q_1, q_2 \in L^\infty(\Omega)$. Then $q_1 \leq q_2$ implies $d(q_1) \leq d(q_2)$.*

Proof. The inequality $q_1 \leq q_2$ implies that $K_{q_1} \leq K_{q_2}$. Hence, the assertion follows from the equivalence of (a) and (c) in Corollary 3.3. □

Lemma 3.10. *Let $q \in L^\infty(\Omega)$, and $k \in \mathbb{R}$.*

(*a*) *There is a countable and discrete set of real values*

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots \to -\infty$$

(*called* ***Neumann eigenvalues***) *such that*

$$(\Delta + k^2 q)u = \lambda u \quad \text{in } \Omega, \qquad \partial_\nu u|_{\partial\Omega} = 0, \tag{12}$$

admits a nontrivial solution (*called a* ***Neumann eigenfunction***) $0 \not\equiv u \in H^1(\Omega)$ *if and only if* $\lambda \in \{\lambda_1, \lambda_2, \ldots\}$, *and there is an orthonormal basis* $(u_1, u_2, \ldots)$ *of* $L^2(\Omega)$ *such that* $u_j \in H^1(\Omega)$ *is a Neumann eigenfunction for* λ_j.

(b) *If λ is not a Neumann eigenvalue, then the problem*

$$(\Delta + k^2 q)u = \lambda u + F \quad \text{in } \Omega, \qquad \partial_\nu u|_{\partial\Omega} = g, \tag{13}$$

has a unique solution $u \in H^1(\Omega)$ for any $F \in L^2(\Omega)$ and $g \in L^2(\partial\Omega)$, and the solution operator is linear and bounded.

(c) *Let* $N_+ := \operatorname{span}\{u_j : \lambda_j > 0\}$. *Then* $\dim(N_+) < \infty$,

$$N_- := \overline{\operatorname{span}\{u_j : \lambda_j \leq 0\}} = \{v \in H^1(\Omega) : v \perp_{L^2} N_+\} \tag{14}$$

is a complement of N_+ (in $H^1(\Omega)$), and

$$\int_\Omega |\nabla v|^2 - k^2 q v^2 \,\mathrm{d}x < 0 \quad \text{for all } v \in N_+, \tag{15}$$

$$\int_\Omega |\nabla v|^2 - k^2 q v^2 \,\mathrm{d}x \geq 0 \quad \text{for all } v \in N_-, \tag{16}$$

where the closure in (14) *is taken with respect to the $H^1(\Omega)$-norm, and $\perp_{L^2}$ denotes orthogonality with respect to the L^2 inner product.*

(d) *$d(q)$ is the number of positive Neumann eigenvalues of $\Delta + k^2 q$; i.e., $d(q) = \dim(N_+)$.*

(e) *0 is a Neumann eigenvalue if and only if $k > 0$ is a resonance frequency.*

Proof. (a) Define $c := k^2\|q\|_{L^\infty(\Omega)} + 1 > 0$ and $R := I - K - k^2 K_q + cK$. Then R is coercive and thus continuously invertible. Using the equivalent variational formulation of (12), we have that $\lambda \in \mathbb{R}$ is a Neumann eigenvalue with Neumann eigenfunction $u \not\equiv 0$ if and only if

$$\int_\Omega (-\nabla u \cdot \nabla v + k^2 quv)\,\mathrm{d}x = \lambda \int_\Omega uv\,\mathrm{d}x \quad \text{for all } v \in H^1(\Omega),$$

which is equivalent to

$$(I - K - k^2 K_q)u = -\lambda K u$$

and thus to

$$Ru = (I - K - k^2 K_q + cK)u = (c - \lambda)Ku. \tag{17}$$

This shows that c cannot be a Neumann eigenvalue since $Ru \not\equiv 0$ for $u \not\equiv 0$. Moreover, using $K = j^* j$, the invertibility of R, and the injectivity of j, we have that (17) is equivalent to

$$\frac{1}{c - \lambda}(ju) = jR^{-1}j^*(ju).$$

This shows that $\lambda \in \mathbb{R}$ is a Neumann eigenvalue with Neumann eigenfunction $u \in H^1(\Omega)$ if and only if $ju \in L^2(\Omega)$ is an eigenfunction of $jR^{-1}j^* : L^2(\Omega) \to L^2(\Omega)$ with eigenvalue $1/(c - \lambda)$. Since j is injective, and every eigenfunction of $jR^{-1}j^*$ lies in the range of j, this is a one-to-one correspondence, and the dimension of the corresponding eigenspaces is the same. Since $jR^{-1}j^*$ is a compact, self-adjoint, positive operator, the assertions in (a) follow from the spectral theorem on self-adjoint compact operators.

(b) This follows from the fact that $I - K - k^2 K_q - \lambda K$ is Fredholm of index 0 and thus continuously invertible if it is injective.

(c) $\dim(N_+) < \infty$ follows from (a). We define

$$N_- := \overline{\operatorname{span}\{u_j : \lambda_j \le 0\}} \quad \text{and} \quad \widetilde{N}_- := \{v \in H^1(\Omega) : v \perp_{L^2} N_+\}.$$

$\widetilde{N}_-$ is closed with respect to the H^1-norm and contains all u_j with $\lambda_j \le 0$, so that $N_- \subseteq \widetilde{N}_-$. To show $N_- = \widetilde{N}_-$, we argue by contradiction. If $N_- \subsetneq \widetilde{N}_-$, then there would exist a $0 \neq v \in \widetilde{N}_-$ with $\langle u_j, v\rangle = 0$ for all u_j with $\lambda_j \le 0$. Using

$$\begin{aligned} 0 = \langle u_j, v\rangle &= \int_\Omega (\nabla u_j \cdot \nabla v + u_j v)\,\mathrm{d}x \\ &= \int_\Omega (\nabla u_j \cdot \nabla v - k^2 q u_j v)\,\mathrm{d}x + \int_\Omega (1 + k^2 q) u_j v\,\mathrm{d}x \\ &= \int_\Omega (1 + k^2 q - \lambda_j) u_j v\,\mathrm{d}x, \end{aligned}$$

and the fact that $\lambda_j \to -\infty$, it would follow that $v \perp_{L^2} u_j$ for all but finitely many u_j. Since $v \perp_{L^2} N_+$, and $(u_1, u_2, \ldots)$ is an orthonormal basis of $L^2(\Omega)$, v must then be a finite combination of u_j with $\lambda_j \le 0$, which would imply that $v = 0$. Hence, $N_- = \widetilde{N}_-$, so that the equality in (14) is proven.

Obviously, $N_+ \cap N_- = 0$ and every $v \in H^1(\Omega)$ can be written as

$$v = \sum_{\lambda_j > 0} \left(\int_\Omega v u_j \, \mathrm{d}x \right) u_j + \left(v - \sum_{\lambda_j > 0} \left(\int_\Omega v u_j \, \mathrm{d}x \right) u_j \right) \in N_+ + N_-,$$

which shows that N_- is a complement of N_+.

To show (15), we use the L^2-orthogonality of the u_j to obtain for all $v = \sum_{\lambda_j > 0} \alpha_j u_j \in N_+$

$$\int_\Omega (|\nabla v|^2 - k^2 q v v) \, \mathrm{d}x = \sum_{\lambda_j > 0} \alpha_j \int_\Omega (\nabla u_j \cdot \nabla v - k^2 q u_j v) \, \mathrm{d}x$$
$$= - \sum_{\lambda_j > 0} \alpha_j \lambda_j \int_\Omega u_j v \, \mathrm{d}x = - \sum_{\lambda_j > 0} \alpha_j^2 \lambda_j \int_\Omega u_j^2 \, \mathrm{d}x < 0.$$

Since every $v \in N_-$ is an $H^1(\Omega)$-limit of finite linear combinations of u_j with $\lambda_j \le 0$, (16) follows with the same argument.

(d) Inequality (15) can be written as

$$\langle (K + k^2 K_q) v, v \rangle > \|v\|^2 \quad \text{for all } v \in N_+.$$

Lemma 3.2(b) implies that the number $d(q)$ of eigenvalues of $K + k^2 K_q$ larger than 1 must be at least $\dim(N_+)$. Likewise, (16) can be written as

$$\langle (K + k^2 K_q) v, v \rangle \le \|v\|^2 \quad \text{for all } v \in N_-.$$

Hence, Lemma 3.2(a) shows that $d(q)$ is at most $\operatorname{codim}(N_-) = \dim(N_+)$.

(e) This is trivial. □

Corollary 3.11. *If $q \in L^\infty(\Omega)$ and $q(x) \le q_{\max} \in \mathbb{R}$ for all $x \in \Omega$ (a.e.), then $d(q) \le d(q_{\max})$, and $d(q_{\max})$ is the number of Neumann eigenvalues of the Laplacian Δ that are larger than $-k^2 q_{\max}$.*

Proof. Obviously, the number of positive Neumann eigenvalues of $\Delta + k^2 q_{\max}$ agrees with the number of Neumann eigenvalues of the Laplacian Δ that are greater than $-k^2 q_{\max}$. Hence, the assertion follows from Lemmas 3.9 and 3.10(d). □

Remark 3.12. One can show, by using constant potentials, that for the Helmholtz equation $\Lambda_{q_2} - \Lambda_{q_1}$ can actually have negative eigenvalues when $q_1 \le q_2$. This shows that in Theorem 3.5 it is indeed necessary to work modulo a finite-dimensional subspace. The details will appear in a subsequent work.

4. Localized potentials for the Helmholtz equation

We now extend the result in [Gebauer 2008] to the Helmholtz equation and prove that we can control the energy terms appearing in the monotonicity relation in spaces of finite codimension. We will first state the result and prove it using a functional analytic relation between operator norms and the ranges of their adjoints in Section 4A. Section 4B then gives an alternative proof that is based on a Runge approximation argument.

4A. ***Localized potentials.*** Our main result on controlling the solutions of the Helmholtz equation in spaces of finite codimension is the following theorem.

Theorem 4.1. *Let $q \in L^\infty(\Omega) \setminus \{0\}$ for which $k > 0$ is not a resonance. Let $B, D \subseteq \overline{\Omega}$ be measurable, $B \setminus \overline{D}$ possess positive measure, and $\overline{\Omega} \setminus \overline{D}$ be connected to Σ.*

Then for any subspace $V \subset L^2(\Sigma)$ with $\dim V < \infty$, there exists a sequence $(g_j)_{j\in\mathbb{N}} \subset V^\perp$ such that

$$\int_B |u_q^{(g_j)}|^2 \,\mathrm{d}x \to \infty \quad \textit{and} \quad \int_D |u_q^{(g_j)}|^2 \,\mathrm{d}x \to 0,$$

where $u_q^{(g_j)} \in H^1(\Omega)$ solves the Helmholtz equation (2) with Neumann boundary data g_j.

The arguments that we will use to prove Theorem 4.1 in this subsection also yield a simple proof for the following elementary result. We formulate it as a theorem since we will utilize it in the next section to control energy terms in monotonicity inequalities for different scattering coefficients.

Theorem 4.2. *Let $q_1, q_2 \in L^\infty(\Omega) \setminus \{0\}$ for which $k > 0$ is not a resonance. If $q_1(x) = q_2(x)$ for all x (a.e.) outside a measurable set $D \subset \Omega$, then there exist constants $c_1, c_2 > 0$ such that*

$$c_1 \int_D |u_1^{(g)}|^2 \,\mathrm{d}x \le \int_D |u_2^{(g)}|^2 \,\mathrm{d}x \le c_2 \int_D |u_1^{(g)}|^2 \,\mathrm{d}x \quad \textit{for all } g \in L^2(\Sigma),$$

where $u_1^{(g)}, u_2^{(g)} \in H^1(\Omega)$ solve the Helmholtz equation (2) with Neumann boundary data g and $q = q_1$, $q = q_2$ respectively.

To prove Theorems 4.1 and 4.2 we will formulate and prove several lemmas. Let us first note that the assertion of Theorem 4.1 already holds if we can prove it for a subset of B with positive measure. We will use the subset $B \cap C$, where C is a small closed ball constructed in the next lemma.

Lemma 4.3. *Let $B, D \subseteq \overline{\Omega}$ be measurable, $B \setminus \overline{D}$ possess positive measure, and $\overline{\Omega} \setminus \overline{D}$ be connected to Σ. Then there exists a closed ball C such that $B \cap C$ has positive measure, $C \cap \overline{D} = \varnothing$, and $\overline{\Omega} \setminus (\overline{D} \cup C)$ is connected to Σ.*

Proof. Let x be a point of Lebesgue density 1 in $B \setminus \overline{D}$. Then the closure C of a sufficiently small ball centered in x will satisfy $B \cap C$ has positive measure, $C \cap \overline{D} = \varnothing$, and $\overline{\Omega} \setminus (\overline{D} \cup C)$ is connected to Σ. □

Now we follow the general approach in [Gebauer 2008]. We formulate the energy terms in Theorem 4.1 as norms of operator evaluations and characterize their adjoints. Then we characterize the ranges of the adjoints using the unique continuation property, and prove Theorem 4.1 using a functional-analytic relation between norms of operator evaluations and ranges of their adjoints.

Lemma 4.4. *Let $q \in L^\infty(\Omega) \setminus \{0\}$ for which $k > 0$ is not a resonance. For a measurable set $D \subset \Omega$ we define*

$$L_D : L^2(\Sigma) \to L^2(D), \quad g \mapsto u|_D,$$

where $u \in H^1(\Omega)$ solves (2). Then L_D is a compact linear operator, and its adjoint satisfies

$$L_D^* : L^2(D) \to L^2(\Sigma), \quad f \mapsto v|_\Sigma,$$

where v solves

$$\Delta v + k^2 q v = f\chi_D, \quad \partial_\nu v|_{\partial\Omega} = 0. \tag{18}$$

Proof. With the operators I, j, and K_q defined as in Section 2A and (4) we have

$$L_D = R_D j (I - K - k^2 K_q)^{-1} \gamma_\Sigma^*,$$

where $R_D : L^2(\Omega) \to L^2(D)$ is the restriction operator $v \to v|_D$. Hence, L_D is a linear compact operator, and its adjoint is

$$L_D^* = \gamma_\Sigma (I - K - k^2 K_q)^{-1} j^* R_D^*.$$

Thus $L_D^* f = v|_\Sigma$, where $v \in H^1(\Omega)$ solves

$$(I - K - k^2 K_q) v = j^* R_D^* f;$$

i.e., for all $w \in H^1(\Omega)$,

$$\int_\Omega (\nabla v \cdot \nabla w - k^2 q v w)\,\mathrm{d}x = \langle (I - K - k^2 K_q) v, w \rangle = \langle j^* R_D^* f, w \rangle = \int_D f w \,\mathrm{d}x,$$

which is the variational formulation equivalent to (18). □

Lemma 4.5. *Let $q \in L^\infty(\Omega) \setminus \{0\}$ for which $k > 0$ is not a resonance. Let $B, D \subseteq \overline{\Omega}$ be measurable and $C \subseteq \overline{\Omega}$ be a closed set such that $B \cap C$ has positive measure, $C \cap \overline{D} = \varnothing$, and $\overline{\Omega} \setminus (\overline{D} \cup C)$ is connected to Σ. Then,*

$$\mathcal{R}(L_{B\cap C}^*) \cap \mathcal{R}(L_D^*) = \{0\}, \tag{19}$$

and $\mathcal{R}(L_{B\cap C}^*)$, $\mathcal{R}(L_D^*) \subset L^2(\Sigma)$ are both dense (and thus in particular infinite-dimensional).

Proof. It follows from the unique continuation property in Theorem 2.4(a) that $L_{B\cap C}$ and L_D are injective. Hence $\mathcal{R}(L_{B\cap C}^*)$ and $\mathcal{R}(L_D^*)$ are dense subspaces of $L^2(\Sigma)$.

The characterization of the adjoint operators in Lemma 4.4 shows that

$$B \cap C \subseteq C \quad \text{implies} \quad \mathcal{R}(L_{B\cap C}^*) \subseteq \mathcal{R}(L_C^*).$$

Hence, (19) follows a fortiori if we can show that

$$\mathcal{R}(L_C^*) \cap \mathcal{R}(L_D^*) = \{0\}.$$

To show this let $h \in \mathcal{R}(L_C^*) \cap \mathcal{R}(L_D^*)$. Then there exist $f_C \in L^2(C)$, $f_D \in L^2(D)$, and $v_C, v_D \in H^1(\Omega)$ such that

$$\Delta v_C + k^2 q v_C = f_C \chi_C, \quad \partial_\nu v|_{\partial\Omega} = 0,$$
$$\Delta v_D + k^2 q v_D = f_D \chi_D, \quad \partial_\nu v|_{\partial\Omega} = 0,$$

and $v_C|_\Sigma = h = v_D|_\Sigma$.

It follows from the unique continuation property in Theorem 2.4(b) that $v_C = v_D$ on the connected set $\Omega \setminus (C \cup \overline{D})$. Hence,

$$v := \begin{cases} v_C = v_D & \text{on } \Omega \setminus (C \cup \overline{D}), \\ v_C & \text{on } \overline{D}, \\ v_D & \text{on } C \end{cases}$$

defines an $H^1(\Omega)$-function solving

$$\Delta v + k^2 q v = 0, \quad \partial_\nu v|_{\partial\Omega} = 0,$$

so that $v = 0$ and thus $h = v_C|_\Sigma = v_D|_\Sigma = v|_\Sigma = 0$. □

Lemma 4.6. *Let X, Y and Z be Hilbert spaces, and $A_1 : X \to Y$ and $A_2 : X \to Z$ be linear bounded operators. Then*

$$\exists c > 0 : \|A_1 x\| \le c\|A_2 x\| \quad \forall x \in X \qquad \textit{if and only if} \qquad \mathcal{R}(A_1^*) \subseteq \mathcal{R}(A_2^*).$$

Proof. This is proven for reflexive Banach spaces in [Gebauer 2008, Lemma 2.5]. Note that one direction of the implication also holds in nonreflexive Banach spaces; see [Gebauer 2008, Lemma 2.4]. □

Lemma 4.7. *Let $V, X, Y \subset Z$ be subspaces of a real vector space Z. If*

$$X \cap Y = \{0\} \quad \textit{and} \quad X \subseteq Y + V,$$

then $\dim(X) \le \dim(V)$.

Proof. Let $(x_j)_{j=1}^m \subset X$ be a linearly independent sequence of m vectors. Then there exist $(y_j)_{j=1}^m \subset Y$ and $(v_j)_{j=1}^m \subset V$ such that $x_j = y_j + v_j$ for all $j = 1, \dots, m$. We will prove the assertion by showing that the sequence $(v_j)_{j=1}^m$ is linearly independent. To this end let $\sum_{j=1}^m a_j v_j = 0$ with $a_j \in \mathbb{R}$, $j = 1, \dots, m$. Then

$$\sum_{j=1}^m a_j x_j = \sum_{j=1}^m a_j (y_j + v_j) = \sum_{j=1}^m a_j y_j \in Y,$$

so that $\sum_{j=1}^m a_j x_j = 0$. Since $(x_j)_{j=1}^m \subset X$ is linearly independent, it follows that $a_j = 0$ for all $j = 1, \dots, m$. This shows that $(v_j)_{j=1}^m$ is linearly independent. □

Proof of Theorem 4.1. Let $q \in L^\infty(\Omega) \setminus \{0\}$ for which $k > 0$ is not a resonance. Let $B, D \subseteq \overline{\Omega}$ be measurable, $B \setminus \overline{D}$ possess positive measure, and $\overline{\Omega} \setminus \overline{D}$ be connected to Σ. Using Lemma 4.3 we obtain a closed set $C \subseteq \overline{\Omega}$ such that $B \cap C$ has positive measure, $C \cap \overline{D} = \varnothing$, and $\overline{\Omega} \setminus (\overline{D} \cup C)$ is connected to Σ.

Let $V \subset L^2(\Sigma)$ be a subspace with $d := \dim(V) < \infty$. Since V is finite-dimensional and thus closed, there exists an orthogonal projection operator $P_V : L^2(\Sigma) \to L^2(\Sigma)$ with

$$\mathcal{R}(P_V) = V, \quad P_V^2 = P_V, \quad \text{and} \quad P_V = P_V^*.$$

From Lemma 4.5, we have that $\mathcal{R}(L_{B\cap C}^*) \cap \mathcal{R}(L_D^*) = 0$ and that $\mathcal{R}(L_{B\cap C}^*)$ is infinite-dimensional. So it follows from Lemma 4.7 that

$$\mathcal{R}(L_{B\cap C}^*) \not\subseteq \mathcal{R}(L_D^*) + V = \mathcal{R}(L_D^*) + \mathcal{R}(P_V^*).$$

Since $B \cap C \subseteq B$ implies that $\mathcal{R}(L_{B\cap C}^*) \subseteq \mathcal{R}(L_B^*)$, and since (using block operator matrix notation)

$$\mathcal{R}((L_D^* \;\; P_V^*)) \subseteq \mathcal{R}(L_D^*) + \mathcal{R}(P_V^*),$$

we obtain that

$$\mathcal{R}(L_B^*) \not\subseteq \mathcal{R}((L_D^* \;\; P_V^*)) = \mathcal{R}\left(\begin{pmatrix} L_D \\ P_V \end{pmatrix}^*\right).$$

It then follows from Lemma 4.6 that there cannot exist a constant $C > 0$ with

$$\|L_B g\|^2 \leq C^2 \left\| \begin{pmatrix} L_D \\ P_V \end{pmatrix} g \right\|^2 = C^2\|L_D g\|^2 + C^2\|P_V g\|^2 \quad \text{for all } g \in L^2(\Sigma).$$

Hence, there must exist a sequence $(\tilde{g}_k)_{k\in\mathbb{N}} \subseteq L^2(\Sigma)$ with

$$\|L_B \tilde{g}_k\| \to \infty \quad \text{and} \quad \|L_D \tilde{g}_k\|, \|P_V \tilde{g}_k\| \to 0.$$

Thus, $g_k := \tilde{g}_k - P_V \tilde{g}_k \in V^\perp \subseteq L^2(\Sigma)$ and

$$\|L_B g_k\| \geq \|L_B \tilde{g}_k\| - \|L_B\|\,\|P_V \tilde{g}_k\| \to \infty \quad \text{and} \quad \|L_D g_k\| \to 0,$$

which shows the assertion. □

Proof of Theorem 4.2. Let $q_1, q_2 \in L^\infty(\Omega)$ for which $k > 0$ is not a resonance, and let $q_1(x) = q_2(x)$ for all x (a.e.) outside a measurable set $D \subset \Omega$. We denote by $L_{q_1,D}$ and $L_{q_2,D}$ the operators from Lemma 4.4 for $q = q_1$ and $q = q_2$. For $f \in L^2(D)$, we then have

$$L^*_{q_1,D} f = v_1|_\Sigma \quad \text{and} \quad L^*_{q_2,D} f = v_2|_\Sigma,$$

where $v_1, v_2 \in H^1(\Omega)$ solve

$$\begin{aligned} \Delta v_1 + k^2 q_1 v_1 &= f\chi_D, \quad \partial_\nu v_1|_{\partial\Omega} = 0,\\ \Delta v_2 + k^2 q_2 v_2 &= f\chi_D, \quad \partial_\nu v_2|_{\partial\Omega} = 0. \end{aligned}$$

Since this also implies

$$\begin{aligned} \Delta v_1 + k^2 q_2 v_1 &= f\chi_D + k^2(q_2 - q_1)v_1, \quad \partial_\nu v_1|_{\partial\Omega} = 0,\\ \Delta v_2 + k^2 q_1 v_2 &= f\chi_D + k^2(q_1 - q_2)v_2, \quad \partial_\nu v_2|_{\partial\Omega} = 0, \end{aligned}$$

and $q_1 - q_2$ vanishes (a.e.) outside D, it follows that

$$v_1|_\Sigma = L^*_{q_2,D}(f + k^2(q_2 - q_1)v_1) \quad \text{and} \quad v_2|_\Sigma = L^*_{q_1,D}(f + k^2(q_1 - q_2)v_2).$$

Hence, $\mathcal{R}(L^*_{q_1,D}) = \mathcal{R}(L^*_{q_2,D})$, so that the assertion follows from Lemma 4.6. □

4B. *Localized potentials and Runge approximation.* In this subsection we give an alternative proof of Theorem 4.1 that is based on a Runge approximation argument that characterizes whether a given function $\varphi \in L^2(O)$ on a measurable subset $O \subseteq \Omega$ can be approximated by functions in a subspace of solutions of the Helmholtz equation in Ω. Throughout this subsection let $q \in L^\infty(\Omega) \setminus \{0\}$ for which $k > 0$ is not a resonance. We will prove the following theorem.

Theorem 4.8. *Let $D \subseteq \Omega$ be a measurable set and $C \subset \Omega$ be a closed ball for which $C \cap \overline{D} = \varnothing$, and $\overline{\Omega} \setminus (C \cup \overline{D})$ is connected to Σ.*

Then for any subspace $V \subset L^2(\Sigma)$ with $\dim V < \infty$, there exists a function $\varphi \in L^2(C \cup \overline{D})$ that can be approximated (in the $L^2(C \cup \overline{D})$-norm) by solutions $u \in H^1(\Omega)$ of

$$(\Delta + k^2 q)u = 0 \quad \textit{in } \Omega \qquad \textit{with} \qquad \partial_\nu u|_{\partial\Omega\setminus\Sigma} = 0, \quad \partial_\nu u|_\Sigma \in V^\perp,$$

and satisfies

$$\varphi|_{\overline{D}} \equiv 0 \quad \textit{and} \quad \varphi|_B \not\equiv 0$$

for all subsets $B \subseteq C$ with positive measure.

Before we prove Theorem 4.8, let us first show that it implies Theorem 4.1.

Corollary 4.9. *Let $B, D \subseteq \overline{\Omega}$ be measurable, $B \setminus \overline{D}$ possess positive measure, and $\overline{\Omega} \setminus \overline{D}$ be connected to Σ. Then for any subspace $V \subset L^2(\Sigma)$ with $\dim V < \infty$, there exists a sequence $(g_j)_{j\in\mathbb{N}} \subset V^\perp$ such that*

$$\int_B |u_q^{(g_j)}|^2 \,\mathrm{d}x \to \infty \quad \textit{and} \quad \int_D |u_q^{(g_j)}|^2 \,\mathrm{d}x \to 0,$$

where $u_q^{(g_j)} \in H^1(\Omega)$ solves the Helmholtz equation (2) with Neumann boundary data g_j.

Proof. As in Lemma 4.3, we can find a closed ball $C \subset \Omega$ so that $B \cap C$ has positive measure, $C \cap \overline{D} = \varnothing$, and $\overline{\Omega} \setminus (\overline{D} \cup C)$ is connected to Σ. Using Theorem 4.8, there exists $\varphi \in L^2(C \cup \overline{D})$ and a sequence of solutions $(\tilde{u}^{(j)})_{j\in\mathbb{N}} \subset H^1(\Omega)$ of $(\Delta + k^2 q)\tilde{u}^{(j)} = 0$ in Ω with $\partial_\nu \tilde{u}^{(j)}|_{\partial\Omega\setminus\Sigma} = 0$, $\partial_\nu \tilde{u}^{(j)}|_\Sigma \in V^\perp$,

$$\|\tilde{u}^{(j)}|_{B\cap C}\|_{L^2(B\cap C)} \to \|\varphi\|_{L^2(B\cap C)} > 0 \quad \text{and} \quad \|\tilde{u}^{(j)}|_{\overline{D}}\|_{L^2(\overline{D})} \to 0.$$

Obviously, the scaled sequence

$$g^{(j)} := \frac{\partial_\nu \tilde{u}^{(j)}}{\sqrt{\|\tilde{u}^{(j)}|_{\overline{D}}\|_{L^2(\overline{D})}}} \in V^\perp$$

satisfies the assertion. ☐

To prove Theorem 4.8, we start with an abstract characterization showing whether a given function $\varphi \in L^2(O)$ on a measurable set $O \subseteq \Omega$ is a limit of functions from a subspace of solutions of the Helmholtz equation in Ω. For the sake of readability, we write $v\chi_O \in L^2(\Omega)$ for the zero extension of a function $v \in L^2(O)$, and we write the dual pairing on $H^{-1/2}(\partial\Omega) \times H^{1/2}(\partial\Omega)$ as an integral over $\partial\Omega$.

Lemma 4.10. *Let $O \subseteq \Omega$ be measurable. Let $H \subseteq H^1(\Omega)$ be a (not necessarily closed) subspace of solutions of $(\Delta + k^2 q)u = 0$ in Ω.*

A function $\varphi \in L^2(O)$ can be approximated on O by solutions $u \in H$ in the sense that

$$\inf_{u\in H} \|\varphi - u\|_{L^2(O)} = 0$$

if and only if $\int_O \varphi v \,\mathrm{d}x = 0$ for all $v \in L^2(O)$ for which the solution $w \in H^1(\Omega)$ of

$$(\Delta + k^2 q)w = v\chi_O \quad \textit{and} \quad \partial_\nu w|_{\partial\Omega} = 0 \tag{20}$$

satisfies $\int_{\partial\Omega} \partial_\nu u|_{\partial\Omega} w|_{\partial\Omega} \,\mathrm{d}s = 0$ for all $u \in H$.

Proof. Let

$$\mathcal{R} := \{u|_O : u \in H\} \subseteq L^2(O).$$

Let $v \in L^2(O)$ and $w \in H^1(\Omega)$ solve (20). Then $v \in \mathcal{R}^\perp$ if and only if, for all $u \in H$,

$$0 = \int_O uv \,\mathrm{d}x = \int_\Omega u(\Delta + k^2 q)w \,\mathrm{d}x = \int_\Omega w(\Delta + k^2 q)u \,\mathrm{d}x - \int_{\partial\Omega} \partial_\nu u|_{\partial\Omega} w|_{\partial\Omega} \,\mathrm{d}s = -\int_{\partial\Omega} \partial_\nu u|_{\partial\Omega} w|_{\partial\Omega} \,\mathrm{d}s.$$

Hence, the assertion follows from $\overline{\mathcal{R}} = (\mathcal{R}^\perp)^\perp$ (where orthogonality and closures are taken with respect to the $L^2(O)$ inner product). □

Now we characterize the functions w appearing in Lemma 4.10 for a setting that will be considered in the proof of Theorem 4.8.

Lemma 4.11. *Let V be a finite-dimensional subspace of $L^2(\Sigma)$, and $O \subset \Omega$ be a closed set for which the complement $\overline{\Omega} \setminus O$ is connected to Σ.*

We define the spaces

$$W := \{w \in H^1(\Omega) : \exists v \in L^2(O) \text{ s.t. } (\Delta + k^2 q)w = v\chi_O,\ \partial_\nu w|_{\partial\Omega} = 0,\ w|_\Sigma \in V\},$$
$$W_0 := \{w \in H^1(\Omega) : \exists v \in L^2(O) \text{ s.t. } (\Delta + k^2 q)w = v\chi_O,\ \partial_\nu w|_{\partial\Omega} = 0,\ w|_\Sigma = 0\}.$$

Then the codimension $d := \dim(W/W_0)$ of W_0 in W is at most $\dim(V)$; i.e., there exist functions $w_1, \ldots, w_d \in W$ such that every $w \in W$ can be written as

$$w = w_0 + \sum_{j=1}^{d} a_j w_j,$$

with (w-dependent) $w_0 \in W_0$ and $a_1, \ldots, a_d \in \mathbb{R}$.

Proof. W_0 is the kernel of the restricted trace operator

$$\gamma_\Sigma|_W : W \to V, \quad w \mapsto w|_\Sigma.$$

Hence, the codimension of W_0 as a subspace of W is

$$\dim(W/W_0) = \dim(\mathcal{R}(\gamma_\Sigma|_W)) \leq \dim(V),$$

which proves the assertion. □

Proof of Theorem 4.8. Let $D \subseteq \Omega$ be a measurable set and $C \subset \Omega$ be a closed ball for which $C \cap \overline{D} = \varnothing$ and $\overline{\Omega} \setminus (C \cup \overline{D})$ is connected to Σ. Let V be a finite-dimensional subspace of $L^2(\Sigma)$.

To apply Lemma 4.10, we set $O := C \cup \overline{D}$ and

$$H := \{u \in H^1(\Omega) : (\Delta + k^2 q)u = 0 \text{ in } \Omega,\ \partial_\nu u|_{\partial\Omega\setminus\Sigma} = 0,\ \partial_\nu u|_\Sigma \in V^\perp\}.$$

Then $w \in H^1(\Omega)$ satisfies (20) and

$$\int_{\partial\Omega} \partial_\nu u|_{\partial\Omega} w|_{\partial\Omega}\, \mathrm{d}s = 0$$

for all $u \in H$if and only if $w \in W$, with W defined in Lemma 4.11. Hence, by Lemma 4.10, a function $\varphi \in L^2(C \cup \overline{D})$ can be approximated by solutions $u \in H$ if and only if

$$\int_{C\cup\overline{D}} \varphi(\Delta + k^2 q)w\, \mathrm{d}x = 0 \quad \text{for all } w \in W. \tag{21}$$

Thus, the assertion of Theorem 4.8 follows if we can show that there exists $\varphi \in L^2(C \cup \overline{D})$ that satisfies (21) and vanishes on D but not on any subset of C having positive measure.

To construct such a φ, we first note that the Helmholtz equation (2) on Ω is uniquely solvable for all Neumann data $g \in L^2(\Sigma)$, and by unique continuation, linearly independent Neumann data yield solutions whose restrictions to the open ball C° are linearly independent. Hence, there exists an infinite number of linearly independent solutions

$$\varphi_j \in H^1(C^\circ) \qquad \text{with} \qquad (\Delta + k^2 q)\varphi_j = 0 \quad \text{in } C^\circ, \quad j \in \mathbb{N}. \tag{22}$$

We extend φ_j by zero on $\overline{D} \cup \partial C$ to $\varphi_j \in L^2(O)$.

Every $w_0 \in W_0$, with W_0 from Lemma 4.11, must possess zero Cauchy data $w_0|_{\partial C} = 0$ and $\partial_\nu w_0|_{\partial C} = 0$ by unique continuation. Hence, for all $w_0 \in W_0$, and $j \in \mathbb{N}$,

$$\int_O \varphi_j (\Delta + k^2 q) w_0 \, dx = \int_C \varphi_j (\Delta + k^2 q) w_0 \, dx = \int_{\partial C} (\varphi_j|_{\partial C} \partial_\nu w_0|_{\partial C} - \partial_\nu \varphi_j|_{\partial C} w_0|_{\partial C}) \, ds = 0.$$

Moreover, by a dimensionality argument, there must exist a nontrivial finite linear combination φ of the infinitely many linearly independent φ_j such that

$$\int_O \varphi(\Delta + k^2 q) w_k \, dx = 0$$

for the finitely many functions $w_1, \dots, w_d \in W$ from Lemma 4.11. Thus, using Lemma 4.11, we have constructed a function $\varphi \in L^2(O)$ with $\varphi|_{\overline{D}} \equiv 0$, $\varphi|_{C^\circ} \not\equiv 0$, and

$$\int_O \varphi(\Delta + k^2 q) w \, dx = 0 \quad \text{for all } w \in W = W_0 + \operatorname{span}\{w_1, \dots, w_d\}.$$

Moreover, φ solves (22), so that the unique continuation result from measurable sets in Theorem 2.4 also yields that $\varphi|_B \not\equiv 0$ for all $B \subseteq C^\circ$ with positive measure. Since ∂C is a null set, the latter also holds for all $B \subseteq C$ with positive measure. As explained above, the assertion of Theorem 4.8 now follows from Lemma 4.10. □

5. Local uniqueness for the Helmholtz equation

We are now able to prove the first main result in this work, announced as Theorem 1.1 in the Introduction, and extend the local uniqueness result in [Harrach and Ullrich 2017] to the case of negative potentials and $n \geq 2$.

As in Section 2A, let $\Omega \subset \mathbb{R}^n$, $n \geq 2$, denote a bounded Lipschitz domain, and let $\Sigma \subseteq \partial\Omega$ be an arbitrarily small, relatively open part of the boundary $\partial\Omega$. For $q_1, q_2 \in L^\infty(\Omega)$ let

$$\Lambda(q_1), \Lambda(q_2) : L^2(\Sigma) \to L^2(\Sigma), \quad \Lambda(q_1) : g \mapsto u_1|_\Sigma, \quad \Lambda(q_2) : g \mapsto u_2|_\Sigma,$$

be the Neumann-to-Dirichlet operators for the Helmholtz equation

$$(\Delta + k^2 q)u = 0 \quad \text{in } \Omega, \qquad \partial_\nu u|_{\partial\Omega} = \begin{cases} g & \text{on } \Sigma, \\ 0 & \text{else,} \end{cases} \tag{23}$$

with $q = q_1$ and $q = q_2$ respectively, and let $k > 0$ be such that it is not a resonance for q_1 or q_2.

Theorem 5.1. *Let $q_1 \leq q_2$ in a relatively open set $O \subseteq \overline{\Omega}$ that is connected to Σ. Then*

$$q_1|_O \not\equiv q_2|_O \quad \text{implies} \quad \Lambda(q_1) \neq \Lambda(q_2).$$

Moreover, in that case, $\Lambda(q_2) - \Lambda(q_1)$ has infinitely many positive eigenvalues.

Proof. If $q_1|_O \not\equiv q_2|_O$ then there exists a subset $B \subseteq O$ with positive measure, and a constant $c > 0$ such that $q_2(x) - q_1(x) \geq c$ for all $x \in B$ (a.e.). From the monotonicity inequality in Theorem 3.5 we have that $\Lambda(q_2) - \Lambda(q_1) \geq_{\text{fin}} A$, where

$$A : L^2(\Sigma) \to L^2(\Sigma), \qquad \int_\Sigma hAg \,\mathrm{d}s = \int_\Omega k^2(q_2 - q_1) u_1^{(g)} u_1^{(h)} \,\mathrm{d}x.$$

Note that $A = S_1^* j^* k^2 M_{q_1 - q_2} j S_1$, where $S_1 : g \mapsto u_1^{(g)}$ is the solution operator and $j : H^1(\Omega) \to L^2(\Omega)$ is the compact inclusion, so A is indeed a compact, self-adjoint linear operator on $L^2(\Sigma)$.

We will now prove the assertion by contradiction and assume that $\Lambda(q_2) - \Lambda(q_1) \leq_{\text{fin}} 0$. Then, the transitivity result in Lemma 3.4 gives that $A \leq_{\text{fin}} 0$. By the characterization in Corollary 3.3, there would exist a finite-dimensional subspace $V \subseteq L^2(\partial\Omega)$, with

$$\begin{aligned} 0 &\geq \int_\Omega k^2(q_2 - q_1)|u_1^{(g)}|^2 \,\mathrm{d}x \\ &= \int_O k^2(q_2 - q_1)|u_1^{(g)}|^2 \,\mathrm{d}x + \int_{\Omega \setminus O} k^2(q_2 - q_1)|u_1^{(g)}|^2 \,\mathrm{d}x \\ &\geq c \int_B k^2 |u_1^{(g)}|^2 \,\mathrm{d}x - C \int_{\Omega \setminus O} k^2 |u_1^{(g)}|^2 \,\mathrm{d}x \end{aligned}$$

for all $g \in V^\perp$, where $C := (\|q_1\|_{L^\infty(\Omega)} + \|q_2\|_{L^\infty(\Omega)})$ and $u_1^{(g)}$ solves (23) with $q = q_1$.

However, using the localized potentials from Theorem 4.1 with $D := \overline{\Omega} \setminus O$, there must exist a Neumann datum $g \in V^\perp$ with

$$c \int_B k^2 |u_1^{(g)}|^2 \,\mathrm{d}x > C \int_{\Omega \setminus O} k^2 |u_1^{(g)}|^2 \,\mathrm{d}x,$$

which contradicts the above inequality. Hence, $\Lambda(q_2) - \Lambda(q_1)$ must have infinitely many negative eigenvalues, and in particular $\Lambda(q_2) \neq \Lambda(q_1)$. □

Proof of Theorem 1.1. The result is an immediate consequence of Theorem 5.1. □

Theorem 5.1 shows that two scattering coefficient functions can be distinguished from knowledge of the partial boundary measurements if their difference is of definite sign in a neighborhood of Σ (or any open subset of Σ since $\Lambda(\Sigma)$ determines the boundary measurements on all smaller parts). This definite sign condition is satisfied for piecewise-analytic functions, see, e.g., [Harrach and Ullrich 2013, Theorem A.1], but the authors are not aware of other named function spaces, with less regularity, where infinite oscillations between positive and negative values when approaching the boundary can be ruled out. In the following corollary the term piecewise-analytic is understood with respect to a partition in finitely many subdomains with piecewise C^∞-boundaries; see [Harrach and Ullrich 2013] for a precise definition.

Corollary 5.2. *If $q_1 - q_2$ is piecewise-analytic on Ω then*

$$\Lambda(q_1) = \Lambda(q_2) \quad \textit{if and only if} \quad q_1 = q_2.$$

Proof. This follows from Theorem 1.1 and [Harrach and Ullrich 2013, Theorem A.1]. □

6. Detecting the support of a scatterer

We will now show that an unknown scatterer, where the refraction index is either higher or lower than an otherwise homogeneous background value, can be reconstructed by simple monotonicity comparisons.

6A. ***Scatterer detection by monotonicity tests.*** As before, let $\Omega \subset \mathbb{R}^n$, $n \geq 2$, be a bounded domain with Lipschitz boundary. The domain is assumed to contain an open set (the scatterer) $D \subseteq \Omega$ with $\overline{D} \subset \Omega$ and connected complement $\Omega \setminus \overline{D}$. We assume that the scattering index satisfies $q(x) = 1$ in $\Omega \setminus D$ (a.e.) and that there exist constants $q_{\min}, q_{\max} \in \mathbb{R}$ so that either

$$1 < q_{\min} \leq q(x) \leq q_{\max} \quad \text{for all } x \in D \text{ (a.e.)}$$

or

$$q_{\min} \leq q(x) \leq q_{\max} < 1 \quad \text{for all } x \in D \text{ (a.e.)}.$$

$\Lambda(q)$ denotes the Neumann-to-Dirichlet operator for the domain containing the scatterer, and $\Lambda(1)$ is the Neumann-to-Dirichlet operator for a homogeneous domain with $q \equiv 1$. For both cases, we assume that $k > 0$ is not a resonance.

For an open set $B \subseteq \Omega$ (e.g., a small ball), we define the operator

$$T_B : L^2(\Sigma) \to L^2(\Sigma), \qquad \int_\Sigma g T_B h \, \mathrm{d}s := \int_B k^2 u_1^{(g)} u_1^{(h)} \, \mathrm{d}x,$$

where $u_1^{(g)}, u_1^{(h)} \in H^1(\Omega)$ solve (2) with $q \equiv 1$ and Neumann boundary data g and h respectively. Obviously, T_B is a compact self-adjoint linear operator.

The following two theorems show that D can be reconstructed by comparing $\Lambda(q) - \Lambda(1)$ with T_B in the sense of the Loewner order up to finitely many eigenvalues introduced in Section 3A.

Theorem 6.1. *Let*

$$1 < q_{\min} \leq q(x) \leq q_{\max} \quad \textit{for all } x \in D \textit{ (a.e.)},$$

and let $d(q_{\max})$ be defined as in Lemma 2.1 (which also equals the number of Neumann eigenvalues of the Laplacian Δ that are larger than $-k^2 q_{\max}$; see Corollary 3.11).

(a) *If $B \subseteq D$ then*

$$\alpha T_B \leq_{d(q_{\max})} \Lambda(q) - \Lambda(1) \quad \textit{for all } \alpha \leq q_{\min} - 1.$$

(b) *If $B \not\subseteq D$ then, for all $\alpha > 0$, $\Lambda(q) - \Lambda(1) - \alpha T_B$ has infinitely many negative eigenvalues.*

Theorem 6.2. *Let*

$$q_{\min} \leq q(x) \leq q_{\max} < 1 \quad \textit{for all } x \in D \textit{ (a.e.)},$$

and let $d(1)$ be defined as in Lemma 2.1 (which also equals the number of Neumann eigenvalues of the Laplacian Δ that are larger than $-k^2$; see Corollary 3.11).

(a) *If $B \subseteq D$ then there exists $\alpha_{\max} > 0$ such that*

$$\alpha T_B \leq_{d(1)} \Lambda(1) - \Lambda(q) \quad \textit{for all } \alpha \leq \alpha_{\max}.$$

(b) *If $B \not\subseteq D$ then, for all $\alpha > 0$, $\Lambda(1) - \Lambda(q) - \alpha T_B$ has infinitely many negative eigenvalues.*

6B. *Proof of Theorems 6.1 and 6.2.* We prove both results by combining the monotonicity relations and localized potentials results from the last subsections.

Proof of Theorem 6.1. By the monotonicity relation in Theorem 3.5 there exists a subspace $V \subset L^2(\Sigma)$ with $\dim(V) \leq d(q) \leq d(q_{\max})$ (see Corollary 3.11) and

$$\int_\Sigma g(\Lambda(q) - \Lambda(1))g \, \mathrm{d}s \geq \int_\Omega k^2(q-1)|u_1^{(g)}|^2 \, \mathrm{d}x \quad \text{for all } g \in V^\perp.$$

If $B \subseteq D$ and $\alpha \leq q_{\min} - 1$, then $q - 1 \geq \alpha\chi_B$, so that for all $g \in L^2(\Sigma)$

$$\int_\Omega k^2(q-1)|u_1^{(g)}|^2 \, \mathrm{d}x \geq \int_B k^2\alpha|u_1^{(g)}|^2 \, \mathrm{d}x = \alpha \int_\Sigma g T_B g \, \mathrm{d}s.$$

Hence, if $B \subseteq D$ and $\alpha \leq q_{\min} - 1$, then

$$\int_\Sigma g(\Lambda(q) - \Lambda(1))g \, \mathrm{d}s \geq \alpha \int_\Sigma g T_B g \, \mathrm{d}s \quad \text{for all } g \in V^\perp,$$

which proves (a).

To prove (b) by contradiction, let $B \not\subseteq D$, $\alpha > 0$, and assume that

$$\Lambda(q) - \Lambda(1) \geq_{\text{fin}} \alpha T_B. \tag{24}$$

Using the monotonicity relation in Remark 3.6 together with Theorem 4.2, there exists a finite-dimensional subspace $V \subset L^2(\Sigma)$ and a constant $C > 0$, so that for all $g \in V^\perp$

$$\int_\Sigma g(\Lambda(q) - \Lambda(1))g \, \mathrm{d}s \leq \int_D k^2(q-1)|u_q^{(g)}|^2 \, \mathrm{d}x \leq C \int_D k^2(q-1)|u_1^{(g)}|^2 \, \mathrm{d}x. \tag{25}$$

Combining (24) and (25) using the transitivity result from Lemma 3.4, there exists a finite-dimensional subspace $\widetilde{V} \subset L^2(\Sigma)$ with

$$\alpha \int_B k^2|u_1^{(g)}|^2 \, \mathrm{d}x \leq C \int_D k^2(q-1)|u_1^{(g)}|^2 \, \mathrm{d}x \quad \text{for all } g \in \widetilde{V}^\perp.$$

However, this is contradicted by the localized potentials result in Theorem 4.1, which guarantees the existence of a sequence $(g_j)_{j\in\mathbb{N}} \subset \widetilde{V}^\perp$ with

$$\int_B |u_1^{(g_j)}|^2 \, \mathrm{d}x \to \infty \quad \text{and} \quad \int_D |u_1^{(g_j)}|^2 \, \mathrm{d}x \to 0.$$

Hence, $\Lambda(q) - \Lambda(1) - \alpha T_B$ cannot have only finitely many negative eigenvalues. □

Proof of Theorem 6.2. The proof is analogous to that of Theorem 6.1. We state it for the sake of completeness. Let

$$q_{\min} \le q(x) \le q_{\max} < 1 \quad \text{for all } x \in D \text{ (a.e.).}$$

If $B \subseteq D$, then by the monotonicity relation in Remark 3.6, together with Theorem 4.2, we have

$$\begin{aligned}\int_\Sigma g(\Lambda(q)-\Lambda(1))g\,\mathrm{d}s \le_{d(1)} \int_\Omega k^2(q-1)|u_q^{(g)}|^2\,\mathrm{d}x &\le -\int_D k^2(1-q_{\max})|u_q^{(g)}|^2\,\mathrm{d}x \\ &\le -c(1-q_{\max})\int_D k^2|u_1^{(g)}|^2\,\mathrm{d}x \le -c(1-q_{\max})\int_B k^2|u_1^{(g)}|^2\,\mathrm{d}x \\ &= -c(1-q_{\max})\int_\Sigma gT_Bg\,\mathrm{d}s,\end{aligned}$$

with a constant $c > 0$ from Theorem 4.2. This shows that $B \subseteq D$ implies

$$\alpha T_B \le_{d(1)} \Lambda(1)-\Lambda(q) \quad \text{for all } \alpha \le c(1-q_{\max}) =: \alpha_{\max},$$

so that (a) is proven.

To prove (b) by contradiction, let $B \not\subseteq D$, $\alpha > 0$, and assume that

$$\Lambda(1)-\Lambda(q) \ge_{\text{fin}} \alpha T_B. \tag{26}$$

By the monotonicity relation in Theorem 3.5, we have

$$\int_\Sigma g(\Lambda(1)-\Lambda(q))g\,\mathrm{d}s \le_{\text{fin}} \int_D k^2(1-q)|u_1^{(g)}|^2\,\mathrm{d}x. \tag{27}$$

Combining (26) and (27) using the transitivity result from Lemma 3.4, we have

$$\alpha\int_B k^2|u_1^{(g)}|^2\,\mathrm{d}x \le_{\text{fin}} \int_D k^2(1-q)|u_1^{(g)}|^2\,\mathrm{d}x.$$

However, this is contradicted by Theorem 4.1, which guarantees (for each finite-dimensional space $V \subset L^2(\Sigma)$) the existence of a sequence $(g_j)_{j\in\mathbb{N}} \subset V^\perp$ with

$$\int_B |u_0^{(g_j)}|^2\,\mathrm{d}x \to \infty \quad \text{and} \quad \int_D |u_0^{(g_j)}|^2\,\mathrm{d}x \to 0.$$

Hence, $\Lambda(1)-\Lambda(q)-\alpha T_B$ cannot have only finitely many negative eigenvalues, which shows (b). □

6C. *Remarks and extensions.* We finish this section with some remarks on possible extensions of our results. Theorems 6.1 and 6.2 hold with analogous proofs also for the case that the homogeneous background scattering index is replaced by a known inhomogeneous function $q_0 \in L^\infty(\Omega)$. Using the concept of the inner and outer support from [Harrach and Ullrich 2013] (see also [Kusiak and Sylvester 2003; Gebauer and Hyvönen 2008; Harrach and Seo 2010] for the origins of this concept), we can also treat the case where $\Omega \setminus \overline{D}$ is not connected or where there is no clear jump of the scattering index. The monotonicity tests will then determine D up to the difference of the inner and outer support. Moreover, the so-called indefinite case that the domain contains scatterers with higher and lower refractive indices can be treated by shrinking a large test region analogously to [Harrach and Ullrich 2013]; see also [Garde and Staboulis 2019].

Acknowledgements

Pohjola and Salo were supported by the Academy of Finland (Finnish Centre of Excellence in Inverse Problems Research, grant number 284715) and by an ERC Starting Grant (grant number 307023).

References

[Alessandrini 1990] G. Alessandrini, "Singular solutions of elliptic equations and the determination of conductivity by boundary measurements", *J. Differential Equations* **84**:2 (1990), 252–272. MR Zbl

[Alessandrini 2012] G. Alessandrini, "Strong unique continuation for general elliptic equations in 2D", *J. Math. Anal. Appl.* **386**:2 (2012), 669–676. MR Zbl

[Ammari and Uhlmann 2004] H. Ammari and G. Uhlmann, "Reconstruction of the potential from partial Cauchy data for the Schrödinger equation", *Indiana Univ. Math. J.* **53**:1 (2004), 169–183. MR Zbl

[Arnold and Harrach 2013] L. Arnold and B. Harrach, "Unique shape detection in transient eddy current problems", *Inverse Problems* **29**:9 (2013), art. id. 095004. MR Zbl

[Astala and Päivärinta 2006] K. Astala and L. Päivärinta, "Calderón's inverse conductivity problem in the plane", *Ann. of Math.* (2) **163**:1 (2006), 265–299. MR Zbl

[Astala et al. 2009] K. Astala, T. Iwaniec, and G. Martin, *Elliptic partial differential equations and quasiconformal mappings in the plane*, Princeton Math. Series **48**, Princeton Univ. Press, 2009. MR Zbl

[Barth et al. 2017] A. Barth, B. Harrach, N. Hyvönen, and L. Mustonen, "Detecting stochastic inclusions in electrical impedance tomography", *Inverse Problems* **33**:11 (2017), art. id. 115012. MR Zbl

[Brander et al. 2018] T. Brander, B. Harrach, M. Kar, and M. Salo, "Monotonicity and enclosure methods for the p-Laplace equation", *SIAM J. Appl. Math.* **78**:2 (2018), 742–758. MR Zbl

[Brühl 2001] M. Brühl, "Explicit characterization of inclusions in electrical impedance tomography", *SIAM J. Math. Anal.* **32**:6 (2001), 1327–1341. MR Zbl

[Brühl and Hanke 2000] M. Brühl and M. Hanke, "Numerical implementation of two noniterative methods for locating inclusions by impedance tomography", *Inverse Problems* **16**:4 (2000), 1029–1042. MR Zbl

[Calderón 1980] A. P. Calderón, "On an inverse boundary value problem", pp. 65–73 in *Seminar on numerical analysis and its applications to continuum physics* (Rio de Janeiro, 1980), edited by W. H. Meyer and M. A. Raupp, Coleção Atas **12**, Soc. Brasil. Mat., Rio de Janeiro, 1980. MR

[Calderón 2006] A. P. Calderón, "On an inverse boundary value problem", *Comput. Appl. Math.* **25**:2-3 (2006), 133–138. MR Zbl

[Caro and Rogers 2016] P. Caro and K. M. Rogers, "Global uniqueness for the Calderón problem with Lipschitz conductivities", *Forum Math. Pi* **4** (2016), art. id. e2. MR Zbl

[Druskin 1998] V. Druskin, "On the uniqueness of inverse problems from incomplete boundary data", *SIAM J. Appl. Math.* **58**:5 (1998), 1591–1603. MR Zbl

[de Figueiredo and Gossez 1992] D. G. de Figueiredo and J.-P. Gossez, "Strict monotonicity of eigenvalues and unique continuation", *Comm. Partial Differential Equations* **17**:1-2 (1992), 339–346. MR Zbl

[Garde 2018] H. Garde, "Comparison of linear and non-linear monotonicity-based shape reconstruction using exact matrix characterizations", *Inverse Probl. Sci. Eng.* **26**:1 (2018), 33–50. MR Zbl

[Garde and Staboulis 2017] H. Garde and S. Staboulis, "Convergence and regularization for monotonicity-based shape reconstruction in electrical impedance tomography", *Numer. Math.* **135**:4 (2017), 1221–1251. MR Zbl

[Garde and Staboulis 2019] H. Garde and S. Staboulis, "The regularized monotonicity method: detecting irregular indefinite inclusions", *Inverse Probl. Imaging* **13**:1 (2019), 93–116. Zbl

[Gebauer 2008] B. Gebauer, "Localized potentials in electrical impedance tomography", *Inverse Probl. Imaging* **2**:2 (2008), 251–269. MR Zbl

[Gebauer and Hyvönen 2008] B. Gebauer and N. Hyvönen, "Factorization method and inclusions of mixed type in an inverse elliptic boundary value problem", *Inverse Probl. Imaging* **2**:3 (2008), 355–372. MR Zbl

[Griesmaier and Harrach 2018] R. Griesmaier and B. Harrach, "Monotonicity in inverse medium scattering on unbounded domains", *SIAM J. Appl. Math.* **78**:5 (2018), 2533–2557. MR Zbl

[Guillarmou and Tzou 2013] C. Guillarmou and L. Tzou, "The Calderón inverse problem in two dimensions", pp. 119–166 in *Inverse problems and applications: inside out, II*, edited by G. Uhlmann, Math. Sci. Res. Inst. Publ. **60**, Cambridge Univ. Press, 2013. MR Zbl

[Haberman and Tataru 2013] B. Haberman and D. Tataru, "Uniqueness in Calderón's problem with Lipschitz conductivities", *Duke Math. J.* **162**:3 (2013), 497–516. MR Zbl

[Hadi and Tsouli 2001] I. E. Hadi and N. Tsouli, "Strong unique continuation of eigenfunctions for p-Laplacian operator", *Int. J. Math. Math. Sci.* **25**:3 (2001), 213–216. MR Zbl

[Harrach 2009] B. Harrach, "On uniqueness in diffuse optical tomography", *Inverse Problems* **25**:5 (2009), art. id. 055010. MR Zbl

[Harrach 2012] B. Harrach, "Simultaneous determination of the diffusion and absorption coefficient from boundary data", *Inverse Probl. Imaging* **6**:4 (2012), 663–679. MR Zbl

[Harrach 2013] B. Harrach, "Recent progress on the factorization method for electrical impedance tomography", *Comput. Math. Methods Med.* **2013** (2013), art. id. 425184. MR Zbl

[Harrach 2019] B. Harrach, "Uniqueness and Lipschitz stability in electrical impedance tomography with finitely many electrodes", *Inverse Problems* **35**:2 (2019), art. id. 024005. Zbl

[Harrach and Lin 2017] B. Harrach and Y.-H. Lin, "Monotonicity-based inversion of the fractional Schrödinger equation, I: Positive potentials", preprint, 2017. To appear in *SIAM J. Math. Anal.* arXiv

[Harrach and Lin 2019] B. Harrach and Y.-H. Lin, "Monotonicity-based inversion of the fractional Schrödinger equation, II: General potentials and stability", preprint, 2019. arXiv

[Harrach and Meftahi 2019] B. Harrach and H. Meftahi, "Global uniqueness and Lipschitz-stability for the inverse Robin transmission problem", *SIAM J. Appl. Math.* **79**:2 (2019), 525–550. MR Zbl

[Harrach and Minh 2016] B. Harrach and M. N. Minh, "Enhancing residual-based techniques with shape reconstruction features in electrical impedance tomography", *Inverse Problems* **32**:12 (2016), art. id. 125002. MR Zbl

[Harrach and Minh 2018] B. Harrach and M. N. Minh, "Monotonicity-based regularization for phantom experiment data in electrical impedance tomography", pp. 107–120 in *New trends in parameter identification for mathematical models* (Rio de Janeiro, 2017), edited by B. Hofmann et al., Birkhäuser, Cham, 2018. MR Zbl

[Harrach and Seo 2010] B. Harrach and J. K. Seo, "Exact shape-reconstruction by one-step linearization in electrical impedance tomography", *SIAM J. Math. Anal.* **42**:4 (2010), 1505–1518. MR Zbl

[Harrach and Ullrich 2013] B. Harrach and M. Ullrich, "Monotonicity-based shape reconstruction in electrical impedance tomography", *SIAM J. Math. Anal.* **45**:6 (2013), 3382–3403. MR Zbl

[Harrach and Ullrich 2015] B. Harrach and M. Ullrich, "Resolution guarantees in electrical impedance tomography", *IEEE Trans. Med. Imaging* **34**:7 (2015), 1513–1521. Zbl

[Harrach and Ullrich 2017] B. Harrach and M. Ullrich, "Local uniqueness for an inverse boundary value problem with partial data", *Proc. Amer. Math. Soc.* **145**:3 (2017), 1087–1095. MR Zbl

[Harrach et al. 2015] B. Harrach, E. Lee, and M. Ullrich, "Combining frequency-difference and ultrasound modulated electrical impedance tomography", *Inverse Problems* **31**:9 (2015), art. id. 095003. MR Zbl

[Harrach et al. 2017] B. Harrach, V. Pohjola, and M. Salo, "The monotonicity method for inverse scattering", *Oberwolfach Rep.* **24** (2017), 57–60.

[Harrach et al. 2018] B. Harrach, Y.-H. Lin, and H. Liu, "On localizing and concentrating electromagnetic fields", *SIAM J. Appl. Math.* **78**:5 (2018), 2558–2574. MR Zbl

[Harrach et al. 2019] B. Harrach, V. Pohjola, and M. Salo, "Dimension bounds in monotonicity methods for the Helmholtz equation", preprint, 2019. To appear in *SIAM J. Math. Anal.* arXiv

[Hörmander 1985] L. Hörmander, *The analysis of linear partial differential operators, III: Pseudodifferential operators*, Grundlehren der Math. Wissenschaften **274**, Springer, 1985. MR Zbl

[Ikehata 1998] M. Ikehata, "Size estimation of inclusion", *J. Inverse Ill-Posed Probl.* **6**:2 (1998), 127–140. MR Zbl

[Ikehata 1999] M. Ikehata, "Identification of the shape of the inclusion having essentially bounded conductivity", *J. Inverse Ill-Posed Probl.* **7**:6 (1999), 533–540. MR Zbl

[Imanuvilov et al. 2010] O. Y. Imanuvilov, G. Uhlmann, and M. Yamamoto, "The Calderón problem with partial data in two dimensions", *J. Amer. Math. Soc.* **23**:3 (2010), 655–691. MR Zbl

[Imanuvilov et al. 2015] O. Y. Imanuvilov, G. Uhlmann, and M. Yamamoto, "The Neumann-to-Dirichlet map in two dimensions", *Adv. Math.* **281** (2015), 578–593. MR Zbl

[Isakov 1988] V. Isakov, "On uniqueness of recovery of a discontinuous conductivity coefficient", *Comm. Pure Appl. Math.* **41**:7 (1988), 865–877. MR Zbl

[Isakov 2007] V. Isakov, "On uniqueness in the inverse conductivity problem with local data", *Inverse Probl. Imaging* **1**:1 (2007), 95–105. MR Zbl

[Kang et al. 1997] H. Kang, J. K. Seo, and D. Sheen, "The inverse conductivity problem with one measurement: stability and estimation of size", *SIAM J. Math. Anal.* **28**:6 (1997), 1389–1405. MR Zbl

[Kenig and Salo 2013] C. Kenig and M. Salo, "The Calderón problem with partial data on manifolds and applications", *Anal. PDE* **6**:8 (2013), 2003–2048. MR Zbl

[Kenig and Salo 2014] C. Kenig and M. Salo, "Recent progress in the Calderón problem with partial data", pp. 193–222 in *Inverse problems and applications* (Irvine, CA, 2012/Hangzhou, 2012), edited by P. Stefanov et al., Contemp. Math. **615**, Amer. Math. Soc., Providence, RI, 2014. MR Zbl

[Kenig et al. 2007] C. E. Kenig, J. Sjöstrand, and G. Uhlmann, "The Calderón problem with partial data", *Ann. of Math.* (2) **165**:2 (2007), 567–591. MR Zbl

[Kirsch 1998] A. Kirsch, "Characterization of the shape of a scattering obstacle using the spectral data of the far field operator", *Inverse Problems* **14**:6 (1998), 1489–1512. MR Zbl

[Kirsch and Grinberg 2008] A. Kirsch and N. Grinberg, *The factorization method for inverse problems*, Oxford Lecture Series Math. Appl. **36**, Oxford Univ. Press, 2008. MR Zbl

[Kohn and Vogelius 1984] R. Kohn and M. Vogelius, "Determining conductivity by boundary measurements", *Comm. Pure Appl. Math.* **37**:3 (1984), 289–298. MR Zbl

[Kohn and Vogelius 1985] R. V. Kohn and M. Vogelius, "Determining conductivity by boundary measurements, II: Interior results", *Comm. Pure Appl. Math.* **38**:5 (1985), 643–667. MR Zbl

[Krupchyk and Uhlmann 2016] K. Krupchyk and G. Uhlmann, "The Calderón problem with partial data for conductivities with $3/2$ derivatives", *Comm. Math. Phys.* **348**:1 (2016), 185–219. MR Zbl

[Kusiak and Sylvester 2003] S. Kusiak and J. Sylvester, "The scattering support", *Comm. Pure Appl. Math.* **56**:11 (2003), 1525–1548. MR Zbl

[Lax 2002] P. D. Lax, *Functional analysis*, Wiley, New York, 2002. MR Zbl

[Maffucci et al. 2016] A. Maffucci, A. Vento, S. Ventre, and A. Tamburrino, "A novel technique for evaluating the effective permittivity of inhomogeneous interconnects based on the monotonicity property", *IEEE Trans. Components Packaging Manuf. Tech.* **6**:9 (2016), 1417–1427.

[Nachman 1996] A. I. Nachman, "Global uniqueness for a two-dimensional inverse boundary value problem", *Ann. of Math.* (2) **143**:1 (1996), 71–96. MR Zbl

[Reed and Simon 1972] M. Reed and B. Simon, *Methods of modern mathematical physics, I: Functional analysis*, Academic Press, New York, 1972. MR Zbl

[Regbaoui 2001] R. Regbaoui, "Unique continuation from sets of positive measure", pp. 179–190 in *Carleman estimates and applications to uniqueness and control theory* (Cortona, Italy, 1999), edited by F. Colombini and C. Zuily, Progr. Nonlinear Differential Equations Appl. **46**, Birkhäuser, Boston, 2001. MR Zbl

[Su et al. 2017] Z. Su, L. Udpa, G. Giovinco, S. Ventre, and A. Tamburrino, "Monotonicity principle in pulsed eddy current testing and its application to defect sizing", pp. art. id. 48–02 in *Int. Applied Computational Electromagnetics Society Symposium* (Florence, 2017), IEEE, Piscataway, NJ, 2017.

[Sylvester and Uhlmann 1987] J. Sylvester and G. Uhlmann, "A global uniqueness theorem for an inverse boundary value problem", *Ann. of Math.* (2) **125**:1 (1987), 153–169. MR Zbl

[Tamburrino and Rubinacci 2002] A. Tamburrino and G. Rubinacci, "A new non-iterative inversion method for electrical resistance tomography", *Inverse Problems* **18**:6 (2002), 1809–1829. MR Zbl

[Tamburrino et al. 2016] A. Tamburrino, Z. Su, S. Ventre, L. Udpa, and S. S. Udpa, "Monotonicity based imaging method in time domain eddy current testing", pp. 1–8 in *Electromagnetic nondestructive evaluation, XIX* (Sendai, Japan, 2015), edited by N. Yusa et al., Studies Appl. Electromag. Mech. **41**, IOS Press, Amsterdam, 2016.

[Ventre et al. 2017] S. Ventre, A. Maffucci, F. Caire, N. Le Lostec, A. Perrotta, G. Rubinacci, B. Sartre, A. Vento, and A. Tamburrino, "Design of a real-time eddy current tomography system", *IEEE Trans. Magnetics* **53**:3 (2017), art. id. 6200408.

[Zhou et al. 2018] L. Zhou, B. Harrach, and J. K. Seo, "Monotonicity-based electrical impedance tomography for lung imaging", *Inverse Problems* **34**:4 (2018), art. id. 045005. MR Zbl

Received 27 Nov 2017. Revised 20 Jul 2018. Accepted 20 Nov 2018.

BASTIAN HARRACH: harrach@math.uni-frankfurt.de
Institute of Mathematics, Goethe University Frankfurt, Frankfurt, Germany

VALTER POHJOLA: valter.pohjola@gmail.com
Research Unit of Mathematical Sciences, University of Oulu, Oulu, Finland

MIKKO SALO: mikko.j.salo@jyu.fi
Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä, Finland

ANALYSIS AND PDE
Vol. 12, No. 7, 2019
dx.doi.org/10.2140/apde.2019.12.1773

SOLUTIONS OF THE 4-SPECIES QUADRATIC REACTION-DIFFUSION SYSTEM ARE BOUNDED AND C^∞-SMOOTH, IN ANY SPACE DIMENSION

M. CRISTINA CAPUTO, THIERRY GOUDON AND ALEXIS F. VASSEUR

We establish the boundedness of solutions of reaction-diffusion systems with quadratic (in fact slightly superquadratic) reaction terms that satisfy a natural entropy dissipation property, in any space dimension $N > 2$. This bound implies the C^∞-regularity of the solutions. This result extends the theory which was restricted to the two-dimensional case. The proof heavily uses De Giorgi's iteration scheme, which allows us to obtain local estimates. The arguments rely on duality reasoning in order to obtain new estimates on the total mass of the system, both in the $L^{(N+1)/N}$ norm and in a suitable weak norm. The latter uses C^α regularization properties for parabolic equations.

1. Introduction

This paper is mainly concerned with the system of reaction-diffusion equations

$$\begin{gathered} \partial_t a_i - \nabla\cdot(D_i\nabla a_i) = Q_i(a), \quad i\in\{1,2,3,4\},\ t\ge 0,\ x\in\mathbb{R}^N, \\ Q_i(a) = (-1)^{i+1}(a_2a_4 - a_1a_3), \end{gathered} \tag{1}$$

with initial condition

$$a|_{t=0} = a^0 = (a_1^0, a_2^0, a_3^0, a_4^0). \tag{2}$$

This system arises in chemistry where four species interact according to the reactions

$$A_1 + A_3 \rightleftharpoons A_2 + A_4,$$

the unknowns $(t,x)\mapsto a_i(t,x)$ in (1) being the local mass concentrations of the species labeled by $i\in\{1,2,3,4\}$: $\int_{\mathbb{R}^N} a_i(t,x)\,\mathrm{d}x$ is interpreted as the mass of the constituent i at time t. It is thus physically relevant to consider initial data a_i^0 which are nonnegative integrable functions. The reactants are subjected to space diffusion and the diffusion coefficients depend on the considered species. In full generality, D_i can be a function of the space variable with values in the space of $N\times N$ matrices. Throughout this paper, we restrict to the case of scalar and constant matrices

$$D_i(x) = d_i\,\mathbb{I}, \quad d_i > 0 \text{ constant},$$

with coefficients that satisfy

$$0 < \delta_\star \le d_i \le \delta^\star. \tag{3}$$

Assuming that the initial data are smooth, say $a_i^0 \in C^\infty(\mathbb{R}^N)$, existence-uniqueness of smooth and nonnegative solutions for (1)–(2) can be justified at least on a small time interval, by using a standard

MSC2010: 35K45, 35B65, 35K57.
Keywords: reaction-diffusion systems, global regularity, blow-up methods.

fixed-point reasoning; see for instance [Goudon and Vasseur 2010, Proposition A.2] or [Pierre 2010, Lemma 1.1]. Global existence of weak solutions is established in [Desvillettes et al. 2007]. We address the question of the boundedness of the solutions, which will imply that solutions are globally defined and remain infinitely smooth [Goudon and Vasseur 2010, Proposition A.1].

The difficulty comes from the fact we are dealing with different diffusion coefficients. As already noticed in [Goudon and Vasseur 2010], the question becomes trivial when all the D_i's vanish: in this case, we are concerned with a mere system of ODEs which clearly satisfies a maximum principle. The answer is also immediate when all the diffusion coefficients are equal to the same constant $d_i = \delta_\star$. Indeed, in this situation, the total mass

$$M(t,x) = \sum_{i=1}^{4} a_i(t,x)$$

satisfies the heat equation $\partial_t M = \delta_\star \Delta M$, which, again, easily leads to a maximum principle. In the general situation, one may wonder whether or not the system has the explosive behavior of nonlinear heat equations [Weissler 1985]. Counterexamples of systems with polynomial nonlinearities presented in [Pierre and Schmitt 1997] show that this question is relevant and nontrivial; see also [Pierre 2010, Theorem 4.1]. We refer the reader to the survey [Pierre 2010] for a general presentation of the problem, further references, and many deep comments on the mathematical difficulties raised by such systems.

Two properties are crucial for the analysis of the problem. First of all, system (1) conserves mass:

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{i=1}^{4} \int_{\mathbb{R}^N} a_i \,\mathrm{d}x = 0. \tag{4}$$

Second of all, it dissipates entropy:

$$\sum_{i=1}^{4} Q_i(a) \ln(a_i) = -(a_2a_4 - a_1a_3) \ln\left(\frac{a_2a_4}{a_1a_3}\right) \le 0. \tag{5}$$

These properties suggest to consider more general systems, involving more reactants and possibly more intricate nonlinearities. To be more specific, we extend the discussion to systems that read

$$\begin{gathered} \partial_t a_i - \nabla\cdot(D_i \nabla a_i) = Q_i(a), \quad i \in \{1,\dots,p\},\ t \ge 0,\ x \in \mathbb{R}^N, \\ Q_i : a \in \mathbb{R}^p \longmapsto \mathbb{R}^p, \end{gathered} \tag{6}$$

endowed with the initial condition

$$a|_{t=0} = a^0 = (a_1^0, \dots, a_p^0), \tag{7}$$

where the reaction term fulfills the following conditions:

(h1) There exists $\mathscr{Q} > 0$ and $q > 0$ such that for any $a \in \mathbb{R}^p$ and $i \in \{1,\dots,p\}$, we have $|\nabla_a Q_i(a)| \le \mathscr{Q}|a|^{q-1}$.

(h2) For any $i \in \{1,\dots,p\}$, if $a_i \le 0$ then $Q_i(a) \le 0$.

(h3) $\sum_{i=1}^{p} Q_i(a) = 0$.

(h4) $\sum_{i=1}^{p} Q_i(a) \ln(a_i) \le 0$.

Assumption (h1) governs the growth of the nonlinearity. In what follows, we will be concerned with quadratic and superquadratic growth: $q \geq 2$ (but q is not necessarily assumed to be an integer). Assumption (h2) relies on the preservation of nonnegativity of the solutions, and it is thus physically relevant. Assumptions (h3) and (h4) imply mass conservation and entropy dissipation, respectively. Note that the entropy dissipation actually provides an estimate on (nonlinear) derivatives of the unknown since it leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{p}\int_{\mathbb{R}^N} a_i \ln(a_i)\,\mathrm{d}x + 4\delta_\star \sum_{i=1}^{p}\int_{\mathbb{R}^N} |\nabla\sqrt{a_i}|^2\,\mathrm{d}x \leq 0. \tag{8}$$

In view of (h3) and (h4), it is thus natural to consider initial data such that

$$\begin{gathered} a_i^0 : x \in \mathbb{R}^N \longmapsto a_i^0(x) \geq 0, \\ \sup_{i\in\{1,\dots,p\}} \int_{\mathbb{R}^N} a_i^0(1+\ln(a_i^0)+|x|)\,\mathrm{d}x = \mathscr{M}^0 < \infty. \end{gathered} \tag{9}$$

We refer the reader to Proposition 2.1 below for a more precise statement in terms of a priori estimates. It means that the initial concentrations have finite mass and entropy. The moment condition controls the spreading of the mass. However, while (8) has a clear physical meaning, it does not provide enough estimates for the analysis of the problem: note that with $u, u\ln(u) \in L^1$ and $\nabla\sqrt{u} \in L^2$, it is still not clear how the nonlinear term $Q(u)$ can make sense in $\mathscr{D}'$! For this reason, a notion of *renormalized* solutions is introduced in [Fischer 2015], and existence of solutions in this framework can be established. Entropy dissipation plays also a central role in the analysis of the asymptotic trend towards equilibrium [Desvillettes and Fellner 2006; Fellner et al. 2016; Pierre et al. 2017a; 2017b].

In the specific quadratic and two-dimensional case ($q = 2$, $N = 2$) the question is fully answered in [Goudon and Vasseur 2010]: starting from $L^\infty \cap C^\infty$ initial data, the solution remains bounded and smooth and the problem is globally well-posed. In fact [loc. cit.] proves a *regularizing* effect: with data satisfying (9) only, the solution becomes *instantaneously* bounded and smooth, which implies global well-posedness. The proof in [loc. cit.] relies on De Giorgi's approach [1957]; it uses entropy dissipation, see (8), to get a nonlinear control on level sets of the solution, which eventually leads to the L^∞ bound. The result is extended for higher space dimensions in [Cañizo et al. 2014], which handles, with different techniques, the quadratic case when the diffusion coefficients are close enough to the same constant (how small the distance between the d_j's should be depends on the space dimension, in a explicit way; see also [Fellner et al. 2016; Pierre et al. 2017a]), and in [Caputo and Vasseur 2009], which handles subquadratic nonlinearities ($q < 2$ in (h1), not necessarily integer). Two ingredients are crucial in the approach of [loc. cit.]:

- First, [loc. cit.] uses systematically rescaled quantities

$$a_i^{(\epsilon)}(s, y) = \epsilon^{2/(q-1)} a_i(t + \epsilon^2 s, x + \epsilon y), \tag{10}$$

with $\epsilon > 0$: $a^{(\epsilon)}$ satisfies the same evolution equation as a. Note that in the quadratic case ($q = 2$), for $N = 2$, the rescaling leaves invariant the natural norms of the problem $\|a\|_{L^\infty(0,\infty;L^1(\mathbb{R}^2))}$ and $\|\nabla\sqrt{a}\|_{L^2((0,\infty)\times\mathbb{R}^2)}$.

• Second, the parabolic regularity is obtained by adapting De Giorgi's techniques, and by working with a certain norm of the rescaled unknown which becomes small as $\epsilon \to 0$. It turns out that the necessary estimate holds in a weak sense. Namely, one has to consider the set of distributions

$$T \in \mathscr{D}'((0,T)\times\mathbb{R}^N) \text{ such that } T = \Delta\Phi, \text{ with } \Phi \in L^\infty((0,\infty)\times\mathbb{R}^N).$$

The corresponding rescaled norm behaves like $\mathscr{O}(\epsilon^{(4-2q)/(q-1)})$, which indeed tends to 0 as $\epsilon \to 0$ for subquadratic nonlinearities $q < 2$. The idea of using such a weak norm also appeared in the regularity analysis for the Navier–Stokes equation [Vasseur 2010]. We also refer the reader to [Caffarelli and Vasseur 2010; Vasseur 2007] for further applications of De Giorgi's techniques to the analysis of fluid mechanics systems and to [Alonso et al. 2016; Goudon and Urrutia 2016] for the study of models for populations dynamics governed by "chemotactic-like" mechanisms. This approach is also useful for the analysis of the preservation of bounds by numerical schemes when solving nonlinear convection-diffusion systems [Chainais-Hillairet et al. 2017]. In the reasoning adopted in [Caputo and Vasseur 2009], a special role is played by the total mass $M = \sum_{i=1}^p a_i$, which satisfies the diffusion equation

$$\partial_t M - \Delta(dM) = 0, \quad d(t,x) = \frac{\sum_{i=1}^p d_i a_i(t,x)}{\sum_{i=1}^p a_i(t,x)}, \tag{11}$$

where, by virtue of (3), the diffusion coefficient d satisfies

$$0 < \delta_\star \le d(t,x) \le \delta^\star.$$

This relation can be used to establish, through an elegant duality argument, an estimate in $L^2((0,T)\times\mathbb{R}^N)$; see [Pierre and Schmitt 1997; Desvillettes et al. 2007]. This estimate is a key for proving the global existence of weak solutions for the quadratic problem (1)–(2) in [Desvillettes et al. 2007]: at least, it is worth pointing out that with this L^2 estimate the right-hand side $Q_i(a)$ in (1) makes sense, while the estimates based on the mass conservation and entropy dissipation were not enough. However, the L^2 estimate does not shrink the rescaled solutions $a^{(\epsilon)}$ as $\epsilon \to 0$ and it is thus not enough to provide global boundedness and regularity. This is where we can take advantage of using a weak norm.

In the present work, we wish to fill the gap in the boundedness theory and to provide a complete answer for the quadratic case in *any* dimension. In fact, our analysis also covers higher nonlinearities, but with an implicit condition on the growth condition. Our main results can be stated as follows.

Theorem 1.1. *Let $N \in \mathbb{N}$, with $N \ge 3$. For any initial data $a^0 = (a_1^0, a_2^0, a_3^0, a_4^0)$ in $(C^\infty(\mathbb{R}^N) \cap L^\infty(\mathbb{R}^N))^4$ such that $a_i(x) \ge 0$ for any $x \in \mathbb{R}^n$ and $i \in \{1,\dots,4\}$, there exists a unique, globally defined, solution $a = (a_1, a_2, a_3, a_4)$ to* (1)–(2) *which is nonnegative, bounded on* $[0,T]\times\mathbb{R}^N$ *for any* $0 < T < \infty$, *and C^∞-smooth.*

Theorem 1.2. *Let $N \in \mathbb{N}$, with $N \ge 3$. Consider a system* (6) *satisfying* (h1)–(h4). *There exists $\nu_0 > 0$ depending on N, $\delta_\star$ and $\delta^\star$ such that if* (h1) *holds with $2 \le q \le 2 + \nu_0 \le 2(N+1)/N$, then for any nonnegative $a^0 \in C^\infty(\mathbb{R}^N;\mathbb{R}^p) \cap L^\infty(\mathbb{R}^N;\mathbb{R}^p)$, there exists a unique, globally defined, solution a to* (6)–(7) *which is nonnegative, bounded on* $[0,T]\times\mathbb{R}^N$ *for any* $0 < T < \infty$, *and C^∞-smooth.*

Theorem 1.1 thus appears as a consequence of Theorem 1.2. The extra power ν_0 allowed on the nonlinearities depends on N, $\delta_\star$ and $\delta^\star$ in an implicit way and our method does not provide any precise estimate. It seems unlikely that it can correspond to a physically relevant threshold. The problem of regularity remains open for higher nonlinearities. The proof still follows the De Giorgi strategy, and relies on a refinement of the weak norm estimate obtained in [Caputo and Vasseur 2009] (which, though, remains a crucial ingredient of the proof). To be more specific, we are going to upgrade the L^∞ estimate to a C^α estimate, working with the set of distributions

$$T \in \mathscr{D}'((0,T)\times\mathbb{R}^N) \text{ such that } T = \Delta\Phi, \text{ with } \Phi \in L^\infty(0,\infty;C^\alpha(\mathbb{R}^N)),$$

for a certain regularity coefficient $0 < \alpha \le 1$. This is combined with a $L^{(N+1)/N}$ estimate on the total mass, obtained through a duality argument. This argument is directly inspired by the derivation of elliptic estimates in [Fabes and Stroock 1984] and it appears as a dual version of the Alexandrof–Bakelman–Pucci–Krylov–Tso (ABPKT) estimate [Alexandrof 1966; Bakelman 1961; Pucci 1966; Krylov 1976; Tso 1985]. We point out that, contrarily to the approach in [Caputo and Vasseur 2009], we do not use here the bounds derived from the entropy dissipation (8).

The paper is organized as follows. In Section 2, we give an overview of the main steps of the proof. Section 3 is concerned with the weak estimate on the total mass. It relies on a Hölderian regularity analysis for parabolic equations. This is combined with a duality argument which uses crucially the nonnegativity of the solution. Section 4 is devoted to a complementary estimate in a suitable Lebesgue space, which, again, relies on a duality approach. Section 5 explains how the arguments combine to end the proof of the main results. The paper is completed by a quite long appendix which details how the De Giorgi machinery arises in the justification of the intermediate steps of the proof.

Remark 1.3. The key estimates of the proof involve the bound from below $0 < \delta_\star \le d(t,x)$; hence the result cannot be extended to cases with degenerate diffusion coefficients. That the total mass M satisfies the diffusion equation (11) uses crucially the fact that the diffusion coefficients are given by *scalar* matrices. It also uses the assumption that the coefficients d_i are constant; otherwise an additional convection term $\nabla\cdot(uM)$ arises with the velocity field $(t,x)\mapsto u(t,x)$ having components

$$u_j(t,x) = \frac{\sum_{i=1}^p a_i(t,x)\,\partial_{x_j} d_i(x)}{\sum_{i=1}^p a_i(t,x)}.$$

It is likely that, up to suitable technical requirements on the d_i's, the analysis could cover such a situation as well. Our analysis can be adapted to handle problems in bounded domains with Neumann boundary conditions; the situation of Dirichlet conditions is more subtle since there are difficulties to obtain useful estimates up to the boundaries [Pierre et al. 2017a].

2. Main steps of the proof

***A priori estimates*: *boundedness*, *global existence and regularity of the solutions*.** In what follows, we are going to establish several a priori estimates satisfied by the solutions of (6). To this end, we will perform various manipulations, such as integrations by parts, permutations of integrals and derivation, etc. These manipulations apply to the smooth solutions of the problem that can be shown to exist on a small enough time interval; see [Goudon and Vasseur 2010, Proposition A.2]. They equally apply to solutions of suitable approximations of the problem (6). The construction of such an approximation (by regularizing data, coefficients, cutting-off the nonlinearities…) can be a delicate issue in order to preserve the structural features of the original equation, and to admit a globally defined smooth solution. We refer the reader on this issue to [Desvillettes et al. 2007]. As it will be clear in the forthcoming discussion, the estimates we are going to derive do not depend on the regularization parameter, but only on N, $\delta_\star$, $\delta^\star$, and $\mathscr{Q}$, p, q (see (h1)), which, eventually, allows us to conclude by getting rid of the regularization parameter. The very first estimate is a direct consequence of the mass conservation and entropy dissipation properties of the system. The following claim, see [Goudon and Vasseur 2010, Proposition 2.1], applies without any restriction on the number of species p, the degree of nonlinearity q nor on the space dimension N.

Proposition 2.1 [Goudon and Vasseur 2010]. *Assume* (h1)–(h4). *Let* $a_0 = (a_1^0, \dots, a_p^0)$, *with nonnegative components*, *satisfy* (9). *Then*, *for any* $0 < T < \infty$, *there exists* $0 < C(T) < \infty$ *such that*

$$\sup_{0\le t\le T} \left\{ \sum_{i=1}^{p} \int_{\mathbb{R}^N} a_i (1 + |x| + |\ln(a_i)|)(t,x)\,\mathrm{d}x \right\} + \sum_{i=1}^{p} \int_0^T \int_{\mathbb{R}^N} |\nabla \sqrt{a_i}|^2(s,x)\,\mathrm{d}x\,\mathrm{d}s + \sum_{i=1}^{p} \int_0^T \int_{\mathbb{R}^N} Q_i(a)\ln(a_i)\,\mathrm{d}x\,\mathrm{d}s \le C(T).$$

The entropy dissipation (8) tells us that $\sum_{i=1}^{p} \int_{\mathbb{R}^N} a_i \ln(a_i)(t,x)\,\mathrm{d}x$ is a nonincreasing function of the time variable. However, this quantity has no sign. To make this information a useful estimate, involving the nonnegative quantities $a_i|\ln(a_i)|$ we need a control on the first-order space moments $\int_{\mathbb{R}^N} |x| a_i(t,x)\,\mathrm{d}x$. We refer the reader to [Goudon and Vasseur 2010] for details. This estimate will not be used in our reasoning; nevertheless the entropy dissipation still has a crucial role in the proof of Theorems 1.1 and 1.2. By the way note that the counterexamples of systems that produce blow up in [Pierre and Schmitt 1997] very likely do not satisfy the entropy dissipation property.

As said above, for data in $C^\infty \cap L^\infty(\mathbb{R}^N)$, we can construct a C^∞ and bounded solution defined on a small enough interval. Let $T_{\max}$ be the lifespan of such a solution. Standard bootstrapping arguments tell us that if $T_{\max} < \infty$ then we have

$$\limsup_{t\to T_{\max}} \|a(t,\cdot)\|_{L^\infty(\mathbb{R}^N)} = +\infty.$$

In what follows, we are going to obtain a uniform bound satisfied by $\|a(t,\cdot)\|_{L^\infty(\mathbb{R}^N)}$ on the time interval $[0, T_{\max})$, depending only on $T_{\max}$ and the assumptions on the data, which thus contradicts the occurrence of a blow-up of the solution in finite time. Therefore, the L^∞ estimate implies that the lifespan of the solutions of (6)–(7) is infinite. Moreover, boundedness also implies the regularity of the solution, by a bootstrap argument; see [Goudon and Vasseur 2010, Proposition A.1].

The key intermediate statements. The main ingredient consists in showing that the local boundedness can be obtained from a local estimate in L^r, with $r > 1$; see [Caputo and Vasseur 2009, Proposition 4].

We thus work on balls

$$B_\rho = \{x \in \mathbb{R}^N : |x| \le \rho\}.$$

Lemma 2.2 (De Giorgi-type lemma, [Caputo and Vasseur 2009]). *We suppose that* $2 \le q < 2(N+1)/N$. *We also suppose that* (h1)–(h4) *holds. Let a be a nonnegative solution to* (6) *on* $(-1,0)\times B_1$. *Then, for any* $r > 1$, *there exists a universal constant* $\delta_r > 0$ *such that, if* $a = (a_1, \cdots, a_p)$ *satisfies*

$$\sum_{i=1}^{p} \|a_i\|_{L^r((-1,0)\times B_1)} \le \delta_r,$$

then $0 \le a_i(0,0) \le 1$ *for* $i \in \{1,\dots,p\}$.

The proof relies on De Giorgi's techniques [1957]; see also [Alikakos 1979] for an alternative approach. For the sake of completeness we describe the main steps in Appendix B; it is also important to detail this proof since this is where the entropy dissipation plays a central role. At first sight this information does not look very useful since the natural estimates for (6)–(7) in Proposition 2.1 do not involve L^r norms for an exponent r larger than 1. However, we will be able to identify further estimates, which shrink for the rescaled solutions (10) as $\epsilon \to 0$. Namely, we will find that $r = (N+1)/N$ plays a specific role since the rescaled total mass satisfies $\lim_{\epsilon\to 0} \|M^{(\epsilon)}\|_{L^{(N+1)/N}((-1,0)\times B_1)} = 0$. Thus, for ϵ small enough the rescaled solution fulfills the criterion in Lemma 2.2.

Lemma 2.3. *There exists* $\epsilon_0 > 0$ *and* $\nu_0 > 0$ *depending on* N, $\delta_\star$ *and* $\delta^\star$ *such that if* (h1) *holds with* $2 \le q \le 2+\nu_0 \le 2(N+1)/N$, *then for all* $0 < \epsilon \le \epsilon_0$ *we have*

$$\sum_{i=1}^{p} \|a_i^{(\epsilon)}\|_{L^{(N+1)/N}((-1,0)\times B_1)} \le \delta,$$

with $\delta = \delta_{(N+1)/N}$ *as defined in Lemma 2.2.*

Coming back to the original variables, we obtain the L^∞ estimate.

Corollary 2.4. *Let* ϵ_0 *be defined in Lemma 2.3. Then, for all* $T_{\max}/2 < t < T_{\max}$, *we have*

$$\sum_{i=1}^{p} \|a_i(t,\cdot)\|_{L^\infty(\mathbb{R}^N)} \le \epsilon_0^{-2/(q-1)}.$$

Proof. Pick x_0 in $\mathbb{R}^N$ and $t_0 \in (T_{\max}/2, T_{\max})$. Applying Lemma 2.2 to $a^{(\epsilon_0)}$ yields

$$0 \le \sum_{i=1}^{p} a_i(t_0,x_0) = \epsilon_0^{-2/(q-1)} \sum_{i=1}^{p} a_i^{(\epsilon_0)}(0,0) \le \epsilon_0^{-2/(q-1)}. \qquad \square$$

Having this statement at hand allows us to conclude the proof of Theorem 1.2. Let $2 \le q \le 2+\nu_0 \le 2(N+1)/N$. Let $a = (a_1,\dots,a_p)$ be a solution to (6)–(7), and let $T_{\max}$ be the lifespan of a. Assume that $T_{\max}$ is finite. Then, for each $i \in \{1,\dots,p\}$, Corollary 2.4 tells us that $a_i(t,\cdot)$ is uniformly bounded for all $T_{\max}/2 < t < T_{\max}$ and thus the sup norm does not blow up as $t \to T_{\max}$. This contradicts the fact that $T_{\max}$ is the maximal time of existence of a smooth solution of (6)–(7). $\square$

Therefore the cornerstone of the proof consists in proving Lemma 2.3 and identifying the specific role played by the norm $L^{(N+1)/N}$. The argument is two-fold and it uses the diffusion equation (11) satisfied by the total mass $M(t,x)=\sum_{i=1}^{p} a_i(t,x)$. On the one hand, we shall show that the norm $L^{(N+1)/N}$ of M can be controlled by means of the norm $L^\infty(0,\infty;L^1(\mathbb{R}^N))$. On the other hand, we shall obtain a new estimate on a *weak norm* of M, which will allow us to conclude that

$$\lim_{\epsilon\to 0}\|M^{(\epsilon)}\|_{L^\infty(0,\infty;L^1(\mathbb{R}^N))}=0,\quad \text{with } M^{(\epsilon)}(s,y)=\epsilon^{2/(q-1)}M(t+\epsilon^2 s,x+\epsilon y).$$

This analysis is based on duality arguments and regularization properties of parabolic equations. Accordingly, we can conclude to the shrinking as $\epsilon\to 0$ of the $L^{(N+1)/N}$ norm of the rescaled solutions.

3. Weak norm estimates on the total mass and shrinking of the rescaled total mass

Our approach relies on the following statement.

Proposition 3.1. *Let $\Phi:(0,T)\times\mathbb{R}^N\to\mathbb{R}$ such that*

(a) *Φ lies in $L^\infty((0,T)\times\mathbb{R}^N)$;*

(b) *$\Delta\Phi=M\geq 0$;*

(c) *Φ satisfies $\partial_t\Phi-d\Delta\Phi=0$ on $(0,T)\times\mathbb{R}^N$, with a coefficient $d:(0,T)\times\mathbb{R}^N\to\mathbb{R}$ satisfying $0<\delta_\star\leq d(t,x)\leq\delta^\star<\infty$ for a.e. $(t,x)\in(0,T)\times\mathbb{R}^N$.*

Then, there exists $\alpha\in(0,1]$ such that $\Phi\in C^{[\alpha/2,\alpha]}([t_0,T]\times\mathbb{R}^N)$ for any $t_0>0$, which means that we can find $C>0$ such that, for any $(t,x)\in[t_0,T]\times\mathbb{R}^N$ and $(\tau,h)\in\mathbb{R}\times\mathbb{R}^N$ with $t+\tau\geq t_0$, we have

$$\frac{|\Phi(t+\tau,x+h)-\Phi(t,x)|}{|\tau|^{\alpha/2}+|h|^\alpha}\leq C\|\Phi\|_{L^\infty}.$$

This Hölder regularity estimate for nonconservative parabolic equations dates back to [Krylov and Safonov 1979; 1980]. In fact, the result of those papers does not need the sign property (b). However, as it will be explained below, this sign property naturally appears for the system under consideration, and it plays a further crucial role throughout the analysis. Let us explain the interest of this statement for our purpose. As said above, the total mass M satisfies the diffusion equation (11). Of course, by definition, M is a nonnegative function which lies in $L^\infty(0,\infty;L^1(\mathbb{R}^N))$. Let Φ satisfy $\Delta\Phi=M\geq 0$. Since $d(t,x)$ is bounded above by $\delta^\star$, Φ also satisfies the evolution equation

$$\partial_t\Phi-\delta^\star\Delta\Phi=(d-\delta^\star)\Delta\Phi=(d-\delta^\star)M\leq 0.$$

This observation is the cornerstone of the analysis performed in [Caputo and Vasseur 2009]. In particular, we will make use of the following crucial property established in Proposition 11 and Corollary 12 of that paper.

Proposition 3.2. *Let $N\in\mathbb{N}$, with $N\geq 3$. Let $\Phi=\Delta^{-1}M$ with M the total mass associated to a solution of* (6). *Then, we have*

$$\|\Phi\|_{L^\infty((0,T)\times\mathbb{R}^N)}\leq\|\Phi(0,\cdot)\|_{L^\infty(\mathbb{R}^N)}\leq K_N\,\|M(0,\cdot)\|_{L^\infty(\mathbb{R}^N)}^{1-2/N}\,\|M(0,\cdot)\|_{L^1(\mathbb{R}^N)}^{2/N},$$

where $K_N>0$ is a certain universal constant, which only depends on the space dimension.

Proposition 3.1 thus strengthens the results of [Caputo and Vasseur 2009] in the sense that it provides, beyond the L^∞ estimate on Φ, a Hölder-regularity estimate. Since the estimate in Proposition 3.2 is not evident at first sight, we give the main steps of the proof in Appendix C for the sake of completeness. We shall use the following consequence of Proposition 3.1, which is precisely the estimate that allows us to go beyond the subquadratic nonlinearities dealt with in [loc. cit.].

Lemma 3.3. *Let M be a nonnegative solution of* (11), *and let $\Phi = \Delta^{-1}M$. Let $t \geq t_0 > 0$ and $x \in \mathbb{R}^N$. For $\epsilon > 0$, we set $M^{(\epsilon)}(s,y) = \epsilon^{2/(q-1)}M(t+\epsilon^2 s, x+\epsilon y)$. We suppose that $M^{(\epsilon)}$ lies in $L^\infty(-4,0;L^1(\mathbb{R}^N))$. Then, there exists $c>0$ and $0<\alpha\leq 1$* (*provided by Proposition 3.1*), *depending only on N, $\delta_\star$ and $\delta^\star$, such that for any $0<\epsilon\leq\sqrt{t_0}/2$,*

$$\sup_{-4\leq s\leq 0}\int_{B_2} M^{(\epsilon)}(s,y)\,\mathrm{d}y \leq c\|\Phi\|_{L^\infty}\,\epsilon^{\alpha-2+2/(q-1)}.$$

Proof. Let $\zeta \in C_c^\infty(\mathbb{R}^N)$ be such that $\mathrm{supp}(\zeta)\subset B_2$ and $\zeta(x)=1$ for any $x\in B_1$. Since $M^{(\epsilon)}\geq 0$, we get

$$\begin{aligned}\int_{B_1} M^{(\epsilon)}(s,y)\,\mathrm{d}y &\leq \int_{B_2}\zeta M^{(\epsilon)}(s,y)\,\mathrm{d}y = \int_{B_2}\zeta\Delta\Phi^{(\epsilon)}(s,y)\,\mathrm{d}y\\ &\leq \int_{B_2}\Delta\zeta(y)(\Phi^{(\epsilon)}(s,y)-\Phi^{(\epsilon)}(0,0))\,\mathrm{d}y.\end{aligned}$$

By virtue of Proposition 3.1, we can write

$$\begin{aligned}\int_{B_1} M^{(\epsilon)}(s,y)\,\mathrm{d}y &\leq \epsilon^{-2+2/(q-1)}\int_{B_2}\Delta\zeta(y)(\Phi(t+\epsilon^2 s, x+\epsilon y)-\Phi(t,x))\,\mathrm{d}y\\ &\leq C\|\zeta\|_{W^{2,\infty}(\mathbb{R}^N)}\|\Phi\|_{L^\infty}\epsilon^{\alpha-2+2/(q-1)}\end{aligned}$$

for any $s\in(-4,0)$ and $0<\epsilon^2<t_0/4$. In the first inequality, the exponent $2/(q-1)$ comes from the rescaling that defines $M^{(\epsilon)}$, and the exponent -2 comes from the relation

$$M^{(\epsilon)}(s,y) = \frac{1}{\epsilon^2}\Delta(\Phi^{(\epsilon)}(t+\epsilon^2 s, x+\epsilon y)). \qquad \square$$

The information in Lemma 3.3 is relevant when the exponent $\alpha-2+2/(q-1)$ is positive. This implies a restriction on $q\leq 1+2/(2-\alpha)$, where we remind the reader that $\alpha\in(0,1]$ depends on $N,\delta_\star,\delta^\star$ and we note that the bound from above increases to 3 as $\alpha\to 1$. This combines with the constraint $q\leq 2(N+1)/N$, which is of different nature; see Lemma 2.2, Lemma 2.3 and Lemma B.2.

As indicated above, the Hölder estimate in Proposition 3.1 is a standard result due to [Krylov and Safonov 1979; 1980]. For the sake of completeness, we provide in Appendix D an alternative proof, fully based on De Giorgi's arguments. Note however that this analysis uses the additional assumption (b), which appears naturally in the problem under consideration.

4. $L^{(N+1)/N}$ estimate on the total mass

This section is devoted to the proof of the following statement.

Proposition 4.1. *There exists a constant $K > 0$ (depending on N, $\delta_\star$, $\delta^\star$) such that, $M \geq 0$ being a solution of* (11) *in $Q_2 = (-4,0) \times B_2$ and defining $Q_1 = (-1,0) \times B_1$, we have*

$$\|M\|_{L^{(N+1)/N}(Q_1)} \leq K \sup_{-4 \leq t \leq 0} \int_{B_2} M(t,x)\,\mathrm{d}x.$$

Proof. Let f be in $C_c^\infty(Q_1)$ such that

$$\|f\|_{L^{N+1}(Q_1)} \leq 1.$$

We consider the solution of the end-value problem

$$\begin{aligned} \partial_t u + d\Delta u &= f \quad \text{in } (0,T) \times \mathbb{R}^N, \\ u(T,x) &= 0, \quad u|_{\partial B_2} = 0. \end{aligned} \tag{12}$$

We start by reminding the reader of the Alexandrof–Bakelman–Pucci–Krylov–Tso (ABPKT) inequality [Alexandrof 1966; Bakelman 1961; Pucci 1966; Krylov 1976; Tso 1985]: there exists a constant $\mathscr{C} > 0$ such that

$$\sup_{(t,x)\in Q_2} |u(t,x)| \leq \mathscr{C}\, \|f\|_{L^{N+1}(Q_2)}. \tag{13}$$

In order to obtain an estimate on the $L^{(N+1)/N}(Q_1)$ norm of M, a solution of (11), we proceed by duality, bearing in mind the definition

$$\|M\|_{L^{(N+1)/N}(Q_1)} = \sup\left\{ \left| \iint_{Q_1} Mf\,\mathrm{d}x\,\mathrm{d}t \right| : f \in C_c^\infty(Q_1),\ \|f\|_{L^{N+1}(Q_1)} \leq 1 \right\}.$$

Let ζ be a cut-off function: $\zeta \in C_c^\infty(B_{3/2})$, $\zeta(x) = 1$ for any $x \in B_1$, and $0 \leq \zeta(x) \leq 1$ for any $x \in \mathbb{R}^N$. Note that

$$\iint_{Q_2} \zeta Mf\,\mathrm{d}x\,\mathrm{d}t = \iint_{Q_1} Mf\,\mathrm{d}x\,\mathrm{d}t,$$

since $\operatorname{supp}(f) \subset Q_1$. We compute this integral by using (12):

$$\begin{aligned} \iint_{Q_2} \zeta Mf\,\mathrm{d}x\,\mathrm{d}t &= \iint_{Q_2} \zeta M(\partial_t u + d\Delta u)\,\mathrm{d}x\,\mathrm{d}t \\ &= \int_{-2}^{0} \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_{B_2} \zeta Mu\,\mathrm{d}x \right) \mathrm{d}t - \iint_{Q_2} \zeta u \Delta(dM)\,\mathrm{d}x\,\mathrm{d}t + \iint_{Q_2} \zeta M d\Delta u\,\mathrm{d}x\,\mathrm{d}t \\ &= \int_{B_2} \zeta Mu(0,x)\,\mathrm{d}x - 2\iint_{Q_2} dM \nabla\zeta \cdot \nabla u\,\mathrm{d}x\,\mathrm{d}t - \iint_{Q_2} udM\Delta\zeta\,\mathrm{d}x\,\mathrm{d}t. \end{aligned}$$

We have used several integrations by parts where the boundary terms vanish owing to the fact that $\operatorname{supp}(\zeta) \subset B_{3/2} \subset B_2$. The integrand of the penultimate term in the right-hand side can be rewritten as $\sqrt{dM}\nabla u \cdot \sqrt{dM}\nabla\zeta$, and then we use the Cauchy–Schwarz inequality and the Young inequality

$$ab = \sqrt{\kappa}a \frac{b}{\sqrt{\kappa}} \leq \frac{1}{2}\left(\kappa a^2 + \frac{b^2}{\kappa} \right).$$

We thus arrive at the following estimate:

$$\left|\iint_{Q_2} \zeta f M \,\mathrm{d}x\,\mathrm{d}t\right| \le \left|\int_{B_2} \zeta M u(0,x)\,\mathrm{d}x\right| + \kappa \iint_{Q_2} dM|\nabla u|^2\,\mathrm{d}x\,\mathrm{d}t \\ + \frac{1}{\kappa}\iint_{Q_2} dM|\nabla\zeta|^2\,\mathrm{d}x\,\mathrm{d}t + \left|\iint_{Q_2} udM\Delta\zeta\,\mathrm{d}x\,\mathrm{d}t\right|, \quad (14)$$

where $\kappa \in (0,1)$ is a parameter that will be determined later on. Inspired from [Fabes and Stroock 1984, proof of Theorem 2.1], in order to estimate the second integral in the right-hand side, we use the elementary relation

$$|\nabla u|^2 = \tfrac{1}{2}\Delta(u^2) - u\Delta u.$$

Going back to (12), we are thus led to

$$d|\nabla u|^2 = \tfrac{1}{2}d\Delta(u^2) + \tfrac{1}{2}\partial_t(u^2) - uf.$$

The advantage of this formulation relies on the fact that, denoting by ν the outward unit normal on ∂B_2,

$$u|_{\partial B_2} = u^2|_{\partial B_2} = 0, \quad \nabla u^2\cdot\nu|_{\partial B_2} = 2u\nabla u\cdot\nu|_{\partial B_2} = 0,$$

which allows us to perform further integration by parts. We get

$$\begin{aligned}\iint_{Q_2} & dM|\nabla u|^2\,\mathrm{d}x\,\mathrm{d}t \\ &= \frac{1}{2}\iint_{Q_2} dM\Delta(u^2)\,\mathrm{d}x\,\mathrm{d}t + \frac{1}{2}\iint_{Q_2} M\partial_t(u^2)\,\mathrm{d}x\,\mathrm{d}t - \iint_{Q_2} Muf\,\mathrm{d}x\,\mathrm{d}t \\ &= -\frac{1}{2}\iint_{Q_2}\nabla(dM)\cdot\nabla(u^2)\,\mathrm{d}x\,\mathrm{d}t + \frac{1}{2}\int_{B_2} Mu^2(0,x)\,\mathrm{d}x - \frac{1}{2}\iint_{Q_2}\Delta(dM)u^2\,\mathrm{d}x\,\mathrm{d}t - \iint_{Q_2} Muf\,\mathrm{d}x\,\mathrm{d}t \\ &= \frac{1}{2}\int_{B_2} Mu^2(0,x)\,\mathrm{d}x - \iint_{Q_2} Muf\,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

For the last term, since $\mathrm{supp}(f) \subset Q_1$, the integral actually reduces over Q_1 only. The Hölder inequality then yields

$$\begin{aligned}\left|\iint_{Q_2} Muf\,\mathrm{d}x\,\mathrm{d}t\right| &= \left|\iint_{Q_1} Muf\,\mathrm{d}x\,\mathrm{d}t\right| \\ &\le \|u\|_{L^\infty(Q_1)}\|M\|_{L^{(N+1)/N}(Q_1)}\|f\|_{L^{N+1}(Q_1)} \\ &\le \mathscr{C}\,\|f\|^2_{L^{N+1}(Q_1)}\|M\|_{L^{(N+1)/N}(Q_1)},\end{aligned}$$

by using (13). Additionally, still by using (13) and $\mathrm{supp}(f) \subset Q_1$, we get

$$\begin{aligned}\frac{1}{2}\iint_{Q_2} Mu^2(0,x)\,\mathrm{d}x &\le \frac{1}{2}\|u\|^2_{L^\infty(Q_2)}\|M\|_{L^\infty(-4,0;L^1(Q_2))} \\ &\le \mathscr{C}^2\,\|f\|^2_{L^{N+1}(Q_1)}\|M\|_{L^\infty(-4,0;L^1(Q_2))}.\end{aligned}$$

The last two terms in the right-hand side of (14) are estimated as follows: we get

$$\iint_{Q_2} dM|\nabla\zeta|^2 \,\mathrm{d}x\,\mathrm{d}t \le 4\delta^\star \|\zeta\|^2_{W^{1,\infty}(B_2)} \|M\|_{L^\infty((-4,0);L^1(B_2))}$$

and

$$\begin{aligned}\left|\iint_{Q_2} udM\Delta\zeta \,\mathrm{d}x\,\mathrm{d}t\right| &\le 4\delta^\star \|\zeta\|_{W^{2,\infty}(B_2)} \|u\|_{L^\infty(Q_2)} \|M\|_{L^\infty((-4,0);L^1(B_2))}\\ &\le 4\delta^\star \|\zeta\|_{W^{2,\infty}(B_2)} \mathscr{C} \|f\|_{L^{N+1}(Q_1)} \|M\|_{L^\infty((-4,0);L^1(B_2))}.\end{aligned}$$

The first integral in the right-hand side of (14) is dominated by

$$\|u\|_{L^\infty(Q_2)} \|M\|_{L^\infty(-4,0;L^1(B_2))} \le \mathscr{C} \|f\|_{L^{N+1}(Q_2)} \|M\|_{L^\infty(-4,0;L^1(B_2))}.$$

Finally, we have found a constant $C > 0$ such that for any $f \in C_c^\infty(Q_1)$, with $\|f\|_{L^{N+1}(Q_1)} \le 1$, we have

$$\left|\iint_{Q_1} fM \,\mathrm{d}x\,\mathrm{d}t\right| \le C\Big(\Big(1+\kappa+\frac{1}{\kappa}\Big)\|M\|_{L^\infty(-4,0;L^1(B_2))} + \kappa \|M\|_{L^{(N+1)/N}(Q_1)}\Big).$$

Taking the supremum over such f's makes the dual norm $L^{(N+1)/N}(Q_1)$ appear. We choose κ small enough, so that $1-\kappa C > 1$, and we conclude that

$$\|M\|_{L^{(N+1)/N}(Q_1)} \le \frac{C(1+\kappa+1/\kappa)}{1-\kappa C} \|M\|_{L^\infty(-4,0;L^1(B_2))}$$

holds. □

5. End of proof of Theorem 1.2: proof of Lemma 2.3

Let $0 < \epsilon_0 < \sqrt{T_{\max}/2}$. For each component $a_i^{(\epsilon)}$, Proposition 4.1 gives

$$\|a_i^{(\epsilon)}\|_{L^{(N+1)/N}(Q_1)} \le \|M^{(\epsilon)}\|_{L^{(N+1)/N}(Q_1)} \le K\ \|M^{(\epsilon)}\|_{L^\infty(-4,0;L^1(B_2))}. \tag{15}$$

Next, Lemma 3.3, yields

$$\|M^{(\epsilon)}\|_{L^\infty(-4,0;L^1(B_2))} \le c\|\Phi\|_{L^\infty} \epsilon^{\alpha-2+2/(q-1)}. \tag{16}$$

Combining (15) and (16) with Proposition 3.2 leads to

$$\sum_{i=1}^{p} \|a_i^{(\epsilon)}\|_{L^{(N+1)/N}(Q_1)} \le \mathscr{K}\ \|a^0\|^{1-2/N}_{L^\infty(\mathbb{R}^N)} \|a^0\|^{2/N}_{L^1(\mathbb{R}^N)}\ \epsilon^{\alpha-2+2/(q-1)} \tag{17}$$

for a constant $\mathscr{K}$ which depends on p and N. This information is useful as far as the degree of nonlinearities is such that the exponent remains positive, which means $q \le 2+\alpha/(2-\alpha)$. It ends the proof of Lemma 2.3.

As explained in Section 2, having at hand this property of the rescaled solution we go back to the original unknown, and we deduce the L^∞ bound of the solution, see Corollary 2.4. Theorem 1.2, and therefore Theorem 1.1 too, is fully justified. □

Remark 5.1. The estimates discussed above differ from [Caputo and Vasseur 2009] (see specifically Corollary 14 and Lemma 15), and in particular the smallness condition on ϵ_0 does not involve the initial entropy (9).

Appendix A: A basic iteration lemma

The De Giorgi approach leads us to construct sequences, based on energy-entropy estimates, where the parameter of the sequence controls level sets of the solution and space-time localization. Roughly speaking, we obtain a nonlinear control of the n-th level by the $(n-1)$-th level. Namely, if u_n characterizes a level set associated to a value $\eta_n > 0$, over a domain Q_n, we obtain inequalities like $u_n \le \Lambda u_{n-1}^{\gamma}$. We can finally conclude to a local property of the solution, letting n go to ∞ by using the following simple result.

Lemma A.1. *Let $(u_n)_{n\in\mathbb{N}}$ be a sequence of nonnegative real numbers. We suppose that it satisfies, for any $n \in \mathbb{N}\setminus\{0\}$,*

$$u_n \le \Lambda^n u_{n-1}^{\gamma},$$

where $\Lambda, \gamma > 1$. Then, there exists $\kappa > 0$ (depending on Λ, γ) such that, if $0 \le u_0 \le \kappa$, then $\lim_{n\to\infty} u_n = 0$.

Proof. We set $v_n = \ln(u_n)$ which satisfies

$$v_n \le n \ln(\Lambda) + \gamma v_{n-1},$$

and thus

$$v_n \le \ln(\Lambda) \sum_{j=0}^{n} j \gamma^{n-j} + v_0 \gamma^n \le \gamma^n \ln(\Lambda^{F(\gamma)} u_0),$$

with

$$F(\gamma) = \frac{1}{\gamma} \sum_{j=0}^{\infty} j \Big(\frac{1}{\gamma}\Big)^{j-1} = \frac{1}{\gamma} \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{1}{1-x}\Big)\Big|_{x=1/\gamma} = \frac{1}{\gamma}\Big(\frac{1}{1-1/\gamma}\Big)^2.$$

Therefore v_n tends to $-\infty$, and u_n tends to 0, as $n \to \infty$ provided u_0 is small enough. □

Appendix B: Proof of Lemma 2.2

The proof is based on the De Giorgi techniques [1957] and it is reminiscent of the method introduced by Alikakos [1979]. We exploit the dissipative properties of the system by considering the following nonnegative, nondecreasing, convex, and C^1 function

$$H(z) = \begin{cases} (1+z)\ln(1+z) - z & \text{if } z \ge 0, \\ 0 & \text{if } z \le 0. \end{cases}$$

Let us introduce the sequences, for $j \in \mathbb{N}$,

$$k_j = 1 - 2^{-j}, \quad t_j = \tfrac{1}{4} + 2^{-j-2}.$$

Henceforth, we set $\mathcal{B}_j = B_{t_j}$ and $\mathcal{Q}_j = (-t_j, 0) \times \mathcal{B}_j$. Note that

$$B\big(0, \tfrac{1}{4}\big) \subset \mathcal{B}_j \subset \mathcal{B}_{j-1} \subset B\big(0, \tfrac{1}{2}\big),$$
$$\big(-\tfrac{1}{4}, 0\big) \times B\big(0, \tfrac{1}{4}\big) \subset \mathcal{Q}_j \subset \mathcal{Q}_{j-1} \subset \big(-\tfrac{1}{2}, 0\big) \times B\big(0, \tfrac{1}{2}\big).$$

We also introduce a family of cut-off functions that satisfies the properties

$$\zeta_j : \mathbb{R}^N \to [0, \infty), \quad \zeta_j \in C_c^\infty(\mathbb{R}^N),$$
$$0 \le \zeta_j(x) \le 1,$$
$$\zeta_j(x) = 1 \text{ for } x \in \mathcal{B}_j, \quad \zeta_j(x) = 0 \text{ for } x \in \mathbb{R}^N \setminus \mathcal{B}_{j-1},$$

and

$$\sup_{l,m\in\{1,\dots,N\},\ x\in\mathbb{R}^N} |\partial^2_{l,m}\zeta_j(x)| \le C\, 2^{2j} \quad \text{for a certain constant } C > 0.$$

Lemma B.1. *There exists a constant* $\widehat{C} > 0$, *which depends only on* $\delta_\star$, $\delta^\star$, *and on* (h1)–(h4), *such that for any solution* $a = (a_1, \dots, a_p)$ *of* (6) *and any* $\eta \in [0, 1]$, *we have*

$$\sup_{-t_j \le t \le 0} \sum_{i=1}^p \int_{\mathcal{B}_j} H(a_i - \eta)(t, x)\,\mathrm{d}x + 4\delta_\star \sum_{i=1}^p \iint_{\mathcal{Q}_j} \big|\nabla_x \sqrt{1 + [a_i - \eta]_+}\big|^2(\tau, x)\,\mathrm{d}x\,\mathrm{d}\tau$$
$$\le \widehat{C}\Bigg(2^{2j} \sum_{i=1}^p \int_{-t_{j-1}}^0 \int_{\mathcal{B}_{j-1}} H(a_i - \eta)(s, x)\,\mathrm{d}x\,\mathrm{d}s$$
$$+ \sum_{i=1}^p \int_{-t_{j-1}}^0 \int_{\mathcal{B}_{j-1}} (1 + [a_i - \eta]_+)^{q-1} \ln(1 + [a_i - \eta]_+)(\tau, x)\,\mathrm{d}x\,\mathrm{d}\tau\Bigg).$$

Proof. Multiply (6) by $\zeta_j H'(a_i - \eta)$, integrate over $\mathcal{B}_{j-1}$ and sum. We get

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{i=1}^p \int_{\mathcal{B}_{j-1}} \zeta_j H(a_i - \eta)\,\mathrm{d}x = \sum_{i=1}^p \int_{\mathcal{B}_{j-1}} d_i \Delta a_i\, H'(a_i - \eta)\zeta_j\,\mathrm{d}x + \sum_{i=1}^p \int_{\mathcal{B}_{j-1}} Q_i(a) H'(a_i - \eta)\zeta_j\,\mathrm{d}x. \tag{18}$$

The first term in the right-hand side of (18) can be written as

$$-\sum_{i=1}^p \int_{\mathcal{B}_{j-1}} d_i |\nabla a_i|^2\, H''(a_i - \eta)\zeta_j\,\mathrm{d}x + \sum_{i=1}^p \int_{\mathcal{B}_{j-1}} d_i H(a_i - \eta)\Delta\zeta_j\,\mathrm{d}x,$$

where, on the one hand,

$$\sum_{i=1}^p \int_{\mathcal{B}_{j-1}} d_i |\nabla_x a_i|^2\, H''(a_i - \eta)\zeta_j\,\mathrm{d}x \ge 4\delta_\star \sum_{i=1}^p \int_{\mathcal{B}_j} \big|\nabla_x \sqrt{1 + [a_i - \eta]_+}\big|^2\,\mathrm{d}x,$$

and, on the other hand,

$$\sum_{i=1}^p \int_{\mathcal{B}_{j-1}} d_i H(a_i - \eta)\Delta\zeta_j\,\mathrm{d}x \le C\delta^\star 2^{2j} \int_{\mathcal{B}_{j-1}} H(a_i - \eta)\,\mathrm{d}x.$$

For the second term in the right-hand side of (18), we get

$$\begin{aligned}\sum_{i=1}^{p}\int_{\mathcal{B}_{j-1}} Q_i(a)H'(a_i-\eta)\zeta_j\,\mathrm{d}x &= \sum_{i=1}^{p}\int_{\mathcal{B}_{j-1}} (Q_i(a)-Q_i(1+[a-\eta]_+))\ln(1+[a_i-\eta]_+)\,\zeta_j\,\mathrm{d}x \\ &\qquad+\underbrace{\sum_{i=1}^{p}\int_{\mathcal{B}_{j-1}} Q_i(1+[a-\eta]_+)\ln(1+[a_i-\eta]_+)\zeta_j\,\mathrm{d}x}_{\leq 0 \text{ by (h4)}} \\ &\leq 2p\mathcal{Q}\sum_{i=1}^{p}\int_{\mathcal{B}_{j-1}} (1+[a_i-\eta]_+)^{q-1}\ln(1+[a_i-\eta]_+)\,\mathrm{d}x.\end{aligned}$$

The last estimate is a consequence of (h1) and of the elementary inequality

$$|1+[a-\eta]_+-a|\leq 1+|[a-\eta]_+-a|\leq 1+\eta\leq 2;$$

see [Goudon and Vasseur 2010, proof of Lemma 3.1] or [Caputo and Vasseur 2009, Lemma 3]. We arrive at

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{p}\int_{\mathcal{B}_{j-1}} \zeta_j H(a_i-\eta)\,\mathrm{d}x + 4\delta_\star\sum_{i=1}^{p}\int_{\mathcal{B}_j}\big|\nabla_x\sqrt{1+[a_i-\eta]_+}\big|^2\,\mathrm{d}x \\ &\qquad\leq C\delta^\star 2^{2j}\sum_{i=1}^{p}\int_{\mathcal{B}_{j-1}} H(a_i-\eta)\,\mathrm{d}x + 2p\mathcal{Q}\sum_{i=1}^{p}\int_{\mathcal{B}_{j-1}}(1+[a_i-\eta]_+)^{q-1}\ln(1+[a_i-\eta]_+)\,\mathrm{d}x.\end{aligned}$$

We integrate this relation over (s,t), with $-t_j\leq t\leq 0$ and $-t_{j-1}\leq s\leq t_j$, and next we average with respect to $s\in(-t_{j-1},-t_j)$, taking into account that $t_{j-1}-t_j=2^{-j-2}$. We obtain

$$\begin{aligned}&\sum_{i=1}^{p}\int_{\mathcal{B}_j} H(a_i-\eta)(t,x)\,\mathrm{d}x + 4\delta_\star\sum_{i=1}^{p}\int_{-t_j}^{t}\int_{\mathcal{B}_j}\big|\nabla_x\sqrt{1+[a_i-\eta]_+}\big|^2(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau \\ &\quad\leq \sum_{i=1}^{p}\int_{\mathcal{B}_{j-1}}\zeta_j H(a_i-\eta)(t,x)\,\mathrm{d}x \\ &\qquad+4\delta_\star\sum_{i=1}^{p}\frac{1}{2^{-j-2}}\int_{-t_{j-1}}^{-t_j}\int_s^t\int_{\mathcal{B}_j}\big|\nabla_x\sqrt{1+[a_i-\eta]_+}\big|^2(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}s \\ &\quad\leq \sum_{i=1}^{p}\frac{1}{2^{-j-2}}\int_{-t_{j-1}}^{-t_j}\int_{\mathcal{B}_{j-1}}\zeta_j H(a_i-\eta)(s,x)\,\mathrm{d}x\,\mathrm{d}s \\ &\qquad+C\delta^\star 2^{2j}\sum_{i=1}^{p}\frac{1}{2^{-j-2}}\int_{-t_{j-1}}^{-t_j}\int_s^t\int_{\mathcal{B}_{j-1}} H(a_i-\eta)(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}s \\ &\qquad\quad+2p\mathcal{Q}\sum_{i=1}^{p}\frac{1}{2^{-j-2}}\int_{-t_{j-1}}^{-t_j}\int_s^t\int_{\mathcal{B}_{j-1}}(1+[a_i-\eta]_+)^{q-1}\ln(1+[a_i-\eta]_+)(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau\,\mathrm{d}s\end{aligned}$$

$$\leq 2^{j+2}\sum_{i=1}^{p}\int_{-t_{j-1}}^{-t_j}\int_{\mathcal{B}_{j-1}} H(a_i-\eta)(s,x)\,\mathrm{d}x\,\mathrm{d}s + C\delta^\star 2^{2j}\sum_{i=1}^{p}\int_{-t_{j-1}}^{t}\int_{\mathcal{B}_{j-1}} H(a_i-\eta)(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau$$
$$+2p\mathcal{Q}\sum_{i=1}^{p}\int_{-t_{j-1}}^{t}\int_{\mathcal{B}_{j-1}}(1+[a_i-\eta]_+)^{q-1}\ln(1+[a_i-\eta]_+)(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau.$$

We conclude by taking the supremum over $t\in(-t_j,0)$. □

Next, we specify the level set considered in these estimates: we use Lemma B.1 with $\eta=k_j$ and we set

$$\mathcal{U}_j=\left(\sup_{-t_j\leq t\leq 0}\sum_{i=1}^{p}\int_{\mathcal{B}_j} H(a_i-k_j)\,\mathrm{d}x+\sum_{i=1}^{p}\iint_{\mathcal{Q}_j}\left|\nabla_x\sqrt{1+[a_i-k_j]_+}\right|^2\mathrm{d}x\,\mathrm{d}s\right).$$

Lemma B.2. *Let* $2\leq q<2(N+1)/N$. *Then*:

(i) *For any* $r>1$ *there exists a universal constant* $c_r>0$ *such that*

$$\mathcal{U}_0\leq c_r\sum_{i=1}^{p}(\|a_i\|^r_{L^r((-1,0)\times B_1)}+\|a_i\|^{1/2}_{L^r((-1,0)\times B_1)}+\|a_i\|_{L^r((-1,0)\times B_1)}).$$

(ii) *There exists a constant* $\Lambda>1$ *such that*

$$\mathcal{U}_j\leq\Lambda^j\,\mathcal{U}_{j-1}^{1+N/2}$$

for any $j\geq j_0$. *Consequently, there exists* $\delta>0$ *such that* $\mathcal{U}_0\leq\delta$ *implies* $\lim_{j\to\infty}\mathcal{U}_j=0$.

Proof. Throughout the proof, we simply denote by c a constant that depends only on the parameters of the model, and on the Lebesgue exponent, without paying attention to the possible changes of the value of the constant from a line to another.

For proving (i), we go back to the definition

$$\mathcal{U}_0=\left(\sup_{-1/2\leq t\leq 0}\sum_{i=1}^{p}\int_{\mathcal{B}_0} H(a_i)\,\mathrm{d}x+\sum_{i=1}^{p}\iint_{\mathcal{Q}_0}\left|\nabla_x\sqrt{a_i+1}\right|^2\mathrm{d}x\,\mathrm{d}\tau\right),$$

where we remind the reader that $\mathcal{B}_0=B_{1/2}$ and $\mathcal{Q}_0=\left(-\frac{1}{2},0\right)\times B_{1/2}$. We make use of the elementary inequalities

$$H(z)\leq c(z(1+|\ln(z))|), \tag{19}$$
$$\left|\nabla\sqrt{1+a}\right|\leq\left|\nabla\sqrt{a}\right|, \tag{20}$$

which hold for any $z\geq 0$ and any (smooth enough) function $a:\mathbb{R}^N\to[0,\infty)$, respectively. We consider $\zeta_0\in C_c^\infty(\mathbb{R}^N)$, supported in B_1, such that $0\leq\zeta_0(x)\leq 1$ on $\mathbb{R}^N$ and $\zeta_0(x)=1$ on $\mathcal{B}_0$. We get

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{p}\int_{B_1}\zeta_0(x)a_i(t,x)\,\mathrm{d}x=\sum_{i=1}^{p}\int_{B_1} d_i\Delta\zeta_0(x)\,a_i(t,x)\,\mathrm{d}x\leq\delta^\star\|\Delta\zeta_0\|_{L^\infty}\sum_{i=1}^{p}\int_{B_1} a_i(t,x)\,\mathrm{d}x.$$

Let $t \in \left(-\frac{1}{2}, 0\right)$ and $\tau \in (-1, t)$. We integrate over the time interval (τ, t), and then we average over $\tau \in \left(-1, -\frac{1}{2}\right)$. We are led to

$$\sup_{-1/2 \le t \le 0} \sum_{i=1}^{p} \int_{\mathcal{B}_0} a_i(t,x)\,\mathrm{d}x \le c \sum_{i=1}^{p} \int_{-1}^{0} \int_{B_1} a_i(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau.$$

Similarly, the localized version of the entropy dissipation becomes

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{i=1}^{p} \int_{B_1} \zeta_0(x)\, a_i \ln(a_i)\,\mathrm{d}x + \int_{B_1} \zeta_0(x) \frac{d_i |\nabla_x a_i|^2}{a_i}\,\mathrm{d}x = \sum_{i=1}^{p} \int_{B_1} \Delta\zeta_0 d_i (a_i \ln(a_i) - a_i)\,\mathrm{d}x.$$
$$\le \delta^\star \|\Delta\zeta_0\|_{L^\infty} \sum_{i=1}^{p} \int_{B_1} (a_i |\ln(a_i)| + a_i)\,\mathrm{d}x.$$

Again we integrate with respect to the time variable. We shall also use the trick

$$u|\ln(u)| = u\ln(u)\mathbf{1}_{u\ge 1} - u\ln(u)\mathbf{1}_{0\le u<1} \le u\ln(u)\mathbf{1}_{u\ge 1} + \frac{2}{e}\sqrt{u}\mathbf{1}_{0\le u<1},$$

which allows us to dominate

$$u|\ln(u)| \le c(u^r + \sqrt{u}).$$

It follows that

$$\sup_{-1/2 \le t \le 0} \sum_{i=1}^{p} \int_{\mathcal{B}_0} a_i |\ln(a_i)|\,\mathrm{d}x + 4\delta_\star \sum_{i=1}^{p} \int_{-1/2}^{0} \int_{\mathcal{B}_0} \left|\nabla_x \sqrt{a_i}\right|^2 \mathrm{d}x\,\mathrm{d}\tau$$
$$\le c \sum_{i=1}^{p} \left(\int_{-1}^{0} \int_{B_1} (a_i^r + \sqrt{a_i} + a_i)\,\mathrm{d}x\,\mathrm{d}\tau \right)$$
$$\le c \sum_{i=1}^{p} \left(\int_{-1}^{0} \int_{B_1} |a_i|^r\,\mathrm{d}x\,\mathrm{d}\tau + \left(\int_{-1}^{0} \int_{B_1} |a_i|^r\,\mathrm{d}x\,\mathrm{d}\tau \right)^{1/(2r)} \mathrm{meas}(B_1)^{1-1/(2r)} \right.$$
$$\left. + \left(\int_{-1}^{0} \int_{B_1} |a_i|^r\,\mathrm{d}x\,\mathrm{d}\tau \right)^{1/r} \mathrm{meas}(B_1)^{1-1/r} \right),$$

by using the Hölder inequality.

We turn to the proof of (ii). The estimate in Lemma B.1 can be recast as

$$\mathscr{U}_j \le C \left(2^{2j} \sum_{i=1}^{p} \iint_{\mathscr{Q}_{j-1}} H(a_i - k_j)(s,x)\,\mathrm{d}x\,\mathrm{d}s \right.$$
$$\left. + \sum_{i=1}^{p} \iint_{\mathscr{Q}_{j-1}} (1 + [a_i - k_j]_+)^{q-1} \ln(1 + [a_i - k_j]_+)(s,x)\,\mathrm{d}x\,\mathrm{d}s \right). \quad (21)$$

Let us set

$$\Psi(z) = \sqrt{1+z} - 1.$$

For any $\gamma \ge 1,\ \beta > 0$, we can find a constant $c_{\gamma,\beta}$ such that

$$(1+z)^\gamma \ln(1+z) \le c_\beta \Psi(z)^{2(\gamma+\beta)}.$$

Moreover, for $z \geq k_j \geq k_{j-1}$ we have

$$1 \leq \frac{z - k_{j-1}}{k_j - k_{j-1}} = 2^j (z - k_{j-1}).$$

Hence, we can estimate both integrals in the right-hand side of (21) by an expression like

$$\sum_{i=1}^{p} \iint_{\mathcal{Q}_{j-1}} 2^{\gamma j} (1 + [a_i - k_{j-1}]_+)^{\gamma} \ln(1 + [a_i - k_{j-1}]_+) \, \mathrm{d}x \, \mathrm{d}s \\ \leq c_{\gamma,\beta} 2^{\gamma j} \sum_{i=1}^{p} \iint_{\mathcal{Q}_{j-1}} \Psi([a_i - k_{j-1}]_+)^{\gamma+\beta} \, \mathrm{d}x \, \mathrm{d}s.$$

We can play with the exponents γ and β for both terms so that we obtain a common bound from above, and we arrive at

$$\mathscr{U}_j \leq c 2^{4j} \sum_{i=1}^{p} \iint_{\mathcal{Q}_{j-1}} \Psi([a_i - k_{j-1}]_+)^{2(N+2)/N} \, \mathrm{d}x \, \mathrm{d}s.$$

This is possible as far as $2(q-1) \leq 2(N+2)/N$, that is to say $q \leq 2(N+1)/N$. We shall conclude by using an interpolation argument. Indeed, on the one hand, we obviously have

$$\sup_{-t_{j-1} \leq s \leq 0} \int_{\mathcal{B}_{j-1}} |\Psi([a_i - k_{j-1}]_+)|^2 (s, x) \, \mathrm{d}x \leq \mathscr{U}_{j-1},$$

while the Gagliardo–Nirenberg–Sobolev inequality, see [Nirenberg 1959, Theorem, p. 125], yields

$$\int_{-t_{j-1}}^{0} \left(\int_{\mathcal{B}_{j-1}} |\Psi([a_i - k_{j-1}]_+)|^{2N/(N-2)} (s, x) \, \mathrm{d}x \right)^{(N-2)/N} \mathrm{d}s \\ \leq c \iint_{\mathcal{Q}_{j-1}} |\nabla \Psi([a_i - k_{j-1}]_+)|^2 (s, x) \, \mathrm{d}x \, \mathrm{d}s \leq c \mathscr{U}_{j-1}.$$

By using the interpolation

$$\frac{N+2}{N} = \theta \frac{2N}{N-2} + 2(1-\theta), \quad \theta = \frac{N-2}{N} \in (0, 1),$$

we combine these into

$$\iint_{\mathcal{Q}_{j-1}} |\Psi([a_i - k_{j-1}]_+)|^{2(N+2)/N} (s, x) \, \mathrm{d}x \, \mathrm{d}s \\ \leq \int_{-t_{j-1}}^{0} \left(\int_{\mathcal{B}_{j-1}} |\Psi([a_i - k_{j-1}]_+)|^{2N/(N-2)} (s, x) \, \mathrm{d}x \right)^{\theta} \left(\int_{\mathcal{B}_{j-1}} |\Psi([a_i - k_{j-1}]_+)|^2 (s, x) \, \mathrm{d}x \right)^{1-\theta} \mathrm{d}s \\ \leq \mathscr{U}_{j-1}^{1-\theta} \int_{-t_{j-1}}^{0} \left(\int_{\mathcal{B}_{j-1}} |\Psi([a_i - k_{j-1}]_+)|^{2N/(N-2)} (s, x) \, \mathrm{d}x \right)^{(N-2)/N} \mathrm{d}s \leq c \mathscr{U}_{j-1}^{1+2/N}.$$

We conclude by applying Lemma A.1. □

Once we know that $\lim_{j\to\infty}\mathscr{U}_j=0$ we deduce that

$$\lim_{j\to\infty}\frac{1}{t_j}\sum_{i=1}^{p}\iint_{-\mathcal{Q}_j}H(a_i-k_j)\,\mathrm{d}x\,\mathrm{d}t=0\geq 4\sum_{i=1}^{p}\int_{-1/4}^{0}\int_{B(0,1/2)}H(a_i-1)\,\mathrm{d}x\,\mathrm{d}t.$$

It implies that $0\leq a_i(t,x)\leq 1$ holds for a.e. $(t,x)\in\left(-\frac14,0\right)\times B\left(0,\frac14\right)$.

Appendix C: Proof of Proposition 3.2

It is worth giving some hints for the proof of Proposition 3.2, which is fully detailed in [Caputo and Vasseur 2009, Proposition 11, Corollary 12]. Again, the proof heavily relies on duality arguments. The main step consists in showing that

$$\|\Phi(t,\cdot)\|_{L^\infty(\mathbb{R}^N)}\leq\|\Phi(0,\cdot)\|_{L^\infty(\mathbb{R}^N)}. \tag{22}$$

Indeed, we remind the reader that $\Phi(t,x)$ is determined by the convolution formula (for $N>2$)

$$\Phi(t,x)=-C_N\int_{\mathbb{R}^N}\frac{M(t,y)}{|x-y|^{N-2}}\,\mathrm{d}y,$$

where

$$C_N=\frac{1}{(N-2)\sigma_N},$$

with $\sigma_N=2\pi^{N/2}/\Gamma(N/2)$ the measure of the unit sphere of $\mathbb{R}^N$. Thus, given $R>0$, we simply split

$$\Phi(0,x)=-C_N\int_{|x-y|\leq R}\frac{M(0,y)}{|x-y|^{N-2}}\,\mathrm{d}y-C_N\int_{|x-y|>R}\frac{M(0,y)}{|x-y|^{N-2}}\,\mathrm{d}y,$$

which yields

$$|\Phi(0,x)|\leq C_N\|M(0,\cdot)\|_{L^\infty(\mathbb{R}^N)}\frac{\sigma_N R^2}{2}+\frac{C_N}{R^{N-2}}\|M(0,\cdot)\|_{L^1(\mathbb{R}^N)}.$$

Optimizing with respect to R, we get

$$|\Phi(0,x)|\leq K_N\|M(0,\cdot)\|_{L^\infty(\mathbb{R}^N)}^{1-2/N}\|M(0,\cdot)\|_{L^1(\mathbb{R}^N)}^{2/N},$$

where $K_N>0$ depends only on the space dimension $N\geq 3$.

In order to justify (22), we need to introduce a mollified diffusion coefficient. Indeed, as the a_i's are smooth on $[0,T_{\max})\times\mathbb{R}^N$, M is smooth too; thus $(t,x)\mapsto d(t,x)$ is a smooth function, except possibly at the points where $M(t,x)$ vanishes. Given $\mu>0$, we denote by $d_\mu(t,x)$ a smooth function satisfying

$$d_\mu(t,x)=d(t,x)\quad\text{when } M(t,x)\geq\mu,\qquad 0<\delta_\star\leq d_\mu(t,x)\leq\delta^\star.$$

The proof of (22) splits into two steps.

Let $0<T<\infty$. Let $\zeta\in C_c^\infty(\mathbb{R}^N)$ and consider the solution of the *end-value* equation

$$\partial_t\varphi+d_\mu\Delta\varphi=0,\quad \varphi(T,x)=\zeta(x), \tag{23}$$

together with the initial value problem

$$\partial_t \rho - \Delta(d_\mu \rho) = 0, \quad \rho(0,x) = \rho^0(x).$$

We assume that

$$\|\zeta\|_{L^\infty(\mathbb{R}^N)} \le 1.$$

The maximum principle, see for instance [Evans 1998, Theorem 8, Chapter 7], implies

$$\sup_{0\le t\le T} \|\varphi(t,\cdot)\|_{L^\infty(\mathbb{R}^N)} \le \|\zeta\|_{L^\infty(\mathbb{R}^N)} \le 1.$$

We have, by integrating by parts,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N} \rho(t,x)\varphi(t,x)\,\mathrm{d}x &= \int_{\mathbb{R}^N} \big(\partial_t\rho(t,x)\varphi(t,x) + \rho(t,x)\partial_t\varphi(t,x)\big)\,\mathrm{d}x \\
&= \int_{\mathbb{R}^N} \big(\Delta(d_\mu\rho(t,x))\varphi(t,x) - \rho(t,x)d_\mu(t,x)\Delta\varphi(t,x)\big)\,\mathrm{d}x \\
&= -\int_{\mathbb{R}^N} \nabla(d_\mu\rho(t,x))\cdot\nabla\varphi(t,x)\,\mathrm{d}x + \int_{\mathbb{R}^N} \nabla(\rho(t,x)d_\mu(t,x))\cdot\nabla\varphi(t,x)\,\mathrm{d}x \\
&= 0.
\end{aligned}$$

Now, we integrate over $[0,T]$ by using the conditions at $t=0$ for ρ and $t=T$ for φ. It follows that

$$\left|\int_{\mathbb{R}^N} \rho(T,x)\zeta(x)\,\mathrm{d}x\right| = \left|\int_{\mathbb{R}^N} \rho^0(x)\varphi(0,x)\,\mathrm{d}x\right| \le \|\rho^0\|_{L^1(\mathbb{R}^N)}.$$

By virtue of the Hahn–Banach theorem, we conclude that

$$\begin{aligned}
\|\rho(T,\cdot)\|_{L^1(\mathbb{R}^N)} &= \sup\left\{\left|\int_{\mathbb{R}^N} \rho(T,x)\zeta(x)\,\mathrm{d}x\right| : \zeta\in C_c^\infty(\mathbb{R}^N),\ \|\zeta\|_{L^\infty(\mathbb{R}^N)}\le 1\right\} \\
&\le \|\rho^0\|_{L^1(\mathbb{R}^N)}.
\end{aligned}$$

Next, we shall apply similar reasoning in order to make the norm $\|\Delta\zeta\|_{L^1(\mathbb{R}^N)}$ appear. For $0<T<\infty$ and φ a solution of (23), let us set

$$\rho(t,x) = \Delta\varphi(T-t,x),$$

which satisfies

$$\partial_t\rho - \Delta(d_\mu\rho) = 0, \quad \rho(0,x) = \Delta\zeta(x)\in L^1(\mathbb{R}^N).$$

The previous step thus tells us that

$$\|\rho(T,\cdot)\|_{L^1(\mathbb{R}^N)} = \|\Delta\varphi(0,\cdot)\|_{L^1(\mathbb{R}^N)} \le \|\rho(0,\cdot)\|_{L^1(\mathbb{R}^N)} = \|\Delta\zeta\|_{L^1(\mathbb{R}^N)}.$$

Going back to the equation for the total mass, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N} M\varphi(t,x)\,\mathrm{d}x = \int_{\mathbb{R}^N} M(d-d_\mu)\Delta\varphi(t,x)\,\mathrm{d}x.$$

Let $0 < T < T_{\max}$. Integrating over $(0, T)$ yields

$$\begin{aligned}\left|\int_{\mathbb{R}^N} M\varphi(T,x)\,\mathrm{d}x\right| &= \left|\int_{\mathbb{R}^N} \Delta\Phi\varphi(T,x)\,\mathrm{d}x\right| = \left|\int_{\mathbb{R}^N} \Phi(T,x)\Delta\zeta(x)\,\mathrm{d}x\right| \\ &= \left|\int_{\mathbb{R}^N} M\varphi(0,x)\,\mathrm{d}x + \int_0^T\int_{\mathbb{R}^N} M(d-d_\mu)\Delta\varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t\right| \\ &= \left|\int_{\mathbb{R}^N} \Delta\Phi\varphi(0,x)\,\mathrm{d}x + \int_0^T\int_{\mathbb{R}^N} M(d-d_\mu)\Delta\varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t\right| \\ &\le \left|\int_{\mathbb{R}^N} \Phi\Delta\varphi(0,x)\,\mathrm{d}x\right| + \left|+\int_0^T\int_{\mathbb{R}^N} M(d-d_\mu)\Delta\varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t\right| \\ &\le \|\Phi(0,\cdot)\|_{L^\infty(\mathbb{R}^N)}\|\Delta\varphi(0,\cdot)\|_{L^1(\mathbb{R}^N)} + 2T\delta^\star\mu\|\Delta\varphi\|_{L^\infty(0,T;L^1(\mathbb{R}^N))} \\ &\le (\|\Phi(0,\cdot)\|_{L^\infty(\mathbb{R}^N)} + 2T\delta^\star\mu)\|\Delta\zeta\|_{L^1(\mathbb{R}^N)},\end{aligned}$$

where the penultimate inequality holds since

$$|d-d_\mu|M = |d-d_\mu|M\mathbf{1}_{M\le\mu} \le 2\delta^\star\mu.$$

This relation holds for any $\mu > 0$ and $\zeta \in C_c^\infty(\mathbb{R}^N)$. Therefore, we can conclude that (22) holds, which ends the proof. □

Appendix D: Proof of Proposition 3.1 by De Giorgi's approach

For the sake of completeness, we provide here an alternative proof of Proposition 3.1, which, however, uses the additional assumption (b). The interest of this proof is that it entirely relies on energy estimates and De Giorgi's methods, which gives a unified viewpoint on the whole argumentation of the paper. Since the result stated in Proposition 3.1 is standard, the remainder of this section can be safely skipped by the reader not interested in such an alternative proof (the original proof relies on a probabilistic interpretation of the equation and uses arguments from the theory of diffusion processes).

Here and below, given $\rho > 0$, with B_ρ the ball $\{x \in \mathbb{R}^N : |x| \le \rho\}$, we define

$$Q_\rho = (-\rho^2, 0) \times B_\rho.$$

In fact, we shall work within Q_2, considered as a reference domain. From an equation satisfied on Q_2 we wish to establish qualitative properties on a smaller domain, say Q_1 or $Q_{1/2}$. It is also convenient to introduce the domain

$$\widetilde{Q} = \left(-\tfrac{9}{4}, -1\right) \times B_1.$$

We refer the reader to Figure 1; having the picture of the subdomains of Q_2 might be helpful in following the arguments.

The argument for proving Proposition 3.1 relies on a technical lemma that controls oscillations. From now on, for a function φ defined on $\Omega \subset \mathbb{R}^d$, we set

$$\operatorname{osc}(\varphi,\, \Omega) = \sup_{x\in\Omega} \varphi(x) - \inf_{x\in\Omega} \varphi(x).$$

Lemma D.1 (decay of oscillations). *Let Φ satisfy the assumptions of Proposition 3.1. There exists $\lambda \in (0,1)$, which depends only on N and $\delta_\star$, such that*

$$\mathrm{osc}(\Phi, Q_{1/2}) \le \lambda \,\mathrm{osc}(\Phi,\ Q_2).$$

Let us assume temporarily that Lemma D.1 holds true. We pick $(t,x) \in (t_0, T) \times \mathbb{R}^N$, where $0 < t_0 < T < \infty$, and we set

$$\Phi_k(s,y) = \Phi(t + 2^{-2k}s, x + 2^{-k}y),$$

where $k \in \mathbb{N}$ is large enough so that the time variable remains larger than t_0 when $-4 \le s \le 0$; namely, we have

$$k \ge k_0 = \ln\Big(\frac{t-t_0}{4}\Big)\frac{1}{2\ln\big(\frac{1}{2}\big)}.$$

The function Φ_k is defined on Q_2 and it satisfies

$$\partial_s \Phi_k = d_k \Delta_y \Phi_k,$$

where

$$d_k(s,y) = d(t + 2^{-2k}s, x + 2^{-k}y).$$

Moreover, we still have

$$-\|\Phi\|_{L^\infty} \le \Phi_k(s,y) \le +\|\Phi\|_{L^\infty}.$$

Applying Lemma D.1 yields

$$\mathrm{osc}(\Phi_k, Q_{1/2}) \le \lambda\,\mathrm{osc}(\Phi_k,\ Q_2),$$

which can be rewritten as

$$\mathrm{osc}(\Phi(t+\cdot, x+\cdot),\ Q_{2^{-k-1}}) \le \lambda\,\mathrm{osc}(\Phi(t+\cdot, x+\cdot),\ Q_{2^{-k+1}}).$$

We deduce that

$$\mathrm{osc}(\Phi(t+\cdot, x+\cdot),\ Q_{2^{-k}}) \le \sqrt{\lambda}^{k} \times C_0, \quad C_0 = \frac{2}{\sqrt{\lambda}^{k_0}}\|\Phi\|_{L^\infty}.$$

(We should bear in mind the fact that C_0 depends on t_0 through the definition of k_0 and it is proportional to $\|\Phi\|_{L^\infty}$.) Let $x' \in \mathbb{R}^N$ and $t' > t_0$; there exists a unique $k \in \mathbb{N}$ such that $x' - x \in B_{2^{-k+1}} \setminus B_{2^{-k}}$, $2^{-2k} \le |t'-t| \le 2^{-2(k-1)}$. It follows that

$$\frac{|\Phi(t',x') - \Phi(t,x)|}{|t'-t|^{\alpha/2} + |x'-x|^\alpha} \le \frac{C_0}{\sqrt{\lambda}}(\sqrt{\lambda}2^\alpha)^k.$$

If $0 < \sqrt{\lambda} \le \frac{1}{2}$, the right-hand side remains obviously bounded, uniformly with respect to k, for any $0 < \alpha \le 1$; otherwise we choose

$$0 < \alpha = \frac{\ln(1/\sqrt{\lambda})}{\ln(2)} < 1.$$

Hence Proposition 3.1 follows from Lemma D.1. ☐

We are thus left with the task of proving Lemma D.1. To this end, we shall apply the following statement.

Proposition D.2. *Let* $(t,x) \mapsto v(t,x)$ *satisfy*

- *the differential inequality* $\partial_t v - \delta^\star \Delta v \le 0$ *on* Q_2;
- $-1 \le v(t,x) \le +1$ *on* Q_2;
- $\mathrm{meas}(\{(t,x) \in \widetilde{Q} : v(t,x) \le 0\}) \ge \mu\, \mathrm{meas}(\widetilde{Q})$ *for some* $\mu > 0$.

Then, there exists $0 < \eta < 1$ *such that*

$$v(t,x) \le \eta \quad \text{on } Q_{1/2}.$$

The function

$$\widetilde{\Phi}(t,x) = \frac{2}{\mathrm{osc}(\Phi, Q_2)} \left(\Phi(t,x) - \frac{\sup_{Q_2} \Phi + \inf_{Q_2} \Phi}{2} \right)$$

satisfies the first two assumptions of Proposition D.2. Suppose that

$$\mathrm{meas}(\{(t,x) \in Q_2 : \widetilde{\Phi}(t,x) \le 0\}) \ge \frac{\mathrm{meas}(Q_2)}{2}.$$

(Otherwise, we shall apply the same reasoning to $-\widetilde{\Phi}$.) Proposition D.2 tells us that $\widetilde{\Phi}(t,x) \le \eta$ on $Q_{1/2}$, which yields $\mathrm{osc}(\widetilde{\Phi}, Q_{1/2}) \le 1 + \eta$ (since $\inf_{Q_{1/2}} \widetilde{\Phi} \ge -1$), and thus

$$\mathrm{osc}(\Phi, Q_{1/2}) \le \frac{1+\eta}{2} \mathrm{osc}(\Phi, Q_2).$$

It justifies Lemma D.1, with $\lambda = (1+\eta)/2 \in (0,1)$. □

The proof of Proposition D.2 relies on a series of intermediate statements.

Lemma D.3. *Let* $-\infty < a, b < \infty$ *and let* Ω *be a smooth bounded domain in* $\mathbb{R}^N$. *We define* $Q = (a,b) \times \Omega$.

(a) *Let* $u \in L^\infty(a,b; L^2(\Omega)) \cap L^2(a,b; H^1(\Omega))$ *such that*

$$\partial_t u - \delta^\star \Delta u + \mu = 0$$

holds in $\mathscr{D}'(Q)$, *with* μ *a nonnegative measure on* Q. *Let* $F : \mathbb{R} \to \mathbb{R}$ *be a nondecreasing convex function. We assume that* $F(0) = 0$ *and* $F \in W^{1,\infty}_{\mathrm{loc}}(\mathbb{R})$. *Then, there exists a nonnegative measure* ν *such that* $v = F(u)$ *satisfies* $\partial_t v - \delta^\star \Delta v + \nu = 0$ *in* $\mathscr{D}'(Q)$.

(b) *Let* $v \in L^\infty((a,b) \times \Omega) \cap L^2(a,b; H^1(\Omega))$ *be a nonnegative solution of* $\partial_t v - \delta^\star \Delta v + \nu = 0$, *with* ν *a nonnegative measure on* Q. *Then, for any trial function* $\varphi \in C_c^\infty(\Omega)$ *there exists* $C > 0$, *which depends only on* $\delta_\star$, $\|v\|_{L^\infty}$ *and* φ, *such that, for a.e.* $a < s < t < b$, *the following energy inequality holds*:

$$\frac{1}{2} \int_\Omega v^2(t,x) \varphi^2(x)\, \mathrm{d}x + \delta_\star \int_s^t \int_\Omega |\nabla(\phi v)|^2(\tau,x)\, \mathrm{d}x\, \mathrm{d}\tau \le \frac{1}{2} \int_\Omega v^2(s,x) \varphi^2(x)\, \mathrm{d}x + C(t-s).$$

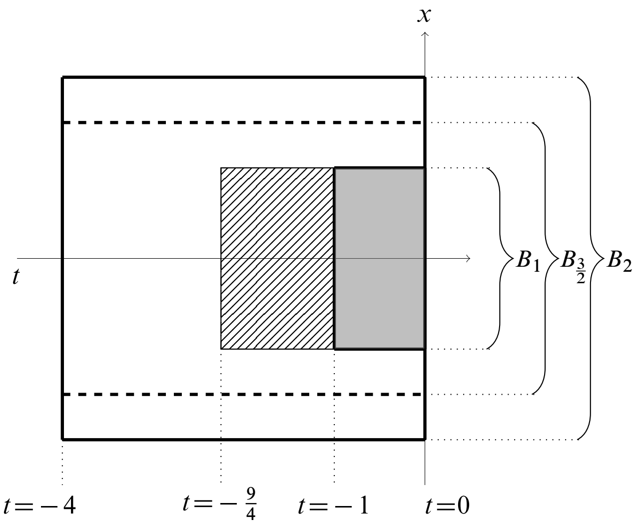

Figure 1. The domains Q_2 (the largest box), $\widetilde{Q}$ (the dashed box) and Q_1 (the gray box).

Proof. Note that $v = F(u)$ also lies in $L^\infty(a,b;L^2(\Omega)) \cap L^2(a,b;H^1(\Omega))$; see, e.g., [Brezis 1983, Proposition IX.5]. Item (a) follows from the computation

$$\partial_t F(u) = -F'(u)\mu + F'(u)\delta_\star \Delta u = \underbrace{-F'(u)\mu - \delta^\star F\text{''}(u)|\nabla u|^2}_{\leq 0} + \delta_\star \Delta F(u).$$

The argument can be made rigorous by working on the weak variational formulation of the equation, with suitable approximation of the solution u.

For proving item (b), we compute

$$\begin{aligned}\tfrac{1}{2}\partial_t(v^2\varphi^2) &= \delta^\star \varphi^2 v \nabla\cdot\nabla v - \nu\varphi^2 v\\ &= \delta^\star \nabla\cdot(\varphi^2 v\nabla v) - \nu\varphi^2 v - \delta^\star \nabla v\cdot\nabla(\varphi^2 v)\\ &= \delta^\star \nabla\cdot(\varphi^2 v\nabla v) - \nu\varphi^2 v - \delta^\star |\nabla(\varphi v)|^2 + \delta^\star v^2|\nabla\varphi|^2.\end{aligned}$$

The second and third terms of the right-hand side are nonpositive; the integral of the last term is dominated by $\delta_\star \|v\|^2_{L^\infty(Q)}\|\varphi\|_{H^1(\Omega)}$. Again a full justification proceeds through an approximation argument. □

For proving Proposition D.2, we shall work with several subdomains of Q_2, as indicated by Figure 1 which might help to follow the arguments.

Lemma D.4. *Let u satisfy $\partial_t u - \delta^\star \Delta u \leq 0$ and $-1 \leq u(t,x) \leq +1$ in Q_2. Let us set*

$$\begin{aligned}\mathscr{A} &= \{(t,x)\in Q_1 : u(t,x) \geq \tfrac{1}{2}\},\\ \mathscr{B} &= \{(t,x)\in \widetilde{Q} : u(t,x) \leq 0\},\\ \mathscr{C} &= \{(t,x)\in Q_1\cup\widetilde{Q} : 0 < u(t,x) < \tfrac{1}{2}\}.\end{aligned}$$

There exists $\alpha > 0$ such that if $\mathrm{meas}(\mathscr{A}) \geq \eta$ *and* $\mathrm{meas}(\mathscr{B}) \geq \frac{1}{2}\,\mathrm{meas}(\widetilde{Q})$, *then* $\mathrm{meas}(\mathscr{C}) \geq \alpha$.

Proof. We argue by contradiction, assuming that a sequence $(u_k)_{k\in\mathbb{N}}$ of solutions of $\partial_t u_k - \delta^\star \Delta u_k \le 0$ in Q_2 satisfies $-1 \le u_k(t,x) \le +1$ and

$$\begin{aligned} &\operatorname{meas}(\mathscr{A}_k) \ge \eta, && \text{with } \mathscr{A}_k = \big\{(t,x) \in Q_1 : u_k(t,x) \ge \tfrac{1}{2}\big\}, \\ &\operatorname{meas}(\mathscr{B}_k) \ge \tfrac{1}{2} \operatorname{meas}(\widetilde{Q}), && \text{with } \mathscr{B}_k = \{(t,x) \in \widetilde{Q} : u_k(t,x) \le 0\}, \\ &\operatorname{meas}(\mathscr{C}_k) \le \tfrac{1}{k}, && \text{with } \mathscr{C}_k = \big\{(t,x) \in Q_1 \cup \widetilde{Q} : 0 < u_k(t,x) < \tfrac{1}{2}\big\}. \end{aligned} \tag{24}$$

We focus our interest on the positive part $v_k = [u_k]_+$, with $[z]_+ = \max(z,0)$, which is still uniformly bounded: $0 \le v_k(t,x) \le 1$. By virtue of Lemma D.3(a), it satisfies

$$\partial_t v_k - \delta^\star \Delta v_k + \mu_k = 0, \tag{25}$$

with μ_k a nonnegative measure. The strategy can be recapped as follows. We shall establish the compactness of v_k in the reduced domain $(-4,0) \times B_{3/2}$. It allows us to assume that v_k converges to a certain function v. Roughly speaking, we are going to show that $v(s,x)$ vanishes on B_1 for certain times $-\frac{3}{2} < s < -1$, which will imply that v vanishes over Q_1. It will eventually lead to a contradiction by considering the behavior of the sets $\mathscr{A}_k, \mathscr{B}_k, \mathscr{C}_k$ as $k \to \infty$.

Let us pick a trial function $\zeta \in C_c^\infty(B_2)$ such that $\zeta(x) = 1$ for any $x \in B_{3/2}$ and $0 \le \zeta(x) \le 1$ for any $x \in \mathbb{R}^N$. By using Lemma D.3(b), we get for $-4 < t_1 < t_2 < 0$

$$\int \zeta^2 |v_k|^2(t_2,x)\,\mathrm{d}x + \delta^\star \int_{t_1}^{t_2}\!\!\int |\nabla(\zeta v_k)|^2(s,x)\,\mathrm{d}x\,\mathrm{d}s \le \int \zeta^2 v_k^2(t_1,x)\,\mathrm{d}x + C(t_2-t_1) \tag{26}$$

for a certain constant $C > 0$. In particular, we have $(\zeta v_k)_{k\in\mathbb{N}}$ is bounded in $L^\infty(-4,0;L^2(B_2)) \cap L^2(-4,0;H^1(B_2))$. Going back to (25), since $\mu_k \ge 0$, $v_k \ge 0$, we observe that

$$\begin{aligned} 0 &\le \int_{t_1}^{t_2}\!\!\int_{B_{3/2}} \mu_k \,\mathrm{d}x\,\mathrm{d}s \le \int_{t_1}^{t_2}\!\!\int_{B_2} \zeta\mu_k \,\mathrm{d}x\,\mathrm{d}s \\ &\le \int_{B_2} \zeta v_k(t_1,x)\,\mathrm{d}x - \delta^\star \int_{t_1}^{t_2}\!\!\int_{B_2} \nabla v_k \cdot \nabla\zeta \,\mathrm{d}x\,\mathrm{d}s \\ &\le \|\zeta\|_{L^1} + 2\delta^\star \|\nabla v_k\|_{L^2(Q_2)} \|\nabla\zeta\|_{L^2(B_2)} \end{aligned}$$

is bounded uniformly with respect to k. Coming back to (25), we deduce that $(\partial_t v_k)_{k\in\mathbb{N}}$ is bounded in $\mathscr{M}^1((-4,0)\times B_{3/2}) + L^2(-4,0;H^{-1}(B_{3/2}))$. By virtue of the Aubin–Lions–Simon lemma [Simon 1987] (in fact we use the extended version [Moussa 2016, Theorem 1], which allows us to deal with measure-valued time derivatives), we conclude that $(v_k)_{k\in\mathbb{N}}$ is compact in $L^2((-4,0)\times B_{3/2})$. We can thus assume that v_k (possibly relabeling sequence) converges to some v in $L^2((-4,0)\times B_{3/2})$. The Bienaymé–Tchebyschev inequality yields

$$\operatorname{meas}\big(\{(t,x) \in ((-4,0)\times B_1) : |v_k(t,x) - v(t,x)| \ge \epsilon\}\big) \le \frac{\|v_k - v\|^2_{L^2((-4,0)\times B_1)}}{\epsilon^2} \xrightarrow{k\to\infty} 0$$

for any $\epsilon > 0$.

Let $(t,x)\in(-4,0)\times B_1$ be such that $\epsilon\le v(t,x)\le\frac{1}{2}-\epsilon$. Then we distinguish the following two cases: either $|v-v_k|(t,x)\ge\epsilon$ or $0\le v_k(t,x)=(v_k-v)(t,x)+v(t,x)\le|v-v_k|(t,x)+v(t,x)\le\frac{1}{2}$. It follows that

$$\begin{aligned}\operatorname{meas}\big(\{(t,x)\in Q_1\cup\widetilde{Q}:\epsilon\le v(t,x)\le\tfrac{1}{2}-\epsilon\}\big)&\le\operatorname{meas}(\{(t,x)\in Q_1\cup\widetilde{Q}:|v-v_k|(t,x)\ge\epsilon\})\\&\qquad+\underbrace{\operatorname{meas}\big(\{(t,x)\in Q_1\cup\widetilde{Q}:0\le v_k(t,x)\le\tfrac{1}{2}\}\big)}_{\operatorname{meas}(\mathscr{C}_k)}\\&\le\operatorname{meas}(\{(t,x)\in Q_1\cup\widetilde{Q}:|v-v_k|(t,x)\ge\epsilon\})+\frac{1}{k},\end{aligned}$$

by using (24). Letting k go to ∞ yields

$$\operatorname{meas}\big(\{(t,x)\in Q_1\cup\widetilde{Q}:\epsilon\le v(t,x)\le\tfrac{1}{2}-\epsilon\}\big)=0.$$

Since this property holds for any ϵ, the monotone convergence property leads to

$$\operatorname{meas}\big(\{(t,x)\in Q_1\cup\widetilde{Q}:0<v(t,x)<\tfrac{1}{2}\}\big)=0.$$

Therefore, we have

$$\text{for a.e. } t\in\big(-\tfrac{9}{4},0\big),\text{ either } v(t,x)=0\text{ or } v(t,x)\ge\tfrac{1}{2}\text{ in } B_1. \tag{27}$$

Similarly, let $(t,x)\in(-4,0)\times B_1$ be such that $v_k(t,x)=0$. We distinguish the following two cases: either $|v-v_k|(t,x)\ge\epsilon$ or $0\le v(t,x)=(v-v_k)(t,x)\le|v-v_k|(t,x)\le\epsilon$. Coming back to (24), we get

$$\begin{aligned}\tfrac{1}{2}\operatorname{meas}(\widetilde{Q})&\le\operatorname{meas}(\mathscr{B}_k)\\&\le\operatorname{meas}(\{(t,x)\in\widetilde{Q}:|v-v_k|(t,x)\ge\epsilon\})+\operatorname{meas}(\{(t,x)\in\widetilde{Q}:v(t,x)\le\epsilon\}).\end{aligned}$$

Letting k go to ∞ we obtain

$$\tfrac{1}{2}\operatorname{meas}(\widetilde{Q})\le\operatorname{meas}(\{(t,x)\in\widetilde{Q}:v(t,x)\le\epsilon\}).$$

By monotone convergence, as $\epsilon\to0$, we arrive at

$$\tfrac{1}{2}\operatorname{meas}(\widetilde{Q})\le\operatorname{meas}(\{(t,x)\in\widetilde{Q}:v(t,x)=0\}).$$

Consequently, we can find a nonnegligible set of times $s\in\big(-\frac{3}{2},-1\big)$ such that $v(s,x)=0$ holds for a.e. $x\in B_1$. Letting k go to ∞ in (25), we obtain $\partial_t v-\delta^\star\Delta v+\nu=0$ on $(-4,0)\times B_{3/2}$, with ν a nonnegative measure. Let $\zeta\in C_c^\infty(B_{3/2})$ be a nonnegative trial function such that $\zeta(x)=1$ for any $x\in B_1$. We apply Lemma D.3(b), and we obtain for a.e. $t\in(s,0)$,

$$\int_{B_1}v^2(t,x)\,\mathrm{d}x\le\int_{B_{3/2}}v^2(t,x)\zeta^2(x)\,\mathrm{d}x\le\int_{B_{3/2}}v^2(s,x)\zeta(x)\,\mathrm{d}x+C(t-s)=C(t-s),$$

where, owing to (27), we also know that the left-hand side is either null or larger than $\operatorname{meas}(B_1)/4$. We deduce that, actually, v vanishes on Q_1. We are going to show that it contradicts (24).

Indeed, let us consider $(t,x)\in Q_1$ such that $v_k(t,x)\ge\frac{1}{2}$. Then, for any $\epsilon>0$, either $|v-v_k|(t,x)\ge\epsilon$ or $v(t,x)=v_k(t,x)+(v-v_k)(t,x)\ge v_k(t,x)-|v-v_k|(t,x)\ge\frac{1}{2}-\epsilon$. With the first property in (24), it follows that

$$\eta \le \operatorname{meas}(\mathscr{A}_k) \le \operatorname{meas}(\{(t,x) \in Q_1 : |v - v_k|(t,x) \ge \epsilon\}) + \operatorname{meas}\big(\{(t,x) \in Q_1 : v(t,x) \ge \tfrac{1}{2} - \epsilon\}\big).$$

Letting k go to ∞ yields

$$\eta \le \operatorname{meas}\big(\{(t,x) \in Q_1 : v(t,x) \ge \tfrac{1}{2} - \epsilon\}\big).$$

Since this inequality holds for any $\epsilon > 0$, we conclude, by monotone convergence, that

$$\eta \le \operatorname{meas}\big(\{(t,x) \in Q_1 : v(t,x) \ge \tfrac{1}{2}\}\big)$$

holds, a contradiction. □

Proof of Proposition D.2. We consider $(t,x) \mapsto v(t,x)$ such that $-1 \le v(t,x) \le +1$,

$$\operatorname{meas}(\{(t,x) \in \widetilde{Q} : v(t,x) \le 0\}) \ge \mu \operatorname{meas}(\widetilde{Q}),$$

and v satisfies $\partial_t v - \delta^\star \Delta v \le 0$ in Q_2. The proof splits into two steps.

Step 1: For $k \in \mathbb{N}$, set

$$v_k(t,x) = 2^k \left(v(t,x) - \left(1 - \frac{1}{2^k}\right)\right).$$

We shall show that the integral

$$\iint_{Q_1} [v_k]_+^2 \,\mathrm{d}x\,\mathrm{d}t$$

can be made as small as we wish, by choosing k large enough. Observe that

$$v_k = 2^k(v-1) + 1 = 2v_{k-1} - 1,$$

which implies that $v_k \le 1$ and

$$\{(t,x) \in \widetilde{Q} : v(t,x) \le 0\} \subset \left\{(t,x) \in \widetilde{Q} : v(t,x) \le 1 - \frac{1}{2^k}\right\} = \{(t,x) \in \widetilde{Q} : v_k(t,x) \le 0\}.$$

Thus, by assumption on v,we have

$$\operatorname{meas}(\{(t,x) \in \widetilde{Q} : v_k(t,x) \le 0\}) \ge \operatorname{meas}(\{(t,x) \in \widetilde{Q} : v(t,x) \le 0\}) \ge \mu \operatorname{meas}(\widetilde{Q}).$$

Let us suppose that, for any $k \in \mathbb{N}$,

$$\iint_{Q_1} [v_k]_+^2 \,\mathrm{d}x\,\mathrm{d}t \ge \delta$$

holds for a certain $\delta > 0$. Since this integral is dominated by

$$\operatorname{meas}(\{(t,x) \in Q_1 : v_k(t,x) \ge 0\}) = \operatorname{meas}\big(\{(t,x) \in Q_1 : v_{k-1}(t,x) \ge \tfrac{1}{2}\}\big),$$

we infer

$$\operatorname{meas}\big(\{(t,x) \in Q_1 : v_{k-1}(t,x) \ge \tfrac{1}{2}\}\big) \ge \delta$$

independently of k. Applying Lemma D.4 yields

$$\operatorname{meas}\big(\{(t,x) \in Q_1 \cup \widetilde{Q} : 0 < v_{k-1}(t,x) < \tfrac{1}{2}\}\big) \ge \alpha,$$

still independently of k. It follows that

$$\begin{aligned}
&\text{meas}(\{(t,x) \in Q_1 \cup \widetilde{Q} : v_k(t,x) \le 0\}) \\
&\qquad = \text{meas}(\{(t,x) \in Q_1 \cup \widetilde{Q} : 2v_{k-1}(t,x) - 1 \le 0\}) \\
&\qquad = \text{meas}(\{(t,x) \in Q_1 \cup \widetilde{Q} : v_{k-1}(t,x) \le 0\}) + \text{meas}\big(\{(t,x) \in Q_1 \cup \widetilde{Q} : 0 < v_{k-1}(t,x) \le \tfrac{1}{2}\}\big) \\
&\qquad \ge \text{meas}(\{(t,x) \in Q_1 \cup \widetilde{Q} : v_{k-1}(t,x) \le 0\}) + \alpha.
\end{aligned}$$

Since $\text{meas}(\{(t,x) \in Q_1 \cup \widetilde{Q} : v_0(t,x) \le 0\}) \ge \text{meas}(\{(t,x) \in \widetilde{Q} : v_0(t,x) \le 0\}) \ge \mu\, \text{meas}(\widetilde{Q})$, this recursion formula leads to

$$\text{meas}(\{(t,x) \in Q_1 \cup \widetilde{Q} : v_k(t,x) \le 0\}) \ge \mu\, \text{meas}(\widetilde{Q}) + k\alpha.$$

However, this cannot occur for any k since the left-hand side is bounded by $\text{meas}(Q_2)$. We conclude that, given $\delta > 0$, there exists $k_\star \in \mathbb{N}$ such that

$$\iint_{Q_1} [v_{k_\star}]_+^2 \,\mathrm{d}x\,\mathrm{d}t \le \delta.$$

Step 2: The second step relies on De Giorgi's analysis. Let us set $w(t,x) = v_{k_\star}(t,x)$. We shall show that, provided δ is small enough (which means $k_\star$ large enough), $w(t,x) \le \frac{1}{2}$ on $Q_{1/2}$. To this end, let us set, for $\ell \in \mathbb{N}$,

$$m_\ell = \frac{1}{2}\left(1 - \frac{1}{2^\ell}\right), \quad w_\ell(t,x) = [w(t,x) - m_\ell]_+, \quad r_\ell = \frac{1}{2}\left(1 + \frac{1}{2^\ell}\right), \quad t_\ell = -r_\ell^2 = -\frac{1}{4}\left(1 + \frac{1}{2^\ell}\right)^2.$$

We are going to work in the domains $Q_{1/2} \subset Q_{r_\ell} \subset Q_1$, which shrink to $Q_{1/2}$ as $\ell \to \infty$; see Figure 2.

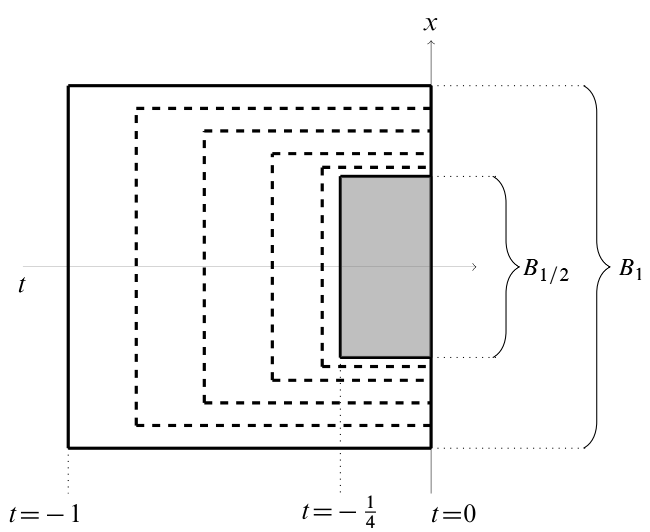

Figure 2. The domains Q_1, Q_{r_ℓ} and $Q_{1/2}$ (the gray box).

We consider a sequence of functions $\zeta_\ell \in C_c^\infty(B_{r_{\ell-1}})$ such that $0 \le \zeta_\ell(x) \le 1$ on $B_{r_{\ell-1}}$ and $\zeta_\ell(x) = 1$ on B_{r_ℓ}. We shall use the basic estimate

$$|\nabla \zeta_\ell(x)| \le C 2^\ell, \quad \frac{1}{t_\ell - t_{\ell-1}} \le C 2^{2\ell}.$$

We already know that $0 \le w_\ell(t,x) \le 1$, by definition. We can apply the energy estimate in Lemma D.3, which reads

$$\begin{aligned}\frac{1}{2}\int_{B_1} w_\ell^2(t,x)\zeta_\ell^2(x)\,\mathrm{d}x + \delta^\star \int_s^t\!\!\int_{B_1} |\nabla(\zeta_\ell w_\ell)|^2(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau \\ \le \frac{1}{2}\int_{B_1} w_\ell^2(s,x)\zeta_\ell^2(x)\,\mathrm{d}x + \delta^\star \int_s^t\!\!\int_{B_1} w_\ell^2 |\nabla \zeta_\ell|^2(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau \end{aligned} \tag{28}$$

for $-1 < s < t_\ell < t < 0$ (note that here we keep explicit the integral in the right-hand side that is roughly estimated by a constant in Lemma D.3). Averaging over $s \in (t_{\ell-1}, t_\ell)$ (and using the fact that the integral of a positive quantity over (s,t) is thus bounded below by the integral over (t_ℓ, t) and above by the integral over $(t_{\ell-1}, t)$) yields

$$\begin{aligned}\frac{1}{2}\int_{B_1} w_\ell^2(t,x)\zeta_\ell^2(x)\,\mathrm{d}x + \delta^\star \int_{t_\ell}^t\!\!\int_{B_1} |\nabla(\zeta_\ell w_\ell)|^2(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau \\ \le \Big(\frac{1}{2} + \delta^\star\Big) C 2^{2\ell} \int_{t_{\ell-1}}^0 \int_{\mathrm{supp}(\zeta_\ell)} |w_\ell|^2(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau.\end{aligned}$$

Let us set

$$\mathscr{U}_\ell = \int_{t_\ell}^0 \int_{B_{r_\ell}} |w_\ell|^2(t,x)\,\mathrm{d}x\,\mathrm{d}t,$$

$$\mathscr{E}_\ell = \sup_{t_\ell \le t \le 0} \int_{B_1} w_\ell^2(t,x)\zeta_\ell^2(x)\,\mathrm{d}x + \int_{t_\ell}^0 \int_{B_1} |\nabla(\zeta_\ell w_\ell)|^2(\tau,x)\,\mathrm{d}x\,\mathrm{d}\tau.$$

We wish to establish a nonlinear recursion for $\mathscr{U}_\ell$, which will allow us to justify that it tends to 0 as $\ell \to \infty$. On the one hand, since

$$w_\ell \le w_{\ell-1} \quad \text{and} \quad \mathrm{supp}(\zeta_\ell) \subset B_{\ell-1},$$

we note that (28) yields

$$\mathscr{E}_\ell \le \Big(2 + \frac{1}{\delta^\star}\Big)\Big(\frac{1}{2} + \delta^\star\Big) C 2^{2\ell} \mathscr{U}_{\ell-1}.$$

On the other hand, we observe that

$$\begin{aligned}\mathscr{U}_\ell &\le \int_{t_\ell}^0 \int_{B_{r_\ell}} |\zeta_\ell w_\ell|^2(t,x)\,\mathrm{d}x\,\mathrm{d}t \\ &\le \Big(\int_{t_\ell}^0 \int_{B_{r_\ell}} |\zeta_\ell w_\ell|^{2(N+2)/N}(t,x)\,\mathrm{d}x\,\mathrm{d}t\Big)^{N/(N+2)} \big(\mathrm{meas}(\{(t,x) \in (t_\ell, 0) \times B_{r_\ell} : \zeta_\ell w_\ell > 0\})\big)^{2/(N+2)},\end{aligned}$$

by using Hölder's inequality. Note that

$$w - m_{\ell-1} = w - m_\ell + \frac{1}{2^{\ell+1}},$$

which leads to

$$\begin{aligned}\text{meas}(\{(t,x)\in(t_\ell,0)\times B_{r_\ell}:\zeta_\ell w_\ell>0\}) &\le \text{meas}(\{(t,x)\in(t_{\ell-1},0)\times B_{r_{\ell-1}}:w_{\ell-1}>2^{-\ell-1}\})\\ &\le 2^{2\ell+2}\mathscr{U}_{\ell-1},\end{aligned}$$

by virtue of the Bienaymé–Tchebyschev inequality. Next, we use the Gagliardo–Nirenberg–Sobolev inequality, see [Nirenberg 1959, Theorem p. 125],

$$\left(\int_{B_{r_\ell}}|\zeta_\ell w_\ell|^{2N/(N-2)}(t,x)\,\mathrm{d}x\right)^{(N-2)/N}\le C_S\int_{B_{r_\ell}}|\nabla(\zeta_\ell w_\ell)|^2(t,x)\,\mathrm{d}x.$$

Mind that we have integrated with respect to the space variable only. We can write

$$\frac{N+2}{N}=\theta\frac{2N}{N-2}+2(1-\theta),\quad \theta=\frac{N-2}{N}\in(0,1),$$

so that

$$\begin{aligned}&\int_{t_\ell}^0\int_{B_{r_\ell}}|\zeta_\ell w_\ell|^{2(N+2)/N}(t,x)\,\mathrm{d}x\,\mathrm{d}t\\ &\qquad\le\int_{t_\ell}^0\left(\int_{B_{r_\ell}}|\zeta_\ell w_\ell|^{2N/(N-2)}(t,x)\,\mathrm{d}x\right)^{\theta}\underbrace{\left(\int_{B_{r_\ell}}|\zeta_\ell w_\ell|^2(t,x)\,\mathrm{d}x\right)^{1-\theta}}_{\le\mathscr{E}_\ell^{1-\theta}}\mathrm{d}t\\ &\qquad\le C_S^\theta\mathscr{E}_\ell^{2-\theta}.\end{aligned}$$

Therefore, gathering all this together, we obtain

$$\mathscr{U}_\ell\le\Lambda^\ell\mathscr{U}_{\ell-1}^{1+2/(N+2)}$$

for a certain constant $\Lambda>1$. Owing to Lemma A.1, we deduce that $\lim_{\ell\to\infty}\mathscr{U}_\ell=0$ provided $\mathscr{U}_0$ is small enough. The smallness condition on $\mathscr{U}_0$ is precisely ensured by the definition $w=v_{k_\star}$ coming from Step 1. Since

$$\frac{1}{|t_\ell|}\int_{t_\ell}^0\int_{B_{r_\ell}}|w_\ell|^2(t,x)\,\mathrm{d}x\,\mathrm{d}t\le\mathscr{U}_\ell,$$

we conclude, by applying Fatou's lemma, that

$$2\iint_{Q_{1/2}}\left[w-\tfrac{1}{2}\right]_+^2(t,x)\,\mathrm{d}x\,\mathrm{d}t\le\liminf_{\ell\to\infty}\frac{1}{t_\ell}\int_{t_\ell}^0\int_{B_{r_\ell}}|w_\ell|^2(t,x)\,\mathrm{d}x\,\mathrm{d}t=0$$

so that, finally, $w(t,x)\le\frac{1}{2}$ holds a.e. on $Q_{1/2}$.

Coming back to the change of unknown,

$$w(t,x)=v_{k_\star}(t,x)=2^{k_\star}\left(v(t,x)-\left(1-\frac{1}{2^{k_\star}}\right)\right)\le\frac{1}{2}$$

becomes

$$v(t,x)\le 1+\frac{1}{2^{k_\star+1}}-\frac{1}{2^{k_\star}}=1-\frac{1}{2^{k_\star+1}}<1.$$

□

Note

After the completion of this work, we learned about results in a similar direction by J. I. Kanel [1990]. This approach shares similar ideas and assumptions, but with different techniques; it has been recently revisited by P. Souplet [2018] to deal with problems endowed with Neumann boundary conditions and nonlinearities with a quadratic growth.

Acknowledgements

Goudon acknowledges the warm welcome of the Math Department of UT-Austin, where a couple of visits have made possible progress on this question.

References

[Alexandrof 1966] A. D. Aleksandrov, "Majorants of solutions of linear equations of order two", *Vestnik Leningrad. Univ.* **21**:1 (1966), 5–25. In Russian; translated in *Amer. Math. Soc. Trans.* **68**:2 (1968), 120–143. MR Zbl

[Alikakos 1979] N. D. Alikakos, "L^p bounds of solutions of reaction-diffusion equations", *Comm. Partial Differential Equations* **4**:8 (1979), 827–868. MR Zbl

[Alonso et al. 2016] R. Alonso, P. Amorim, and T. Goudon, "Analysis of a chemotaxis system modeling ant foraging", *Math. Models Methods Appl. Sci.* **26**:9 (2016), 1785–1824. MR Zbl

[Bakelman 1961] I. J. Bakelman, "On the theory of quasilinear elliptic equations", *Sibirsk. Mat. Ž.* **2** (1961), 179–186. In Russian. MR Zbl

[Brezis 1983] H. Brezis, *Analyse fonctionnelle: théorie et applications*, Masson, Paris, 1983. MR Zbl

[Cañizo et al. 2014] J. A. Cañizo, L. Desvillettes, and K. Fellner, "Improved duality estimates and applications to reaction-diffusion equations", *Comm. Partial Differential Equations* **39**:6 (2014), 1185–1204. MR Zbl

[Caffarelli and Vasseur 2010] L. A. Caffarelli and A. Vasseur, "Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation", *Ann. of Math.* (2) **171**:3 (2010), 1903–1930. MR Zbl

[Caputo and Vasseur 2009] M. C. Caputo and A. Vasseur, "Global regularity of solutions to systems of reaction-diffusion with sub-quadratic growth in any dimension", *Comm. Partial Differential Equations* **34**:10 (2009), 1228–1250. MR Zbl

[Chainais-Hillairet et al. 2017] C. Chainais-Hillairet, B. Merlet, and A. F. Vasseur, "Positive lower bound for the numerical solution of a convection-diffusion equation", pp. 331–339 in *Finite volumes for complex applications, VIII: Methods and theoretical aspects* (Lille, 2017), edited by C. Cancès and P. Omnes, Springer Proc. Math. Stat. **199**, Springer, 2017. MR Zbl

[De Giorgi 1957] E. De Giorgi, "Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari", *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat.* (3) **3** (1957), 25–43. MR Zbl

[Desvillettes and Fellner 2006] L. Desvillettes and K. Fellner, "Exponential decay toward equilibrium via entropy methods for reaction-diffusion equations", *J. Math. Anal. Appl.* **319**:1 (2006), 157–176. MR Zbl

[Desvillettes et al. 2007] L. Desvillettes, K. Fellner, M. Pierre, and J. Vovelle, "Global existence for quadratic systems of reaction-diffusion", *Adv. Nonlinear Stud.* **7**:3 (2007), 491–511. MR Zbl

[Evans 1998] L. C. Evans, *Partial differential equations*, Graduate Studies in Math. **19**, Amer. Math. Soc., Providence, RI, 1998. MR Zbl

[Fabes and Stroock 1984] E. B. Fabes and D. W. Stroock, "The L^p-integrability of Green's functions and fundamental solutions for elliptic and parabolic equations", *Duke Math. J.* **51**:4 (1984), 997–1016. MR Zbl

[Fellner et al. 2016] K. Fellner, E. Latos, and T. Suzuki, "Global classical solutions for mass-conserving, (super-)quadratic reaction-diffusion systems in three and higher space dimensions", *Discrete Contin. Dyn. Syst. Ser. B* **21**:10 (2016), 3441–3462. MR Zbl

[Fischer 2015] J. Fischer, "Global existence of renormalized solutions to entropy-dissipating reaction-diffusion systems", *Arch. Ration. Mech. Anal.* **218**:1 (2015), 553–587. MR Zbl

[Goudon and Urrutia 2016] T. Goudon and L. Urrutia, "Analysis of kinetic and macroscopic models of pursuit-evasion dynamics", *Commun. Math. Sci.* **14**:8 (2016), 2253–2286. MR Zbl

[Goudon and Vasseur 2010] T. Goudon and A. Vasseur, "Regularity analysis for systems of reaction-diffusion equations", *Ann. Sci. Éc. Norm. Supér.* (4) **43**:1 (2010), 117–142. MR Zbl

[Kanel 1990] Y. I. Kanel, "Solvability in the large of a system of reaction-diffusion equations with the balance condition", *Diff. Uravneniya* **26**:3 (1990), 448–458. In Russian; translated in *Diff. Equations* **26**:3 (1990), 331–339. MR Zbl

[Krylov 1976] N. V. Krylov, "Sequences of convex functions, and estimates of the maximum of the solution of a parabolic equation", *Sibirsk. Mat. Ž.* **17**:2 (1976), 290–303. In Russian; translated in *Siberian Math. J.* **17**:2 (1976), 226–236. MR Zbl

[Krylov and Safonov 1979] N. V. Krylov and M. V. Safonov, "An estimate for the probability of a diffusion process hitting a set of positive measure", *Dokl. Akad. Nauk SSSR* **245**:1 (1979), 18–20. In Russian. MR

[Krylov and Safonov 1980] N. V. Krylov and M. V. Safonov, "A property of the solutions of parabolic equations with measurable coefficients", *Izv. Akad. Nauk SSSR Ser. Mat.* **44**:1 (1980), 161–175. In Russian; translated in *Math. USSR-Izv.* **16**:1 (1981), 151–164. MR Zbl

[Moussa 2016] A. Moussa, "Some variants of the classical Aubin–Lions lemma", *J. Evol. Equ.* **16**:1 (2016), 65–93. MR Zbl

[Nirenberg 1959] L. Nirenberg, "On elliptic partial differential equations", *Ann. Scuola Norm. Sup. Pisa* (3) **13**:2 (1959), 115–162. MR Zbl

[Pierre 2010] M. Pierre, "Global existence in reaction-diffusion systems with control of mass: a survey", *Milan J. Math.* **78**:2 (2010), 417–455. MR Zbl

[Pierre and Schmitt 1997] M. Pierre and D. Schmitt, "Blowup in reaction-diffusion systems with dissipation of mass", *SIAM J. Math. Anal.* **28**:2 (1997), 259–269. MR Zbl

[Pierre et al. 2017a] M. Pierre, T. Suzuki, and H. Umakoshi, "Global-in-time behavior of weak solutions to reaction-diffusion systems with inhomogeneous Dirichlet boundary condition", *Nonlinear Anal.* **159** (2017), 393–407. MR Zbl

[Pierre et al. 2017b] M. Pierre, T. Suzuki, and R. Zou, "Asymptotic behavior of solutions to chemical reaction-diffusion systems", *J. Math. Anal. Appl.* **450**:1 (2017), 152–168. MR Zbl

[Pucci 1966] C. Pucci, "Limitazioni per soluzioni di equazioni ellittiche", *Ann. Mat. Pura Appl.* (4) **74**:1 (1966), 15–30. MR Zbl

[Simon 1987] J. Simon, "Compact sets in the space $L^p(0, T; B)$", *Ann. Mat. Pura Appl.* (4) **146** (1987), 65–96. MR Zbl

[Souplet 2018] P. Souplet, "Global existence for reaction-diffusion systems with dissipation of mass and quadratic growth", *J. Evol. Equ.* **18**:4 (2018), 1713–1720. MR Zbl

[Tso 1985] K. Tso, "On an Aleksandrov–Bakel'man type maximum principle for second-order parabolic equations", *Comm. Partial Differential Equations* **10**:5 (1985), 543–553. MR Zbl

[Vasseur 2007] A. F. Vasseur, "A new proof of partial regularity of solutions to Navier–Stokes equations", *Nonlinear Diff. Equations Appl.* **14**:5-6 (2007), 753–785. MR Zbl

[Vasseur 2010] A. Vasseur, "Higher derivatives estimate for the 3D Navier–Stokes equation", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **27**:5 (2010), 1189–1204. MR Zbl

[Weissler 1985] F. B. Weissler, "An L^∞ blow-up estimate for a nonlinear heat equation", *Comm. Pure Appl. Math.* **38**:3 (1985), 291–295. MR Zbl

Received 28 Nov 2017. Revised 22 Aug 2018. Accepted 25 Oct 2018.

M. CRISTINA CAPUTO: caputo@math.utexas.edu
Department of Mathematics, University of Texas at Austin, Austin, TX, United States

THIERRY GOUDON: thierry.goudon@inria.fr
Université Côte d'Azur, Inria, CNRS, LJAD, Nice, France

ALEXIS F. VASSEUR: vasseur@math.utexas.edu
Department of Mathematics, University of Texas at Austin, Austin, TX, United States

ANALYSIS AND PDE
Vol. 12, No. 7, 2019
dx.doi.org/10.2140/apde.2019.12.1805

SPACELIKE RADIAL GRAPHS OF PRESCRIBED MEAN CURVATURE IN THE LORENTZ–MINKOWSKI SPACE

DENIS BONHEURE AND ALESSANDRO IACOPETTI

We investigate the existence and uniqueness of spacelike radial graphs of prescribed mean curvature in the Lorentz–Minkowski space $\mathbb{L}^{n+1}$, for $n \geq 2$, spanning a given boundary datum lying on the hyperbolic space $\mathbb{H}^n$.

1. Introduction

A radial graph is a hypersurface Σ such that each ray emanating from the origin intersects Σ once at most. In the euclidean context the problem of finding radial graphs of prescribed mean curvature has been extensively studied over the years. In the first paper on the subject, Radó [1932] proved that for any given Jordan curve $\Gamma \subset \mathbb{R}^3$, with one-to-one radial projection onto a convex subset of the unit sphere $\mathbb{S}^2$, there exists a minimal graph spanning Γ. Later, Tausch [1981] proved that area-minimizing disk-type hypersurfaces spanning a boundary datum Γ which can be expressed as a radial graph over $\partial\Omega$, where $\Omega \subset \mathbb{S}^n$ is a convex subset, have a local representation as a radial graph. The case of variable mean curvature was investigated by Serrin [1969], and a recent result of radial representation for H-surfaces in cones was given in [Caldiroli and Iacopetti 2016]. Treibergs and Wei [1983] studied the case of closed hypersurfaces, i.e., compact hypersurfaces without boundary. Lopez [2003] and de Lira [2002] studied the case of radial graphs of constant mean curvature.

The Lorentz–Minkowski space, denoted by $\mathbb{L}^{n+1}$, is defined as the vector space $\mathbb{R}^{n+1}$ equipped with the symmetric bilinear form

$$\langle x, y\rangle := x_1 y_1 + \cdots + x_n y_n - x_{n+1} y_{n+1},$$

where $x = (x_1, \dots, x_{n+1})$, $y = (y_1, \dots, y_{n+1}) \in \mathbb{R}^{n+1}$. The bilinear form $\langle \cdot, \cdot \rangle$ is a nondegenerate bilinear form of index 1, see [Spivak 1975, Section A], where the index of a bilinear form on a real vector space is defined as the largest dimension of a negative definite subspace. The modulus of $v \in \mathbb{L}^{n+1}$ is defined as $|v| := \sqrt{|\langle v, v\rangle|}$.

The interest in finding spacelike hypersurfaces of prescribed mean curvature in the Lorentz–Minkowski space comes from the theory of relativity, in which maximal and constant-mean-curvature spacelike

Research partially supported by the project ERC Advanced Grant 2013 no. 339958 Complex Patterns for Strongly Interacting Dynamical Systems COMPAT, by Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM) by FNRS (PDR T.1110.14F and MIS F.4508.14) and by ARC AUWB-2012-12/17-ULB1- IAPAS.
MSC2010: primary 53A10; secondary 35J66, 53C50.
Keywords: prescribed mean curvature, Plateau's problem, H-surfaces, Lorentz–Minkowski space.

hypersurfaces play an important role, see [Bartnik and Simon 1982], where spacelike means that the restriction of the Lorentz metric to the tangent plane, at every point, is positive definite. In the literature, several result are available for spacelike vertical graphs, i.e., hypersurfaces which are expressed as a cartesian graph. Entire maximal spacelike hypersurfaces were studied by Cheng and Yau [1976] and later Treibergs [1982] tackled the general case of entire spacelike hypersurfaces of constant mean curvature. The Dirichlet problem for spacelike vertical graphs in $\mathbb{L}^{n+1}$ was solved by Bartnik and Simon [1982], and Gerhardt [1983] extended those results to the case of vertical graphs contained in Lorentzian manifolds which can be expressed as a product of a Riemannian manifold times an interval. Bayard [2003] studied the more general problem of prescribed scalar curvature. On the contrary, for radial graphs, to our knowledge, the only available result concerns entire spacelike hypersurfaces with prescribed scalar curvature which are asymptotic to the light-cone; see [Bayard and Delanoë 2009].

The geometry of Lorentz–Minkowski spaces plays an in important role in the setting of the problem. A first relevant fact is that there cannot exist spacelike closed hypersurfaces (see Proposition 2.5, or [López 2014] for the case of surfaces in $\mathbb{L}^3$). Therefore $\mathbb{S}^n$-type surfaces are ruled out, and the model hypersurface in $\mathbb{L}^{n+1}$ for describing spacelike radial graphs is the hyperbolic space $\mathbb{H}^n$ (see Definition 2.6). Another important feature of Lorentz–Minkowski spaces is that, given a domain, there exist spacelike hypersurfaces of arbitrarily large (in modulus) mean curvature, see [López 2013], while in the euclidean context this is not true in general. This fact will be crucial in our paper to construct barriers.

We state now the problem. Let Ω be a smooth bounded domain of $\mathbb{H}^n$. For $u:\overline{\Omega}\to\mathbb{R}$, we define the associated radial graph over Ω as the set

$$\Sigma(u) := \{p = e^{u(q)}q \in \mathbb{L}^{n+1} : q \in \overline{\Omega}\}.$$

Let $\mathcal{C}_{\overline{\Omega}}$ be the cone spanned by $\overline{\Omega}$ (minus the origin), i.e., $\mathcal{C}_{\overline{\Omega}} := \{p = \rho q \in \mathbb{L}^{n+1} : q \in \overline{\Omega},\ \rho > 0\}$, and let $H : \mathcal{C}_{\overline{\Omega}} \to \mathbb{R}$.

Definition 1.1. A H-bump (over Ω) is a radial graph Σ whose boundary coincides with $\partial\Omega$ and such that the mean curvature of Σ at every (interior) point equals H.

The Dirichlet problem for spacelike H-bumps is given by

$$\begin{cases} \sum_{i,j=1}^{n}((1-|\nabla u|^2)\delta_{ij}+u_iu_j)u_{ij} = n(1-|\nabla u|^2) - n(1-|\nabla u|^2)^{3/2}e^u H(e^u q) & \text{in } \Omega,\\ |\nabla u|<1 & \text{in } \Omega,\\ u=0 & \text{on } \partial\Omega, \end{cases} \tag{1-1}$$

where u_i, u_{ij} are the covariant derivatives of u, ∇u is the gradient with respect to the Levi-Civita connection of $(\mathbb{H}^n, g)$ (see Section 3), and $g = dx_1\otimes dx_1 + \cdots + dx_n \otimes dx_n - dx_{n+1}\otimes dx_{n+1}$ is the induced Riemannian metric on $\mathbb{H}^n$ (see Section 2).

Definition 1.2. Let $0 < r_1 \le 1 \le r_2$, with $r_1 \ne r_2$. The hyperbolic conical cap of radii r_1, r_2 spanned by $\overline{\Omega}$ is the set

$$\mathcal{C}_{\overline{\Omega}}(r_1, r_2) := \{p = \rho q \in \mathbb{L}^{n+1} : q \in \overline{\Omega},\ r_1 \le \rho \le r_2\}.$$

The main result of our paper is the following existence theorem.

Theorem 1.3. *Let $\alpha \in (0,1)$, $0 < r_1 \le 1 \le r_2$, with $r_1 \neq r_2$. Assume Ω is a bounded domain of $\mathbb{H}^n$ of class $C^{3,\alpha}$ that satisfies a uniform exterior geodesic ball condition. If $H \in C^{1,\alpha}(C_{\overline{\Omega}}(r_1, r_2))$ is positive and satisfies*

(i) *$H(r_1 q) > r_1^{-1}$ and $H(r_2 q) < r_2^{-1}$ for any $q \in \overline{\Omega}$,*

(ii) *$(\partial/\partial\lambda)(\lambda H(\lambda q)) \le 0$ for all $q \in \overline{\Omega}$, $\lambda \in [r_1, r_2]$,*

then there exists a unique solution of problem (1-1) *whose associated radial graph is contained in $\mathcal{C}_{\overline{\Omega}}(r_1, r_2)$.*

Let Ω, r_1, r_2 be in the statement of Theorem 1.3. Let $m \ge 1$, let $\omega : \overline{\Omega} \to \mathbb{R}^+$ be a smooth positive function such that $r_1^{m-1} < \omega < r_2^{m-1}$ and let $H_{m,\omega} : \mathcal{C}_{\overline{\Omega}}(r_1, r_2) \to \mathbb{R}^+$, defined by

$$H_{m,\omega}(x) := \frac{\omega(x/|x|)}{|x|^m}. \tag{1-2}$$

One easily verifies that $H_{m,\omega}$ satisfies the hypotheses (i) and (ii) of Theorem 1.3. In particular, this shows the existence of spacelike radial graphs of prescribed mean curvature even for nonhomogeneous functions H, a case which is not contemplated for instance in [Bayard and Delanoë 2009], where the k-th scalar curvature is prescribed just on $\mathbb{H}^n$.

We remark that (1-1) can be put in divergence form, namely

$$\begin{cases} -\operatorname{div}_{\mathbb{H}^n}(\nabla u/\sqrt{1-|\nabla u|^2}) + n/\sqrt{1-|\nabla u|^2} = n e^u H(e^u q) & \text{in } \Omega, \\ |\nabla u| < 1 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{1-3}$$

where $\operatorname{div}_{\mathbb{H}^n}$ denotes the divergence operator for $(\mathbb{H}^n, g)$. The principal part of this operator appears in the Born–Infeld theory of electromagnetism [1934], which is a particular example of what is usually known as a nonlinear electrodynamics. We therefore stress that Theorem 1.3 provides existence and uniqueness of solutions for some specific Born–Infeld equations in which appear nontrivial nonlinearities involving both the gradient and the function; see also [Bonheure et al. 2016; Bonheure and Iacopetti 2019].

The proof of Theorem 1.3 relies on the combination of several tools. For the existence, we apply a variant of the classical Leray–Schauder fixed point theorem due to Potter [1972]. To this aim, we make use of suitable comparison theorems and we prove fine a priori estimates for the solutions and their gradient. Regarding uniqueness, we take advantage of the Hopf maximum principle as in the version stated by Pucci and Serrin [2004].

We point out that the uniform exterior geodesic ball condition allows us to construct barriers for the gradient of the solutions at the boundary. Such construction strongly depends on the shape of the mean curvature operator for spacelike hypersurfaces in the Lorentz–Minkowski space, and we remark that Theorem 1.3 grants existence of spacelike radial graphs over arbitrarily large and even nonconvex domains of $\mathbb{H}^n$. We note that it is not possible to mimic this construction in the euclidean framework, and in fact the problem of finding radial graphs over proper (possibly nonconvex) domains of $\mathbb{S}^n$ which are not contained in a hemisphere is still open.

Concerning global a priori estimates for the gradient, which is the key step in the proof, we derive a quite complex technical result, see Proposition 8.1, which is inspired from [Gerhardt 1983] and is based on the introduction of an ad hoc differential operator, Stampacchia's truncation method and fine estimates of the L^p-norm of the quantity $\nu(u) = 1/\sqrt{1-|\nabla u|^2}$.

In this paper we also introduce a new definition of admissible couple (Ω, H) and triple (Ω, H, θ), see Definition 4.4, where $\theta \in (0,1)$. This notion of admissibility is very general and works even for nonsmooth domains and just for continuous functions H. However, given a couple (Ω, H), it is not easy in general to verify whether it is admissible or not. In Section 4 we provide trivial examples of admissible couples and in Proposition 4.7 we exhibit a class of functions H such that (Ω, H) is admissible whenever Ω satisfies a uniform exterior geodesic condition. Using the notion of admissible couple, we can extend Theorem 1.3 to a wider class of domains and mean curvature functions.

Theorem 1.4. *Let $\alpha \in (0,1)$, $0 < r_1 \le 1 \le r_2$, with $r_1 \neq r_2$. Assume that Ω is a bounded domain of $\mathbb{H}^n$ of class $C^{3,\alpha}$ and $H \in C^{1,\alpha}(\mathcal{C}_{\overline{\Omega}}(r_1, r_2))$ satisfies conditions* (i) *and* (ii) *of Theorem 1.3. Assume that (Ω, H) is admissible. Then there exists a unique solution of problem* (1-1) *whose associated radial graph is contained in $\mathcal{C}_{\overline{\Omega}}(r_1, r_2)$.*

A further existence result for problem (1-1), under more restrictive assumptions, is as follows.

Theorem 1.5. *Let $\alpha \in (0,1)$ and Ω be a bounded domain of $\mathbb{H}^n$ of class $C^{3,\alpha}$. Assume $\theta \in (0,1)$, $0 < r_1 \le 1 \le r_2$, with $r_1 \neq r_2$, and $H \in C^{1,\alpha}(\mathcal{C}_{\overline{\Omega}}(r_1, r_2))$ satisfies*

(a) *$H(r_1 q) > r_1^{-1}$ and $H(r_2 q) < r_2^{-1}$ for any $q \in \overline{\Omega}$,*

(b) *$(\partial/\partial\lambda)(\lambda H(\lambda q)) < -1/(r_1(\theta - \theta^2/4)^{1/2})$ for all $q \in \overline{\Omega}$, $\lambda \in [r_1, r_2]$,*

(c) *$\|\nabla_0^T H(x)\|_{n+1} < (1-\theta)/(n^{3/2} r_2^2)$ for all $x \in \mathcal{C}_{\overline{\Omega}}(r_1, r_2)$, where $\nabla_0^T H$ is the euclidean tangential component of $\nabla_0 H(x)$ on $T_{x/|x|}\mathbb{H}^n$ (see Definition 6.2), $\nabla_0 H$ is the gradient of H with respect to the euclidean flat metric, and $\|\cdot\|_{n+1}$ is the euclidean norm in $\mathbb{R}^{n+1}$.*

Assume at last that (Ω, H, θ) is admissible according to Definitions 4.4 and 4.10. Then there exists a unique spacelike H-bump contained in $\mathcal{C}_{\overline{\Omega}}(r_1, r_2)$.

We mention this result because the proof quite differs from that of Theorem 1.4 and better shows the differences and difficulties with respect to the euclidean case. The proof is this time based on the classical Leray–Schauder theorem; see for instance [Gilbarg and Trudinger 1977, Theorem 11.3]. The first step is to solve a suitable regularized equation associated to (1-1); see (4-2) and Theorem 5.1. The idea of solving such a regularized equation is taken from [Treibergs 1982], where the author constructs barriers for the gradient at the boundary. The way back to the original Dirichlet problem then uses a gradient maximum principle [Treibergs and Wei 1983, Proposition 6]. In contrast with [Treibergs 1982], we deal here with equations which do not satisfy, in general, a gradient maximum principle [Gilbarg and Trudinger 1977, Theorem 15.1]. In fact, in our case, when passing to local coordinates, we see that the regularized operator associated to (1-1) does not satisfy, in general [loc. cit., condition (15.11)], and the principal part depends both on the gradient and on the domain variables. We refer to Lemma 4.1 below for more details. In order to overcome this difficulty, and eventually deduce a global a priori C^1 estimate, we perform the regularization in a proper way. We then use the admissibility condition to control the gradient at the boundary, whereas we use two different strategies, see Lemma 4.12, for the interior estimate. The first one which is based on the properties of harmonic functions, works only in dimension 2. The other proof works in any dimension and is based on the global gradient bound given by [loc. cit., Theorem 15.2].

Finally, in the spirit of [Treibergs and Wei 1983], we prove a new kind of interior gradient estimate, see Proposition 6.4, so that, under the hypotheses of Theorem 1.5, the solution of the regularized problem is a solution of (1-1). It is important to note that, in contrast to [loc. cit.], since $\mathbb{H}^n$ has negative Ricci curvature and since we deal with hypersurfaces with boundary, the mere gradient estimate of Proposition 6.4 is not sufficient for getting a global a priori C^1-estimate. We refer to Remark 6.5 for more details.

When Ω satisfies a uniform exterior geodesic condition, thanks to Proposition 4.7, Remark 4.8 and Remark 4.9, it is possible to show that the functions given by (1-2) satisfy the hypotheses of Theorem 1.5 for suitable choices of r_1, r_2, m, for ω close to 1 (in the C^1-topology), and for some $\theta_* \in (0,1)$.

As a future goal, it would be natural to investigate if it is possible to remove the monotonicity assumption on H and to extend Theorem 1.3 also to sign-changing mean curvature functions.

The outline of the paper is the following. In Section 2, we fix the notation and we collect some known facts which are useful in the remainder of the paper. In Section 3, we derive the equation for spacelike H-bumps and in Section 4 we prove Proposition 4.7 and some a priori estimates. Section 5 is dedicated to the proof of existence and uniqueness of solutions for the regularized Dirichlet problem associated to problem (1-1). In Section 6, we work out an interior gradient estimate, namely Proposition 6.4, and in Section 7 we prove Theorem 1.5. In Section 8, we prove a global a priori estimate for the gradient. We finally prove Theorems 1.3 and 1.4 in Section 9.

2. Notation and preliminary results

Let $n \geq 2$; we denote by $\mathbb{L}^{n+1}$ the $(n+1)$-dimensional Lorentz–Minkowski space, which is $\mathbb{R}^{n+1}$ equipped with the symmetric bilinear form

$$\langle x, y\rangle := x_1 y_1 + \cdots + x_n y_n - x_{n+1} y_{n+1}.$$

We classify the vectors of $\mathbb{L}^{n+1}$ in three types.

Definition 2.1. A vector $v \in \mathbb{L}^{n+1}$ is said to be

- spacelike if $\langle v, v\rangle > 0$ or $v = 0$;
- timelike if $\langle v, v\rangle < 0$;
- lightlike if $\langle v, v\rangle = 0$ and $v \neq 0$.

The modulus of $v \in \mathbb{L}^{n+1}$ is defined as $|v| := \sqrt{|\langle v, v\rangle|}$. We also denote by $(x, y)_{n+1} = x_1 y_1 + \cdots + x_{n+1} y_{n+1}$ the euclidean scalar product, and by $\|x\|_{n+1} = \sqrt{x_1^2 + \cdots + x_{n+1}^2}$ the euclidean norm in $\mathbb{R}^{n+1}$. Given a vector subspace V of $\mathbb{L}^{n+1}$, we consider the induced metric $\langle \cdot, \cdot\rangle_V$ defined in the natural way

$$\langle v, w\rangle_V := \langle v, w\rangle, \ v, w \in V.$$

According to Definition 2.1 we classify the subspaces of $\mathbb{L}^{n+1}$ as follows.

Definition 2.2. A vector subspace V of $\mathbb{L}^{n+1}$ is said to be

- spacelike if the induced metric is positive definite;
- timelike if the induced metric has index 1;
- lightlike if the induced metric is degenerate.

In this paper, we deal only with hypersurfaces in $\mathbb{L}^{n+1}$, and thus we identify the tangent space of $M \subset \mathbb{L}^{n+1}$ at $p \in M$, denoted by $T_p M$, with a vector subspace of dimension n in $\mathbb{L}^{n+1}$. In particular, by abuse of notation, if $\phi : U \to M$, where U is an open subset of $\mathbb{R}^n$, is a local parametrization, we still use the symbol ∂_i to denote the vector $\partial\phi/\partial x_i$.

Definition 2.3. Let $M \subset \mathbb{L}^{n+1}$ be a hypersurface. We say that M is spacelike (resp. timelike, lightlike) if, for any $p \in M$, the vector subspace $T_p M$ is spacelike (resp. timelike, lightlike). We say that M is a nondegenerate hypersurface if M is spacelike or timelike.

Definition 2.4. A timelike vector $v \in \mathbb{L}^{n+1}$ is said to be future-oriented if $\langle v, E_{n+1}\rangle < 0$ and past-oriented if $\langle v, E_{n+1}\rangle > 0$, where $E_{n+1} := (0, \dots, 0, 1)$.

We observe that for a spacelike (resp. timelike) surface M and $p \in M$, we have the decomposition $\mathbb{L}^{n+1} = T_p M \oplus (T_p M)^\perp$, where $(T_p M)^\perp$ is a timelike (resp. spacelike) subspace of dimension 1; see [López 2014]. A Gauss map is a differentiable map $N : M \to \mathbb{L}^{n+1}$ such that $|N(p)| = 1$ and $N(p) \in (T_p M)^\perp$ for all $p \in M$. If M is spacelike, the Gauss map pointing to the future is a map $N : M \to \mathbb{H}^n$.

We recall now a result which is simple but crucial because it marks a relevant difference between the euclidean geometry and the geometry of Lorentz–Minkowski spaces.

Proposition 2.5. *Let $M \subset \mathbb{L}^{n+1}$ be a compact spacelike, timelike or lightlike hypersurface. Then $\partial M \neq \varnothing$.*

Proof. Assume that $\partial M = \varnothing$ and that M is spacelike (resp. timelike or lightlike). Let $a \in \mathbb{L}^{n+1}$ be a spacelike (resp. timelike) vector. Since M is compact, there exists a minimum (or a maximum) $p_0 \in M$ for the function $f(p) = \langle p, a\rangle$. Since $\partial M = \varnothing$, we know p_0 is a critical point of the function f and thus $\langle v, a\rangle = 0$ for all $v \in T_p M$. Hence $a \in (T_p M)^\perp$, but this gives a contradiction because $(T_p M)^\perp$ is timelike (resp. spacelike or lightlike). □

In other words, the previous result tells us that a closed hypersurface (i.e., compact without boundary) must be degenerate (see Definition 2.3). Therefore closed surfaces are not relevant in the Lorentz–Minkowski space, and this is deeply in contrast to euclidean geometry. For the sake of completeness, we also point out that Proposition 2.5, as well the previous definitions, can be extended to general hypersurfaces; see, e.g., [López 2014, Section 3].

Definition 2.6. The hyperbolic space of center $p_0 \in \mathbb{L}^{n+1}$ and radius $r > 0$ is the hypersurface defined by

$$\mathbb{H}^n(p_0, r) := \{p \in \mathbb{L}^{n+1} : \langle p - p_0, p - p_0\rangle = -r^2,\ \langle p - p_0, E_{n+1}\rangle < 0\},$$

where $E_{n+1} = (0, \dots, 0, 1)$.

From the euclidean point of view, this hypersurface is the "upper sheet" of a hyperboloid of two sheets.

Remark 2.7. The hyperbolic space is a spacelike hypersurface; see [López 2014; Spivak 1975]. In fact, let $v \in T_p\mathbb{H}^n(p_0, r)$ and let $\sigma = \sigma(s)$ be a curve in $\mathbb{H}^n(p_0, r)$ such that $\sigma'(0) = v$. Then, differentiating with respect to s the relation $\langle \sigma(s) - p_0, \sigma(s) - p_0 \rangle = -r^2$ at $s = 0$, we obtain $\langle v, p - p_0 \rangle = 0$. This implies $T_p\mathbb{H}^n(p_0, r) = \mathrm{Span}\{p - p_0\}^{\perp}$. Since $p - p_0$ is a timelike vector, it follows that $\mathbb{H}^n(p_0, r)$ is a spacelike hypersurface. Moreover $N(p) = (p - p_0)/r$ is a Gauss map.

When p_0 is the origin of $\mathbb{L}^{n+1}$, and $r = 1$, the hyperbolic space is denoted by $\mathbb{H}^n$; that is,

$$\mathbb{H}^n := \{(x_1, \dots, x_{n+1}) \in \mathbb{L}^{n+1} : x_1^2 + \cdots + x_n^2 - x_{n+1}^2 = -1,\ x_{n+1} > 0\}.$$

In view of the previous remark, for any $p \in \mathbb{H}^n$, the induced metric on $T_p\mathbb{H}^n$ is positive definite, and hence the tensor $g = dx_1 \otimes dx_1 + \cdots + dx_n \otimes dx_n - dx_{n+1} \otimes dx_{n+1}$ is a Riemannian metric for $\mathbb{H}^n$. Another model for $\mathbb{H}^n$ is the Poincaré model in the unit disk $\mathbb{B}^n := \{y \in \mathbb{R}^n : \|y\|_n < 1\}$, where $\|\cdot\|_n$ is the euclidean norm in $\mathbb{R}^n$. The hyperbolic metric in $\mathbb{B}^n$ is defined by

$$\tilde{g} = \frac{4}{(1 - \|y\|_n^2)^2} \sum_{i=1}^{n} dy_i \otimes dy_i,$$

which is conformally equivalent to the flat metric in $\mathbb{B}^n$. The isometry between $(\mathbb{H}^n, g)$ and $(\mathbb{B}^n, \tilde{g})$ is given by the map $F : \mathbb{H}^n \to \mathbb{B}^n$ defined by

$$F(x) := x_0 - \frac{2(x - x_0)}{\langle x - x_0, x - x_0 \rangle} = \left(\frac{x_1}{1 + x_{n+1}}, \dots, \frac{x_n}{1 + x_{n+1}} \right), \tag{2-1}$$

where $x_0 = (0, \dots, 0, -1) \in \mathbb{R}^{n+1}$; see [Lee 1997, Proposition 3.5]. The map F is also known as hyperbolic stereographic projection, and from a geometrical point of view, F sends a point $x \in \mathbb{H}^n$ to the intersection between the line joining x and x_0 with the hyperplane $\{y \in \mathbb{R}^{n+1} : y_{n+1} = 0\}$.

We conclude this section by recalling a variant of the Leray–Schauder fixed point theorem which will be used in the proof of Theorem 1.3.

Theorem 2.8 (A. J. B. Potter [1972]). *Let X be a locally convex linear Hausdorff topological space and U a closed convex subset of X such that the zero element of X is contained in the interior of U. Let $T : [0, 1] \times U \to X$ be a continuous map such that $T([0, 1] \times U)$ is relatively compact in X. Assume that*

(a) *$T(t, x) \neq x$ for all $x \in \partial U$ and $t \in [0, 1]$;*

(b) *$T(0 \times \partial U) \subset U$.*

Then, there is an element $\bar{x}$ of U such that $\bar{x} = T(1, \bar{x})$.

3. Derivation of the equation

Let Ω be a proper smooth bounded domain of the hyperbolic space $\mathbb{H}^n$. Let us denote by $\mathcal{T}(\Omega)$ the space of tangent vector fields to Ω and denote by ∇^0 the Levi-Civita connection of $\mathbb{L}^{n+1}$. We recall that ∇^0 coincides with the flat connection of $\mathbb{R}^{n+1}$, and we denote by ∇ the induced Levi-Civita connection on Ω. Let u be a smooth function defined on Ω. We denote by du the differential of u and by ∇u the gradient of u, which is the only vector field on Ω such that

$$du(X) = \langle X, \nabla u\rangle \quad \text{for any } X \in \mathcal{T}(\Omega).$$

The second covariant derivative of u is defined as

$$\nabla_{X,Y} u := \nabla_X \nabla_Y u - \nabla_X Y(u) = \nabla_X \nabla_Y u - \nabla_{\nabla_X Y} u \quad \text{for any } X, Y \in \mathcal{T}(\Omega),$$

and the Hessian of u, denoted by $\nabla^2 u$, is the symmetric 2-tensor given by

$$\nabla^2 u\ (X, Y) := \nabla_{X,Y} u \quad \text{for any } X, Y \in \mathcal{T}(\Omega).$$

The Laplacian of u, denoted by Δu, is the trace of the Hessian.

Let $\{e_1, \dots, e_n\}$ be a local orthonormal frame field for Ω and let $\{\omega^1, \dots, \omega^n\}$ be the dual coframe field; i.e., $\omega^i(e_j) = \delta_{ij}$ for any $i, j = 1, \dots, n$. The connection forms ω_{ij}'s defined by

$$\omega_{ij}(X) := \langle \nabla_X e_j, e_i\rangle, \quad X \in \mathcal{T}(\Omega), \tag{3-1}$$

and thus we have

$$\nabla_{e_i} e_j = \sum_{k=1}^{n} \omega_{kj}(e_i) e_k. \tag{3-2}$$

We also recall that the connection forms are skew symmetric, i.e., $\omega_{ij} + \omega_{ji} = 0$, for any $i, j \in \{1, \dots, n\}$. In terms of the dual coframe field the exterior derivative of u (i.e., the differential) can be written as

$$du = \sum_{i=1}^{n} u_i \omega^i,$$

where u_i denotes the covariant derivative $\nabla_{e_i} u$. We will also use the notation ∇_i to denote ∇_{e_i}.

For the second covariant derivatives, taking $X = e_i$, $Y = e_j$ and using (3-1) we have

$$\nabla_{e_i,e_j} u = \nabla_{e_i} u_j - \sum_{k=1}^{n} \omega_{kj}(e_i) u_k. \tag{3-3}$$

From now on we will use the notation u_{ij} to denote $\nabla_{e_i,e_j} u$. In particular the Hessian of u can be written as $u_{ij}\omega_j \otimes \omega_i$ and the Laplacian of u as $\Delta u = \sum_{i=1}^{n} u_{ii}$.

Definition 3.1. Let $A \subset \mathbb{L}^{n+1}$; we define the cone spanned by A as the set

$$\mathcal{C}_A := \{\rho q \in \mathbb{L}^{n+1} : q \in A,\ \rho > 0\}.$$

Remark 3.2. Observe that setting $e_{n+1}(x) := x/|x|$ for $x \in \mathcal{C}_\Omega$, and extending the e_i's as constant along radii, i.e., $e_i(x) = e_i(x/|x|)$, $x \in \mathcal{C}_\Omega$, for $i = 1, \dots, n$, we get that $\{e_1, \dots, e_{n+1}\}$ is a local orthonormal frame field for $\mathcal{C}_\Omega$, where e_{n+1} is the future-oriented unit radial direction, i.e., $\langle e_{n+1}, e_{n+1}\rangle = -1$, $\langle e_{n+1}, E_{n+1}\rangle < 0$. We also observe that by direct computation we have $\nabla^0_i e_{n+1} = e_i$ for any $i = 1, \dots, n$. We remark that by definition $e_{n+1}(q) = q$ for any $q \in \Omega$, and by abuse of notation when writing $\nabla^0_w q$, where $w \in \mathcal{T}(\mathbb{R}^{n+1})$, it will be always understood that we are computing $\nabla^0_w e_{n+1}$ at $x = q$, and $\nabla^0_q w$ will stand for $\nabla^0_{e_{n+1}} w$.

In order to derive the equation of spacelike H-bumps, one can argue as in [Treibergs and Wei 1983, Section 1] with minor adjustments. Indeed, we only need to take into account the changes due to the bilinear form $\langle\cdot,\cdot\rangle$, and the definition of mean curvature for spacelike hypersurfaces [López 2014, Section 3.2]. For the sake of completeness we derive the equation following the scheme of [loc. cit., Section 2].

Let $u \in C^0(\overline{\Omega}) \cap C^2(\Omega)$, let Σ be the associated radial graph and let $\mathcal{Y} : \Omega \to \mathbb{R}^{n+1}$ be the map defined as $\mathcal{Y}(q) := e^{u(q)}q$. From Remark 3.2 it holds that $\nabla^0_i q = e_i$ and thus

$$\nabla^0_i \mathcal{Y} = \nabla^0_i (e^u q) = e^u u_i q + e^u e_i. \tag{3-4}$$

Therefore a local basis for $T_{\mathcal{Y}(q)}\Sigma$ is given by

$$E_i(q) = e^u(e_i + u_i q), \quad i = 1, \dots, n,$$

and the components of the metric are

$$g_{ij} = \langle E_i, E_j\rangle = e^{2u}(\langle e_i, e_j\rangle + u_i u_j \langle q, q\rangle) = e^{2u}(\delta_{ij} - u_i u_j).$$

Since we look for a spacelike hypersurface we must have $|\nabla u|^2 < 1$, and by elementary computations we see that the inverse matrix (g^{ij}) is given by

$$g^{ij} = e^{-2u}\left(\delta_{ij} + \frac{u_i u_j}{1 - |\nabla u|^2}\right). \tag{3-5}$$

For the Gauss map we have

$$N(\mathcal{Y}(q)) = \frac{q + \sum_{i=1}^n u_k e_k}{(1 - |\nabla u|^2)^{1/2}}.$$

Indeed it is elementary to verify that $\langle N(\mathcal{Y}(q)), E_i\rangle = 0$ for any $i = 1, \dots, n$ and

$$\langle N(\mathcal{Y}(q)), N(\mathcal{Y}(q))\rangle = \frac{-1 + |\nabla u|^2}{1 - |\nabla u|^2} = -1.$$

Moreover, as $u = 0$ on $\partial\Omega$, there exists $q_1 \in \Omega$ such that $\nabla u(q_1) = 0$ and by definition $N(\mathcal{Y}(q_1)) = q_1$ and thus $\langle N(\mathcal{Y}(q_1)), E_{n+1}\rangle < 0$. Therefore, since $N \circ \mathcal{Y} \in C^0(\overline{\Omega}, \mathbb{R}^{n+1})$ and Ω is connected, it follows that $N(\mathcal{Y}(\Omega)) \subset \mathbb{H}^n$, so that N is future-oriented. The coefficients of the second fundamental form are given by

$$\sigma_{ij} = \langle N, \nabla^0_i \nabla^0_j \mathcal{Y}\rangle = \frac{e^u(-\delta_{ij} + u_i u_j - u_{ij})}{(1 - |\nabla u|^2)^{1/2}}. \tag{3-6}$$

Indeed, recalling Remark 3.2 and (3-4), by direct computation we have

$$\nabla^0_i(\nabla^0_j \mathcal{Y}) = e^u(u_i u_j q + \nabla^0_i \nabla^0_j u \; q + u_j e_i + u_i e_j + \nabla^0_i e_j).$$

Hence, by using the relations $\langle e_i, e_j\rangle = \delta_{ij}$, $\langle e_i, q\rangle = 0$, and regrouping the terms, we deduce that

$$\langle N, \nabla^0_i \nabla^0_j \mathcal{Y}\rangle = \frac{e^u}{(1 - |\nabla u|^2)^{1/2}}\left(u_i u_j - \nabla^0_i \nabla^0_j u + \langle \nabla^0_i e_j, q\rangle + \sum_{k=1}^n u_k \langle \nabla^0_i e_j, e_k\rangle\right). \tag{3-7}$$

Since $\langle \nabla_i^0 e_j, q\rangle = -\langle e_j, \nabla_i^0 q\rangle = -\langle e_j, e_i\rangle = -\delta_{ij}$ and

$$\nabla_i^0 \nabla_j^0 u - \sum_{k=1}^n u_k \langle \nabla_i^0 e_j, e_k\rangle = \nabla_i \nabla_j u - \sum_{k=1}^n u_k \langle \nabla_i e_j, e_k\rangle = u_{ij},$$

from (3-7) we finally get (3-6).

At the end, from [López 2014, Definition 3.3], the mean curvature of a spacelike hypersurface at $p = \mathcal{Y}(q) \in \Sigma$ is given by

$$nH(\mathcal{Y}(q)) = -\sum_{i,j=1}^n g^{ij}\sigma_{ij}.$$

Therefore, from (3-5) and (3-6), we deduce that u must satisfy the equation

$$\sum_{i,j=1}^n ((1-|\nabla u|^2)\delta_{ij} + u_i u_j)u_{ij} = n(1-|\nabla u|^2) - ne^u(1-|\nabla u|^2)^{3/2} H(\mathcal{Y}(q)).$$

4. A priori estimates

Let $\epsilon > 0$ and let $\eta_\epsilon \in C_0^\infty([0,+\infty))$ be such that $r\eta_\epsilon \in C_0^\infty([0,+\infty))$, $r \mapsto \eta_\epsilon(r)r$, is increasing in $(0, 2/\epsilon)$ and decreasing in $(2/\epsilon, +\infty)$. Assume moreover that

$$\eta_\epsilon(r)r = \begin{cases} r & \text{for } r < 1-\epsilon, \\ 1-\epsilon/2 & \text{for } 1-\epsilon/2 < r < 2/\epsilon, \\ 0 & \text{for } r > 3/\epsilon. \end{cases}$$

We define the regularized equation as

$$\sum_{i,j=1}^n \big((1-\eta_\epsilon^2(|\nabla u|)|\nabla u|^2)\delta_{ij} + \eta_\epsilon^2(|\nabla u|)u_i u_j\big)u_{ij} \\ = n(1-\eta_\epsilon(|\nabla u|)^2|\nabla u|^2)(1-\sqrt{1-\eta_\epsilon(|\nabla u|)^2|\nabla u|^2}e^u H(e^u q)). \quad (4\text{-}1)$$

To simplify the notation we will write $\eta_\epsilon^2|\nabla u|^2$ instead of $\eta_\epsilon^2(|\nabla u|)|\nabla u|^2$. The regularized Dirichlet problem for spacelike H-bumps is

$$\begin{cases} \sum_{i,j=1}^n ((1-\eta_\epsilon^2|\nabla u|^2)\delta_{ij} + \eta_\epsilon^2 u_i u_j)u_{ij} = n(1-\eta_\epsilon^2|\nabla u|^2)(1-\sqrt{1-\eta_\epsilon^2|\nabla u|^2}e^u H(e^u q)) & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \quad (4\text{-}2)$$

We denote by Q_ϵ the operator

$$Q_\epsilon(u) := \sum_{i,j=1}^n ((1-\eta_\epsilon^2|\nabla u|^2)\delta_{ij} + \eta_\epsilon^2 u_i u_j)u_{ij} - n(1-\eta_\epsilon^2|\nabla u|^2) + n(1-\eta_\epsilon^2|\nabla u|^2)^{3/2}e^u H(e^u q).$$

We claim that, in hyperbolic stereographic coordinates, the operator Q_ϵ is uniformly elliptic. This is the content of the next lemma.

Lemma 4.1. *For any $\epsilon \in (0,1)$, the operator Q_ϵ, in hyperbolic stereographic coordinates, is uniformly elliptic with ellipticity constants depending only on ϵ and Ω.*

Proof. Let $F : \mathbb{H}^n \to \mathbb{B}^n$ be the hyperbolic stereographic projection. By definition we have that Ω is mapped into a smooth proper domain $\Lambda = F(\Omega) \Subset \mathbb{B}^n$, $\phi = F^{-1} : \Lambda \to \Omega$ is a global parametrization, and there exist $c_1, c_2 > 0$ depending only on Ω such that

$$c_1 \le \tfrac{1}{4}(1 - \|y\|_n^2)^2 \le c_2 \quad \text{for all } y \in \Lambda. \tag{4-3}$$

Let us set

$$\lambda(y) := \frac{2}{1 - \|y\|_n^2}, \quad y \in \mathbb{B}^n. \tag{4-4}$$

We recall that F is an isometry and the hyperbolic metric in $\mathbb{B}^n$ is $\tilde{g} = \lambda^2 \sum_{i=1}^n dy_i \otimes dy_i$ (see Section 2). In particular, $\langle \partial_i, \partial_j \rangle = \delta_{ij} \lambda^2$, where ∂_i denotes the vector $\partial\phi/\partial y_i$, and the Christoffel symbols of the hyperbolic Levi-Civita connection are given by

$$\Gamma_{ij}^k = \frac{\lambda_i}{\lambda}\delta_{jk} + \frac{\lambda_j}{\lambda}\delta_{ik} - \sum_{l=1}^n \delta_{kl}\frac{\lambda_l}{\lambda}\delta_{ij}, \tag{4-5}$$

where $\lambda_i = \partial\lambda/\partial y_i$. In local coordinates the gradient is given by

$$\nabla u = \sum_{i,j=1}^n \tilde{g}^{ij} \frac{\partial \tilde{u}}{\partial y_i}\partial_j = \lambda^{-2} \sum_{i=1}^n \frac{\partial \tilde{u}}{\partial y_i}\partial_i, \tag{4-6}$$

and thus

$$|\nabla u|^2 = \lambda^{-2} \|\nabla_0 \tilde{u}\|_n^2, \tag{4-7}$$

where $\tilde{u} = u \circ F^{-1}$ and $\nabla_0 \tilde{u}$ is the gradient of $\tilde{u}$ with respect to the euclidean flat metric. Using the well-known expression for the Hessian and the Laplacian in local coordinates we have

$$\nabla^2 u(\partial_i, \partial_j) = \frac{\partial^2 \tilde{u}}{\partial y_j\, \partial y_i} - \sum_{k=1}^n \Gamma_{ji}^k \frac{\partial \tilde{u}}{\partial y_k}, \tag{4-8}$$

$$\Delta u = \sum_{i,j=1}^n \tilde{g}^{ij} \left(\frac{\partial^2 \tilde{u}}{\partial y_i\, \partial y_j} - \sum_{k=1}^n \Gamma_{ij}^k \frac{\partial \tilde{u}}{\partial y_k} \right) = \lambda^{-2} \sum_{i=1}^n \left(\frac{\partial^2 \tilde{u}}{\partial y_i^2} - \sum_{k=1}^n \Gamma_{ii}^k \frac{\partial \tilde{u}}{\partial y_k} \right). \tag{4-9}$$

By using the previous identities and (4-5) we infer that

$$\sum_{i,j=1}^n u_i u_j u_{ij} = \lambda^{-4} \sum_{h,k=1}^n \frac{\partial \tilde{u}}{\partial y_h} \frac{\partial \tilde{u}}{\partial y_k} \frac{\partial^2 \tilde{u}}{\partial y_h\, \partial y_k} + \Phi, \tag{4-10}$$

where Φ is a term which does not involve second-order partial derivatives. From (4-4), (4-7), (4-9) and (4-10) we deduce that the principal part of the operator Q_ϵ, in hyperbolic stereographic coordinates, is

$$\lambda^{-2} \left[\sum_{i,j=1}^n (1 - \eta_\epsilon^2 \lambda^{-2} \|\nabla_0 \tilde{u}\|_n^2) \delta_{ij} \frac{\partial^2 \tilde{u}}{\partial y_i\, \partial y_j} + \eta_\epsilon^2 \lambda^{-2} \frac{\partial \tilde{u}}{\partial y_i} \frac{\partial \tilde{u}}{\partial y_j} \frac{\partial^2 \tilde{u}}{\partial y_i\, \partial y_j} \right],$$

where $\eta_\epsilon = \eta_\epsilon(\lambda^{-1}\|\nabla_0 \tilde{u}\|_n)$. For any $i, j = 1, \dots, n$, we define, for $y \in \Lambda$, $p = (p_1, \dots, p_n) \in \mathbb{R}^n$,

$$\tilde{a}_\epsilon^{ij}(y, p) := \lambda^{-2}[(1 - \eta_\epsilon^2(\lambda^{-1}\|p\|_n)\lambda^{-2}\|p\|_n^2)\delta_{ij} + \eta_\epsilon^2(\lambda^{-1}\|p\|_n)\lambda^{-2} p_i p_j]. \tag{4-11}$$

Now, for any $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$, $y \in \Lambda$, $p \in \mathbb{R}^n$ we claim that

$$c_2 \|\xi\|_n^2 \geq \sum_{i,j=1}^{n} \tilde{a}_\epsilon^{ij}(y,p)\xi_i\xi_j \geq \tfrac{1}{2}\epsilon c_1 \|\xi\|_n^2, \tag{4-12}$$

where the constants c_1, c_2 are given by (4-3). Indeed by the definition of η_ϵ for any $y \in \Lambda$, $p \in \mathbb{R}^n$ it holds that

$$0 \leq \eta_\epsilon^2(\lambda^{-1}\|p\|_n)\lambda^{-2}\|p\|_n^2 \leq \left(1 - \tfrac{1}{2}\epsilon\right)^2$$

and thus

$$\begin{aligned}
\sum_{i,j=1}^{n} \tilde{a}_\epsilon^{ij}(y,p)\xi_i\xi_j &= \lambda^{-2}\left[(1-\eta_\epsilon^2\lambda^{-2}\|p\|_n^2)\|\xi\|_n^2 + \eta_\epsilon^2\lambda^{-2}\left(\sum_{i=1}^{n} p_i\xi_i\right)\left(\sum_{j=1}^{n} p_j\xi_j\right)\right] \\
&= \lambda^{-2}[(1-\eta_\epsilon^2\lambda^{-2}\|p\|_n^2)\|\xi\|_n^2 + \eta_\epsilon^2\lambda^{-2}(p,\xi)_n^2] \\
&\geq \lambda^{-2}(1-\eta_\epsilon^2\lambda^{-2}\|p\|_n^2)\|\xi\|_n^2 \geq \lambda^{-2}\left(1-\left(1-\tfrac{1}{2}\epsilon\right)^2\right)\|\xi\|_n^2 \geq \tfrac{1}{2}c_1\epsilon\|\xi\|_n^2,
\end{aligned}$$

where $(\cdot,\cdot)_n$ denotes the euclidean scalar product in $\mathbb{R}^n$. The proof of the other inequality in (4-12) is similar and we omit the details. □

For $t \in [0, 1]$, we define the operator

$$Q_\epsilon^t(u) := \sum_{i,j=1}^{n} ((1-\eta_\epsilon^2|\nabla u|^2)\delta_{ij} + \eta_\epsilon^2 u_i u_j)u_{ij} - nt(1-\eta_\epsilon^2|\nabla u|^2) + nt(1-\eta_\epsilon^2|\nabla u|^2)^{3/2}e^u H(e^u q).$$

For u such that $|\nabla u|_{\infty,\Omega} < 1$, we also define the operator $\mathcal{Q}^t(u)$ as

$$\mathcal{Q}^t(u) := \sum_{i,j=1}^{n} ((1-|\nabla u|^2)\delta_{ij} + u_i u_j)u_{ij} - nt(1-|\nabla u|^2) + nt(1-|\nabla u|^2)^{3/2}e^u H(e^u q). \tag{4-13}$$

Remark 4.2. By definition, for any fixed $\epsilon \in (0,1)$, if u is such that $|\nabla u|_{\infty,\Omega} \leq 1-\epsilon$, we have $Q_\epsilon^t(u) = \mathcal{Q}^t(u)$ for any $t \in [0,1]$. Moreover, in view of Lemma 4.1 and since the principal parts of Q_ϵ^t, $\mathcal{Q}^t$ are independent of t, they are uniformly elliptic even with respect to t, when passing to hyperbolic stereographic coordinates.

Remark 4.3. As seen in the proof of Lemma 4.1 we can write an explicit expression of the operator Q_ϵ^t in hyperbolic stereographic coordinates defined in the whole $\Lambda = F(\Omega)$. For our purposes we just observe that the transformed operator is of the form

$$\tilde{Q}_\epsilon^t(\tilde{u}) = \sum_{i,j=1}^{n} \tilde{a}_\epsilon^{ij}(y, \nabla_0\tilde{u})\tilde{u}_{ij} + \tilde{b}_{\epsilon,t}(y, u, \nabla_0\tilde{u}),$$

where $\tilde{a}_\epsilon^{ij} = \tilde{a}_\epsilon(y,p) : \Lambda \times \mathbb{R}^n \to \mathbb{R}$ is given by (4-11), and $\tilde{b}_{\epsilon,t} : \Lambda \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is given by

$$\begin{aligned}
\tilde{b}_{\epsilon,t}(y,z,p) := &-(1-\eta_\epsilon^2\lambda^{-2}\|p\|_n^2)\sum_{k=1}^{n} G_k(y)p_k - \eta_\epsilon^2\lambda^{-2}\sum_{h,k,r=1}^{n} G_{hkr}(y)p_h p_k p_r \\
&- nt(1-\eta_\epsilon^2\lambda^{-2}\|p\|_n^2) + nt(1-\eta_\epsilon^2\lambda^{-2}\|p\|_n^2)^{3/2}e^z H(e^z F(y)),
\end{aligned}$$

where λ is defined in (4-4), G_k, G_{hkr} are smooth functions defined in Λ, where $h, k, r \in \{1, \dots, n\}$. We point out that $\tilde{a}_\epsilon^{ij}$ does not depend on z and $\tilde{a}_\epsilon^{ij} = O(1)$, $\tilde{b}_{\epsilon,t} = O(\|p\|_n)$, as $\|p\|_n \to +\infty$, uniformly for $y \in \Lambda$, and z in compact subsets of $\mathbb{R}$. In particular, according to the notation of [Gilbarg and Trudinger 1977], setting $\mathcal{E} := \sum_{ij=1}^n \tilde{a}_\epsilon^{ij} p_i p_j$, we have that $\mathcal{E}$ does not depend on z and $\mathcal{E} = O(\|p\|_n^2)$ as $\|p\|_n \to +\infty$, uniformly for $y \in \Lambda$.

These properties will be useful in the sequel. In addition, since $(\tilde{a}_\epsilon^{ij})$ is symmetric and positive definite, when applying the results of [loc. cit., Section 15], it will be understood that we take $(\tilde{a}_\epsilon^{ij})^* = \tilde{a}_\epsilon^{ij}$ and $c_i = 0$; see [loc. cit., (15.23)].

We define now the class of admissible domains.

Definition 4.4. Let Ω be a bounded domain of $\mathbb{H}^n$ and let $H \in C^0(\mathcal{C}_{\overline{\Omega}})$. We say that (Ω, H) is admissible if there exists a constant $\theta \in (0, 1)$ such that for any $q_0 \in \partial\Omega$ and for any $t \in [0, 1]$, there exist two functions $\varphi_1, \varphi_2 \in C^2(\overline{\Omega})$ satisfying

(i) $\sup_\Omega |\nabla\varphi_i| \le 1 - \theta$ for $i = 1, 2$,

(ii) $\varphi_1(q_0) = 0$ and $\varphi_1(q_0) \le 0$ on $\partial\Omega$,

(iii) $\varphi_2(q_0) = 0$ and $\varphi_2(q_0) \ge 0$ on $\partial\Omega$,

(iv) $\mathcal{Q}^t(\varphi_1) \ge 0$, $\mathcal{Q}^t(\varphi_2) \le 0$ in Ω.

We denote by $\mathcal{A}$ the set of admissible couples (Ω, H). Given $\theta \in (0, 1)$, and given Ω and H as above, we say that (Ω, H, θ) is admissible if (Ω, H) is admissible with constant θ.

Remark 4.5. We observe that $\mathcal{A} \ne \varnothing$. In fact for any given domain $\Omega \subset \mathbb{H}^n$ for any fixed $m > 0$, the function $H(x) = 1/|x|^m$, $x \in \mathcal{C}_{\overline{\Omega}}$, is such that $(\Omega, H) \in \mathcal{A}$. In fact it is easy to see that $Q_t(0) = 0$ for any $t \in [0, 1]$, so that the functions $\varphi_1 = 0$, $\varphi_2 = 0$ satisfy (i)–(iv) for any $\theta \in (0, 1)$. More generally, for any domain Ω and for any function $H \in C^0(\mathcal{C}_{\overline{\Omega}})$ such that $H|_\Omega = 1$, we have $(\Omega, H) \in \mathcal{A}$, and (Ω, H, θ) is admissible for any $\theta \in (0, 1)$.

This condition of admissibility is very general. If we impose some regularity on $\partial\Omega$, and if we assume that H is positive, smooth and not increasing along radii, then every couple (Ω, H) is admissible. This is the content of the next result. We introduce first the following definition.

Definition 4.6. Let Ω be a bounded domain of $\mathbb{H}^n$. We say that Ω satisfies a uniform exterior geodesic ball condition if there exist $\sigma > 0$ and a map $\Xi : \partial\Omega \to \mathbb{H}^n$ of class C^2 such that for any $q_0 \in \partial\Omega$ there exists a geodesic ball in $\mathbb{H}^n$ of radius σ centered at $\xi = \Xi(q_0) \in \mathbb{H}^n \setminus \overline{\Omega}$, and denoted by $B_\sigma(\xi)$, such that $q_0 \in \partial B_\sigma(\xi)$ and $B_\sigma(\xi) \subset \mathbb{H}^n \setminus \overline{\Omega}$.

Proposition 4.7. *Let Ω be a bounded domain of $\mathbb{H}^n$ satisfying a uniform exterior geodesic ball condition. Let $H \in C^1(\mathcal{C}_{\overline{\Omega}})$ be such that $H > 0$ and $(\partial/\partial\lambda)(\lambda H(\lambda q)) \le 0$ for all $q \in \overline{\Omega}$, $\lambda > 0$. Then (Ω, H) is admissible.*

Proof. Let $\mathrm{dist}_{\mathbb{H}^n}(\cdot, \cdot)$ be the geodesic distance in $\mathbb{H}^n$. Let $\sigma > 0$ be the number given by Definition 4.6 for Ω. In particular, by definition, it follows that for any $q_0 \in \partial\Omega$ there exists $\xi = \xi(q_0) \notin \overline{\Omega}$ such that $\mathrm{dist}_{\mathbb{H}^n}(\xi, \partial\Omega) = \mathrm{dist}_{\mathbb{H}^n}(\xi, q_0) = \sigma$.

Let $q_0 \in \partial\Omega,\ t \in [0,1]$ and let $\xi = \xi(q_0)$ satisfying the above properties. Since every geodesic ball of $\mathbb{H}^n$ is geodesically convex, see [Papadopoulos 2005, Section 2.5], we can take $R > 0$ sufficiently large so that $\overline{\Omega}$ is contained in the geodesically convex ball $B_R(\xi)$. We observe that since Ω is bounded and $\mathrm{dist}_{\mathbb{H}^n}(\xi, \partial\Omega) = \sigma$, up to a new choice of a larger R, we can assume that R is uniform with respect to the choice of $q_0 \in \partial\Omega$.

Arguing as in proof of [Gerhardt 1983, Theorem 2.1], we set $|||q||| := \mathrm{dist}_{\mathbb{H}^n}(q, \xi)$ to denote the geodesic distance from ξ and we define

$$\delta^+(q) := \int_{|||q_0|||}^{|||q|||} (1+\gamma(s))^{-1/2}\, ds, \quad \gamma(s) := \alpha e^{\beta s},$$

where α, β are positive constants to be determined later. By construction it holds that $\delta^+ \in C^2(\overline{B_R(\xi)} \setminus \{\xi\})$, $\delta^+ \in C^2(\overline{\Omega}),\ \delta^+(q_0) = 0$ and $\delta^+ \geq 0$ in $\overline{\Omega}$ because of the exterior ball condition.

Let us consider the operator

$$\mathcal{Q}^t_{\mathrm{div}}(u) := -\,\mathrm{div}_{\mathbb{H}^n}\left(\frac{\nabla u}{\sqrt{1-|\nabla u|^2}}\right) + \frac{nt}{\sqrt{1-|\nabla u|^2}} - nte^u H(e^u q)$$

(which is the divergence form of $-\mathcal{Q}^t$). We set

$$A(u) := -\,\mathrm{div}_{\mathbb{H}^n}\left(\frac{\nabla u}{\sqrt{1-|\nabla u|^2}}\right), \quad \nu(u) := \frac{1}{\sqrt{1-|\nabla u|^2}}.$$

We observe that $|\nabla |||q||||| = 1$ for any $q \in B_R(\xi) \setminus \{\xi\}$. This property is known for general manifolds when R is sufficiently small so that $B_R(\xi)$ is contained in a normal neighborhood of ξ; see [Lee 1997, Corollaries 6.9 and 6.11]. In our case, as a consequence of the Cartan–Hadamard theorem, since $\mathbb{H}^n$ has negative sectional curvature it admits global normal coordinates and we are done.

Therefore, since the covariant derivatives of δ^+ are given by $(\delta^+)_i = (1+\gamma)^{-1/2} |||q|||_i$, we obtain that for any $q \in \overline{\Omega}$

$$|\nabla \delta^+| = (1+\gamma)^{-1/2} < 1,$$

and

$$\nu(\delta^+) = \gamma^{-1/2}(1+\gamma)^{1/2}.$$

In addition, by direct computation, see [Gerhardt 1983, (2.14)–(2.16)], it holds that

$$A(\delta^+) = (1+\gamma)^{-1/2}\big(\tfrac{1}{2}\beta - \Delta |||q|||\big)\nu(\delta^+).$$

We observe that $\Delta |||q|||$ is smooth and bounded in compact subsets of $\overline{B_R(\xi)} \setminus \{\xi\}$ and it is singular as $q \to \xi$. Indeed, see [Gerhardt 1983, (2.17)–(2.18)], we have

$$-\Delta |||q||| = -\frac{n-1}{|||q|||} + \Psi, \tag{4-14}$$

where Ψ is a bounded term which is given, in normal coordinates centered at ξ, by

$$\Psi = -\sum_{i,j,k=1}^{n} g^{ij}\, \Gamma^k_{ij}\, |||q|||_k.$$

In particular, in view of the uniform exterior geodesic ball condition, since $\mathrm{dist}_{\mathbb{H}^n}(q, \xi(q_0)) \geq \sigma$ for any $q \in \overline{\Omega}$, for any $q_0 \in \partial\Omega$, from (4-14) we infer that $\Delta |||q|||$ is bounded in $\overline{\Omega}$ by a constant depending only on n, σ, Ω, q_0. In addition, by definition the map $q_0 \mapsto \xi$ is of class $C^2(\partial\Omega, \mathbb{H}^n)$ and thus, by compactness of $\partial\Omega$, it follows that $\Delta |||q|||$ is bounded by a constant depending only n, σ, Ω.

Now, by the previous relations we have

$$\begin{aligned} A(\delta^+) + tn\nu(\delta^+) &= \big[(1+\gamma)^{-1/2}\big(\tfrac{1}{2}\beta - \Delta |||q|||\big) + tn\big](\gamma^{-1/2}(1+\gamma)^{1/2}) \\ &\geq \big(\tfrac{1}{2}\beta - \Delta |||q|||\big)\gamma^{-1/2} = \big(\tfrac{1}{2}\beta - \Delta |||q|||\big)\alpha^{-1/2}e^{-\beta s/2}. \end{aligned}$$

Setting $\overline{H} := \max_{q\in\overline{\Omega}} H(q) > 0$, we can choose β sufficiently large so that $\frac{1}{2}\beta - \Delta |||q||| > 0$ for any $q \in \overline{\Omega}$. With this choice of β we choose α sufficiently small so that $\big(\frac{1}{2}\beta - \Delta |||q|||\big)\alpha^{-1/2}e^{\beta s/2} \geq n\overline{H}$ for any $x \in \overline{\Omega},\ s \in [|||q_0|||, \sup_{q\in\overline{\Omega}} |||q|||]$. Therefore, since $\delta^+ \geq 0$ in $\overline{\Omega}$, and in view of the monotonicity assumption on H, it follows that

$$\begin{aligned} A(\delta^+) + nt\nu(\delta^+) - nte^{\delta^+}H(e^{\delta^+}q) &\geq A(\delta^+) + nt\nu(\delta^+) - nte^{0}H(e^{0}q) \\ &\geq A(\delta^+) + nt\nu(\delta^+) - nt\overline{H} \\ &\geq 0. \end{aligned}$$

Hence, $\mathcal{Q}^t_{\mathrm{div}}(\delta^+) \geq 0$ in $\overline{\Omega}$, which is equivalent to $\mathcal{Q}^t(\delta^+) \leq 0$ in $\overline{\Omega}$, and in addition by construction we have $\delta^+ \geq 0$ on $\partial\Omega$, $\delta^+(x_0) = 0$, and $|\nabla\delta^+| = (1+\gamma)^{-1/2} \leq 1 - \theta_+$ for some number $\theta_+ = \theta_+(\alpha, \beta) \in (0, 1)$.

As pointed out before, in view of the uniform exterior ball condition, $-\Delta |||q|||$ is uniformly bounded by a constant depending only on n, σ, Ω, and by construction $\sup_{q\in\overline{\Omega}} |||q||| \leq R$. Therefore, the numbers α, β can be chosen in a uniform way with respect to the base point $q_0 \in \partial\Omega$ (and also with respect to $t \in [0, 1]$). Hence, there exists $\theta_+ \in (0, 1)$ such that for any $q_0 \in \partial\Omega,\ t \in [0, 1]$, the function $\varphi_2 := \delta^+$ (which depends on the choice of q_0 but not on t) satisfies (i)–(iv) of Definition 4.4 with $\theta = \theta^+$. For the other barrier is suffices to take $\varphi_1 := \delta^-$, where

$$\delta^- := -\int_{|||q_0|||}^{|||q|||} (1+\gamma(s))^{-1/2}\, ds,$$

and to argue as in the previous case. We observe that in this case the choice of α, β has to be made in a different way but it is still uniform with respect to q_0, and t.

In fact

$$\begin{aligned} A(\delta^-) + tn\nu(\delta^-) &= -\gamma^{-1/2}\big(\tfrac{1}{2}\beta - \Delta |||q||| - tn(1+\gamma)^{1/2}\big) \\ &= -\alpha^{-1/2}e^{-\beta s/2}\big(\tfrac{1}{2}\beta - \Delta |||q||| - tn(1+\alpha e^{\beta s})^{1/2}\big). \end{aligned}$$

Taking $\alpha = e^{-\beta \sup_{q\in\overline{\Omega}} |||q|||}$, it follows that $n(1+\alpha e^{\beta s})^{1/2} \leq \sqrt{2}n$ for any $\beta > 0,\ s \in [|||q_0|||, \sup_{q\in\overline{\Omega}} |||q|||]$. With this choice of α, we choose β such that

$$\tfrac{1}{2}\beta - \Delta |||q||| - 2n \geq 0$$

for $x \in \overline{\Omega}$. At the end, we have $A(\delta^-) + tn\nu(\delta^-) \leq 0$, and thus since $H > 0$ it holds that

$$A(\delta^-) + tn\nu(\delta^-) - tne^{\delta^-}H(e^{\delta^-}q) \leq A(\delta^-) + tn\nu(\delta^-) \leq 0 \quad \text{in } \Omega.$$

As before we find a uniform $\theta_- \in (0,1)$ such that for any $q_0 \in \partial\Omega,\ t \in [0,1]$, the function $\varphi_1 := \delta^-$ satisfies $|\nabla\varphi_1| \le 1-\theta_-$ and (ii)–(iv) of Definition 4.4. At the end, choosing $\theta := \min\{\theta_-, \theta_+\}$ we have that for any $q_0 \in \partial\Omega$, $t \in [0,1]$, the functions φ_1, φ_2 satisfy (i)–(iv) of Definition 4.4, and hence (Ω, H) is admissible. □

Remark 4.8. It is important to note that in the previous proof the choice of θ depends only on n, σ, Ω and depends on H just by the number $\overline{H} := \max_{q\in\overline{\Omega}} H(q) > 0$ because of the monotonicity assumption. In particular θ does not depend on the derivatives of H.

If $H \in C^1(C_{\overline{\Omega}}(r_1,r_2))$ we define a canonical extension of H to a mapping on the cone $\mathcal{C}_{\overline{\Omega}}$ in the following way: set

$$h_1(q) := \left[\frac{\partial}{\partial\rho}\rho H(\rho q)\right]_{\rho=r_1}, \quad h_2(q) := \left[\frac{\partial}{\partial\rho}\rho H(\rho q)\right]_{\rho=r_2}$$

and

$$\widehat{H}(\rho q) := \begin{cases} (r_1/\rho)H(r_1 q) + (1-r_1/\rho)h_1(q) & \text{for } \rho \in (0, r_1), \\ H(\rho q) & \text{for } \rho \in [r_1, r_2], \\ (r_2/\rho)H(r_2 q) + (1-r_2/\rho)h_2(q) & \text{for } \rho \in (r_2, +\infty). \end{cases} \tag{4-15}$$

Remark 4.9. It is elementary to check that $\widehat{H} \in C^1(\mathcal{C}_{\overline{\Omega}})$, and if H satisfies $(\partial/\partial\lambda)(\lambda H(\lambda q)) \le 0$ for all $q \in \overline{\Omega},\ \lambda \in [r_1, r_2]$, it follows that

$$\frac{\partial}{\partial\lambda}(\lambda\widehat{H}(\lambda q)) \le 0 \quad \text{for all } q \in \overline{\Omega},\ \lambda > 0.$$

Therefore, since $\widehat{H}(x) = H(x)$ for $x \in \overline{\Omega}$, by Remark 4.8 if Ω satisfies the hypotheses of Proposition 4.7 and H is positive, it follows that $(\widehat{H}, \Omega)$ is admissible with constant which does not depend on the choice of r_1, r_2, and the derivatives of $\widehat{H}$.

In view of the previous remark, the following definition makes sense:

Definition 4.10. Let Ω be a bounded domain of $\mathbb{H}^n$, let $0 < r_1 \le 1 \le r_2$ and let $H \in C^1(C_{\overline{\Omega}}(r_1,r_2))$. We say that (Ω, H) is admissible if $(\Omega, \widehat{H})$ is admissible, where $\widehat{H}$ is the extension of H defined in (4-15), and for $\theta \in (0,1)$ we say that (Ω, H, θ) is admissible if $(\Omega, \widehat{H}, \theta)$ is admissible.

Now we have all the tools to prove the a priori estimates. Let us fix some notation: let $k \in \mathbb{N},\ \alpha \in (0,1)$, and we consider the subspaces $C_0^{k,\alpha}(\overline{\Omega}) := \{u \in C^{k,\alpha}(\overline{\Omega}) : u|_{\partial\Omega} = 0\}$, $C_0^k(\overline{\Omega}) := \{u \in C^k(\overline{\Omega}) : u|_{\partial\Omega} = 0\}$, endowed, respectively, with the usual norms $|\cdot|_{k,\alpha}$, $|\cdot|_k$. We point out that $C_0^{k,\alpha}(\overline{\Omega})$, $C_0^k(\overline{\Omega})$ are closed subspaces of Banach spaces and thus they are Banach too. When needed we will specify also the domain in the norms; otherwise it will be understood that the domain is Ω. Moreover, for the $C^0(\overline{\Omega})$-norm we will use the notation $|\cdot|_\infty$, $|\cdot|_{\infty,\Omega}$, and $\|\cdot\|_\infty$, $\|\cdot\|_{\infty,\Omega}$ when working in the euclidean setting.

We define $\widehat{Q}_\epsilon^t$ as the operator obtained from Q_ϵ^t by replacing H with its extension $\widehat{H}$, and in the class of functions satisfying $|\nabla u|_{\infty,\Omega} < 1$ we define $\widehat{\mathcal{Q}}^t$ as

$$\widehat{\mathcal{Q}}^t(u) := \sum_{i,j=1}^n ((1-|\nabla u|^2)\delta_{ij} + u_i u_j)u_{ij} - nt(1-|\nabla u|^2) + nt(1-|\nabla u|^2)^{3/2} e^u \widehat{H}(e^u q). \tag{4-16}$$

In order to simplify the notation we set $L_{\epsilon,u}u := \sum_{i,j=1}^{n}((1-\eta_\epsilon^2|\nabla u|^2)\delta_{ij}+\eta_\epsilon^2 u_i u_j)u_{ij}$. The first result we prove is about a priori C^0 estimates for solutions of $\widehat{Q}_\epsilon^t(u)=0$.

Lemma 4.11 (a priori C^0 estimates). *Let Ω be a bounded domain and let $r_1 \neq r_2$ such that* $0<r_1\le 1\le r_2$. *Assume that $H \in C^1(\mathcal{C}_{\overline{\Omega}}(r_1,r_2))$ satisfies*

$$H(r_1 q) > r_1^{-1} \quad \textit{and} \quad H(r_2 q) < r_2^{-1} \quad \textit{for any } q \in \overline{\Omega}, \tag{4-17}$$

and

$$\frac{\partial}{\partial\lambda}(\lambda H(\lambda q)) \le 0 \quad \textit{for all } q \in \overline{\Omega},\ \lambda \in [r_1,r_2]. \tag{4-18}$$

For $\epsilon \in (0,1)$, *for every* $t \in [0,1]$, *if $u \in C_0^2(\overline{\Omega})$ is a solution of $\widehat{Q}_\epsilon^t(u)=0$ then*

$$\log r_1 \le u(q) \le \log r_2 \quad \textit{for every } q \in \overline{\Omega}.$$

Proof. Let us observe that since we are assuming (4-17), (4-18), it holds that

$$\widehat{H}(x) > |x|^{-1} \quad \text{if } |x| \leqslant r_1,\ x \in \mathcal{C}_{\overline{\Omega}} \qquad \text{and} \qquad \widehat{H}(x) < |x|^{-1} \quad \text{if } |x| \geqslant r_2,\ x \in \mathcal{C}_{\overline{\Omega}}. \tag{4-19}$$

Let $\epsilon \in (0,1)$, let $t \in [0,1]$ and let $u \in C_0^2(\overline{\Omega})$ such that $\widehat{Q}_\epsilon^t(u)=0$. By definition u is a classical solution of the Dirichlet problem

$$\begin{cases} L_{\epsilon,u}u = nt(1-\eta_\epsilon^2|\nabla u|^2)(1-\sqrt{1-\eta_\epsilon^2|\nabla u|^2}e^u\widehat{H}(e^u q)) & \text{in } \Omega, \\ u=0 & \text{in } \partial\Omega. \end{cases} \tag{4-20}$$

Let $q_0 \in \overline{\Omega}$ such that $u(q_0)=\max_{\overline{\Omega}} u$. Assume by contradiction that $u(q_0) > \log r_2$. Then $q_0 \in \Omega$ because $r_2 \ge 1$ and $u=0$ on $\partial\Omega$. Hence $\nabla u(q_0)=0$, and since q_0 is a maximum point, it holds that $\Delta u(q_0) \le 0$, and by the definition of $L_{u,\epsilon}$ this reads as

$$L_{u,\epsilon}u \le 0.$$

On the other hand it must be that $t>0$ because otherwise if $t=0$ then $u\equiv 0$. Moreover

$$L_{u,\epsilon}u(q_0) = nte^{u(q_0)}\left(\frac{1}{e^{u(q_0)}} - \widehat{H}(e^{u(q_0)}q_0)\right) > 0,$$

because $\widehat{H}(x) < |x|^{-1}$ as $|x|>r_2$. Thus we reach a contradiction. The same argument holds to show that $\min_{\overline{\Omega}} u \ge \log r_1$. ☐

Lemma 4.12 (a priori $C^{1,\alpha}$ estimates). *Let $\epsilon \in (0,1)$ and let Ω be a bounded domain of class C^2. Assume that H satisfies* (4-17), (4-18). *Then, there exist two positive constants M, C and $\alpha_0 \in (0,1)$ such that for all $t \in [0,1]$ if $u \in C_0^2(\overline{\Omega})$ is such that $|\nabla u|_{\infty,\partial\Omega} \le 1-\epsilon$ and is a solution of the equation $\widehat{Q}_\epsilon^t(u)=0$, then*

$$|\nabla u|_{\infty,\Omega} \le M, \quad |u|_{1,\alpha_0} \le C.$$

Proof. Let us fix $\epsilon \in (0,1)$, let $t \in [0,1]$ and let $u=u_t$ be a solution of $\widehat{Q}_\epsilon^t(u)=0$. From Lemma 4.11 we have $\log r_1 \le u \le \log r_2$ and thus by definition u also satisfies $Q_\epsilon^t(u)=0$. Therefore, from now on we can work just with the operator Q_ϵ^t.

Let us set $b_{\epsilon,t}(q,u,\nabla u):=nt(1-\eta_\epsilon^2|\nabla u|^2)(1-\sqrt{1-\eta_\epsilon^2|\nabla u|^2}e^uH(e^uq))$. In view of Remark 4.2, passing to hyperbolic stereographic coordinates, the operator Q_ϵ^t is uniformly elliptic with constants independent of t; moreover, by definition and thanks to Lemma 4.11, the term $b_{\epsilon,t}(q,u,\nabla u)$ is uniformly bounded with respect to t.

Now there are only two possibilities: there exists a constant M independent of t such that $|\nabla u|_{\infty,\Omega}\le M$ for all $t\in[0,1]$ or there exists a subsequence $(t_k)\subset[0,1]$ such that $|\nabla u_{t_k}|_{\infty,\Omega}\to+\infty$ as $k\to+\infty$. We claim that the second case cannot happen. To this end we will give two proofs of this fact; one works only in dimension 2, the other one works in any dimension.

Case of dimension 2: Assume that $|\nabla u_{t_k}|_{\infty,\Omega}\to+\infty$. Let us set $\Omega_k':=\{x\in\overline{\Omega}:|\nabla u_{t_k}|\ge 3/\epsilon\}$, and $q_k\in\Omega_k'$ such that $|\nabla u_{t_k}|_{\infty,\Omega}=|\nabla u_{t_k}(q_k)|$. We observe that Ω_k' is closed and hence is a compact subset of $\overline{\Omega}$, and since $|\nabla u_t|_{\infty,\partial\Omega}\le 1-\epsilon$, we have $\Omega_k'\cap\partial\Omega=\varnothing$ for all k. Let Ω_k'' be the connected component of Ω_k' containing q_k. Consider now the auxiliary problem

$$\begin{cases}\Delta v_{t_k}=nt_k(1-\eta_\epsilon^2|\nabla u_{t_k}|^2)(1-\sqrt{1-\eta_\epsilon^2|\nabla u_{t_k}|^2}e^{u_{t_k}}H(e^{u_{t_k}}q)) & \text{in } \Omega,\\ v_{t_k}=0 & \text{in } \partial\Omega.\end{cases} \tag{4-21}$$

We observe that since u_{t_k} is uniformly bounded, by construction and standard regularity theory we get that v_{t_k} and its gradient are uniformly bounded with respect to k. By definition $w_{t_k}:=u_{t_k}-v_{t_k}$ is harmonic in Ω_k''. Therefore, considering the isometry $F:\mathbb{H}^2\to\mathbb{B}^2$, and since harmonicity is preserved through composition with isometries, see [Hélein and Wood 2008, Section 2.2], we know $\tilde{w}_{t_k}:=F\circ w_{t_k}$ is harmonic in $\widetilde{\Omega}_k'':=F(\Omega_k'')\Subset\mathbb{B}^2$. Now, since the hyperbolic metric $\tilde{g}$ is conformal to the euclidean metric g_0 in $\mathbb{B}^2$ (see Section 2), we have that $\tilde{w}_{t_k}$ is harmonic also in $(\widetilde{\Omega}_k'',g_0)$. We point out that, in general, this fact is false in other dimensions. Hence, since $\tilde{w}_{t_k}$ is harmonic, it follows that also $\nabla_0\tilde{w}_{t_k}$ is harmonic in $\widetilde{\Omega}_k''$, so $\|\nabla_0\tilde{w}_{t_k}\|_n$ achieves its maximum on the boundary, and thus $\|\nabla_0\tilde{w}_{t_k}\|_\infty=\|\nabla_0\tilde{w}_{t_k}\|_{\infty,\partial\widetilde{\Omega}_k''}\to+\infty$ as $k\to+\infty$. On the other hand, by construction and (4-7) we have that

$$\begin{aligned}\|\nabla_0\tilde{w}_{t_k}\|_{\infty,\partial\widetilde{\Omega}_k''}&=\sup_{y\in\partial\widetilde{\Omega}_k''}\|\nabla_0\tilde{w}_{t_k}(y)\|_n=\sup_{q\in\partial\Omega_k''}\frac{4}{(1-\|F(q)\|_n^2)^2}|\nabla w_{t_k}(q)|\\&\le\sup_{q\in\partial\Omega_k''}\frac{4}{(1-\|F(q)\|_n^2)^2}\Big(\frac{3}{\epsilon}+|\nabla v_k(q)|\Big)\end{aligned}$$

is uniformly bounded and thus we get a contradiction.

Case of any dimension $n\ge 2$: Consider $\tilde{u}:=u\circ F^{-1}$, where $F:\mathbb{H}^n\to\mathbb{B}^n$ is the hyperbolic stereographic projection. Then $\tilde{u}$ is a solution of a uniformly elliptic equation which satisfies the hypotheses of [Gilbarg and Trudinger 1977, Theorem 15.2]; see [loc. cit., (i), p. 367]. In fact, thanks to Remark 4.3, writing Q_ϵ^t in local coordinates we see by elementary computations that the natural conditions of [loc. cit., (i), p. 367], are satisfied (uniformly in t). In particular, introducing the operator

$$\delta=\frac{\partial}{\partial z}+\sum_{k=1}^n\|p\|_n^{-2}p_k\frac{\partial}{\partial y_k},$$

we see that $\delta\tilde{a}^{ij}_\epsilon$, $\delta\tilde{b}_{\epsilon,t}$ satisfy, as $\|p\|_n \to +\infty$ (uniformly for $(y,z) \in \Lambda \times [\log r_1, \log r_2]$, and in $t \in [0,1]$), the growth conditions of [loc. cit., (15.36)], and thus the hypotheses of [loc. cit., Theorem 15.2] are satisfied with $c \le 0$.

Thanks to Lemma 4.11 the oscillation of u is uniformly bounded; moreover, since $|\nabla u|_{\infty,\partial\Omega} \le 1-\epsilon$ and the structural conditions are satisfied uniformly in t, we have that the constant given by [loc. cit., Theorem 15.2] is uniformly bounded with respect to t. Hence there exists C independent of t such that $\|\nabla_0 \tilde{u}\|_{\infty,F(\Omega)} \le C$, and hence, in view of (4-7), the same holds for $|\nabla u|_{\infty,\Omega}$. Therefore, it cannot happen that there exists a sequence (t_k) such that $|\nabla u_{t_k}|_{\infty,\Omega} \to +\infty$, and we are done.

Conclusion: From the previous discussion the only possibility is that there exists a constant M such that $|\nabla u|_{\infty,\Omega} \le M$ for all $t \in [0,1]$. From this fact, up to passing to local coordinates, since Q^t_ϵ is uniformly elliptic (with ellipticity constant independent of t) and $b_{\epsilon,t}(q,u,\nabla u)$ is uniformly bounded in t, thanks to [loc. cit., Theorem 13.7], there exists $\alpha_0 \in (0,1)$ and a positive constant C, both depending only on n, Ω, $|\nabla u|_{\infty,\Omega}$, Ω, and the ratio between the uniform bound on $b_{\epsilon,t}$ and the lower ellipticity constant, such that

$$[\nabla u]_{0,\alpha_0} \le C,$$

where $[\,\cdot\,]_{0,\alpha_0}$ denotes the C^{0,α_0} seminorm. At the end, from this fact and Lemma 4.11, we conclude that

$$|u|_{1,\alpha_0} \le C_1$$

for some constant C_1 not depending on t, and the proof is complete. □

5. Existence and uniqueness of solutions for the regularized problem

The aim of this section is to prove the following:

Theorem 5.1. *Let $\alpha \in (0,1)$, $0 < r_1 \le 1 \le r_2$, with $r_1 \ne r_2$, let Ω be a bounded domain of $\mathbb{H}^n$, with boundary of class $C^{2,\alpha}$. Let $H \in C^1(C_{\overline{\Omega}}(r_1,r_2))$ satisfy hypotheses* (i), (ii) *of Theorem 1.3. Assume that (Ω, H) is admissible. Then, there exists $\bar{\epsilon} \in (0,1)$ such that for any $\epsilon \in (0,\bar{\epsilon})$ equation* (4-2) *has a solution. Moreover such a solution is the unique solution of problem* (4-2) *whose associated radial graph is contained in $\mathcal{C}_{\overline{\Omega}}(r_1,r_2)$.*

Proof. We divide the proof into several steps.

Step 1: Choice of $\bar{\epsilon} \in (0,1)$.

Let $\widehat{H}$ be the extension of H defined in (4-15). Since $(\Omega, \widehat{H})$ is admissible we choose $\bar{\epsilon} = \theta$, where θ is given by Definition 4.4.

Let $\epsilon \in (0,1)$ such that $\epsilon < \theta$, let $\alpha_0 \in (0,1)$ be the number given by Lemma 4.12 and set $\beta := \min\{\alpha, \alpha_0\}$. For any fixed $w \in C^{1,\beta}(\overline{\Omega})$ we define the operator $L_{w,\epsilon} : C^{2,\beta}_0(\overline{\Omega}) \to C^{0,\beta}(\overline{\Omega})$ as

$$L_{w,\epsilon} u := \sum_{i,j=1}^{n} ((1-\eta^2_\epsilon |\nabla w|^2)\delta_{ij} + \eta^2_\epsilon w_i w_j) u_{ij}.$$

Step 2: For every $w \in C^{1,\beta}(\overline{\Omega})$ the operator $L_{w,\epsilon}$ is a bijection of $C^{2,\beta}_0(\overline{\Omega})$ onto $C^{0,\beta}(\overline{\Omega})$.

A mapping $u \in C_0^{2,\beta}(\overline{\Omega})$ belongs to the kernel of $L_{w,\epsilon}$ if and only if u solves the Dirichlet problem

$$\begin{cases} \sum_{i,j=1}^{n} ((1-\eta_\epsilon^2|\nabla w|^2)\delta_{ij} + \eta_\epsilon^2 w_i w_j)u_{ij} = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{5-1}$$

Since $L_{w,\epsilon}$ is uniformly elliptic (see Lemma 4.1), by the maximum principle, because $u = 0$ on $\partial\Omega$ we obtain that $u = 0$ in Ω, and this means that $L_{w,\epsilon}$ is injective. In order to prove that $L_{w,\epsilon}$ is onto we use the continuity method. Let $t \in [0, 1]$; we introduce the family of operators

$$\mathcal{L}_{t,w,\epsilon} : C_0^{2,\beta}(\overline{\Omega}) \to C^{0,\beta}(\overline{\Omega})$$

defined by

$$\mathcal{L}_{t,w,\epsilon} = (1-t)\Delta + tL_{w,\epsilon}.$$

We observe that $\mathcal{L}_{0,w,\epsilon} = \Delta$ and for every $f \in C^{0,\beta}(\overline{\Omega})$ the Dirichlet problem

$$\begin{cases} \Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

admits a solution $C^{2,\beta}(\overline{\Omega})$. That is, $\mathcal{L}_{0,w,\epsilon}$ sends $C^{0,\beta}(\overline{\Omega})$ onto $C_0^{2,\beta}(\overline{\Omega})$. Now we claim that there exists a constant $C > 0$ such that

$$|u|_{2,\beta} \le C|\mathcal{L}_{t,w,\epsilon}|_{0,\beta} \tag{5-2}$$

for every $t \in [0, 1]$, for every $u \in C_0^{2,\beta}(\overline{\Omega})$. In view of the method of continuity, this is enough to infer that $\mathcal{L}_{1,w,\epsilon} = L_{w,\epsilon}$ is onto. If (5-2) is false then there exist sequences $(t_k) \subset [0, 1]$ and $(u_k) \subset C_0^{2,\beta}(\overline{\Omega})$ such that

$$|\mathcal{L}_{t,w,\epsilon}|_{0,\beta} \to 0 \quad \text{and} \quad |u_k|_{2,\beta} = 1. \tag{5-3}$$

By compactness, in particular using also the Ascoli–Arzelà theorem, there exist $t \in [0, 1]$ and $u \in C_0^{2,\beta}(\overline{\Omega})$ such that, up to subsequences,

$$t_k \to t \quad \text{and} \quad u_k \to u \quad \text{in } C^2(\overline{\Omega}).$$

By continuity we have $\mathcal{L}_{t,w,\epsilon}u = 0$. Since, up to passing to hyperbolic stereographic coordinates, $\mathcal{L}_{t,w,\epsilon}$ is a convex combination of elliptic operators, it follows that $u = 0$. In particular

$$u_k \to 0 \quad \text{in } C^0(\overline{\Omega}). \tag{5-4}$$

We observe that

$$\mathcal{L}_{t,w,\epsilon}u = \sum_{i,j=1}^{n} a_{t,\epsilon}^{ij} u_{ij},$$

where $a_{t,\epsilon}^{ij} = ((1 - t\eta_\epsilon^2|\nabla w|^2)\delta_{ij} + t\eta_\epsilon^2 w_i w_j)$, and $\mathcal{L}_{t,w,\epsilon}$ is uniformly elliptic; moreover, arguing as in the proof of Lemma 4.1 we see that the ellipticity constants are independent of t. Since the boundary is smooth we can apply global Schauder estimates and we get

$$|u_k|_{2,\beta} \le C(|u_k|_\infty + |\mathcal{L}_{t,w,\epsilon}|_{0,\beta}),$$

with C independent of k. This yields a contradiction with (5-3), (5-4). Hence (5-2) is true and the proof of Step 2 is complete.

Step 3: For every $C > 0$ there exists $K > 0$ such that if $|w|_{1,\beta} \le C$ then $|u|_{2,\beta} \le K|L_{w,\epsilon}u|_{0,\beta}$ for every $u \in C_0^{2,\beta}(\overline{\Omega})$.

We argue by contradiction as in the last part of the proof of Step 2. If the result is false then there exist a bounded sequence (w_k) in $C^{1,\beta}(\overline{\Omega})$ and a sequence (u_k) in $C_0^{2,\beta}(\overline{\Omega})$ such that

$$|u_k|_{2,\beta} = 1 \quad \text{and} \quad |L_{w_k,\epsilon}u_k|_{0,\beta} \to 0. \tag{5-5}$$

By compactness, there exist $w \in C^1(\overline{\Omega})$ and $u \in C_0^2(\overline{\Omega})$ such that, up to subsequences,

$$w_k \to w \quad \text{in } C^1(\overline{\Omega}) \qquad \text{and} \qquad u_k \to u \quad \text{in } C^2(\overline{\Omega}).$$

By continuity we get $L_{w,\epsilon}u = 0$. Then $u = 0$, by Step 2. Taking into account Lemma 4.1 we observe that the operators $L_{w_k,\epsilon}$ are uniformly elliptic with ellipticity constants independent of k. Using standard Schauder estimates we obtain that

$$|u_k|_{2,\beta} \le C_1(|u_k|_\infty + |L_{w_k,\epsilon}u_k|_{0,\beta}),$$

where C_1 is a constant independent of k. Since $u_k \to 0$ in $C^0(\overline{\Omega})$ and by (5-5) we reach a contradiction. The proof of Step 3 is complete.

Step 4: Let (w_k) be a bounded sequence in $C^{1,\beta}(\overline{\Omega})$ and let (f_k) be a bounded sequence in $C^{0,\beta}(\overline{\Omega})$. Then the sequence (u_k) of solutions of

$$\begin{cases} L_{w_k,\epsilon}u_k = f_k & \text{in } \Omega, \\ \qquad u_k = 0 & \text{on } \partial\Omega. \end{cases} \tag{5-6}$$

is bounded in $C^{2,\beta}(\overline{\Omega})$.

The existence of a solution u_k of (5-6) is given by Step 2, and the thesis follows from Step 3.

Step 5: Let us consider the map $T_\epsilon : C_0^{1,\beta}(\overline{\Omega}) \to C_0^{1,\beta}(\overline{\Omega})$, defined as follows: for every $w \in C_0^{1,\beta}(\overline{\Omega})$ we set $T_\epsilon(w) := u$, where $u = u(w,\epsilon)$ is the unique solution of the problem

$$\begin{cases} \sum_{i,j=1}^{n} ((1-\eta_\epsilon^2|\nabla w|^2)\delta_{ij} + \eta_\epsilon^2 w_i w_j)u_{ij} = n(1-\eta_\epsilon^2|\nabla w|^2)(1-\sqrt{1-\eta_\epsilon^2|\nabla w|^2}e^w \widehat{H}(e^w q)) & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{5-7}$$

We claim that T_ϵ is a compact operator.

We first observe T_ϵ is well-defined; in fact, as proved in Step 2, for a given $w \in C_0^{1,\beta}(\overline{\Omega})$, the operator $L_{w,\epsilon}$ is a bijection between $C_0^{2,\beta}(\overline{\Omega})$ and $C^{0,\beta}(\overline{\Omega})$. In addition is T_ϵ is a linear map. It remains to prove that T_ϵ maps bounded families of $C_0^{1,\beta}(\overline{\Omega})$ into relatively compact subsets of $C_0^{1,\beta}(\overline{\Omega})$.

Let (w_λ) be a bounded family of $C_0^{1,\beta}(\overline{\Omega})$; then (u_λ), where $u_\lambda = Tw_\lambda$, is a family of solutions of (5-7). Hence $u_\lambda \in C_0^{2,\beta}(\overline{\Omega})$ and since we assumed that there exists $C > 0$ such that $|w_\lambda|_{1,\beta} \le C$, by Step 3 we have

$$|u_\lambda|_{2,\beta} \le K\big|n(1-\eta_\epsilon^2|\nabla w_\lambda|^2)(1-\sqrt{1-\eta_\epsilon^2|\nabla w_\lambda|^2}e^{w_\lambda}\widehat{H}(e^{w_\lambda}q))\big|_{0,\beta} \le K_1,$$

where K_1 is a positive constant not depending on the family. Hence (u_λ) is uniformly bounded in $C_0^{2,\beta}(\overline{\Omega})$, and in particular by Step 4 and the Ascoli–Arzelà theorem, it is relatively compact in $C_0^{1,\beta}(\overline{\Omega})$. This proves that (u_λ) is relatively compact in $C_0^{1,\beta}(\overline{\Omega})$ and we are done.

Step 6: There exists a constant $C > 0$ such that $|u|_{1,\beta} \le C$ for any $u \in C_0^{1,\beta}(\overline{\Omega})$ satisfying $u = tT_\epsilon u$, where $t \in [0,1]$.

We first observe that by definition and standard elliptic regularity theory any $u \in C_0^{1,\beta}(\overline{\Omega})$ satisfying $u = tT_\epsilon u$ is of class $C_0^{2,\beta}(\overline{\Omega})$ and satisfies $\widehat{Q}_\epsilon^t(u) = 0$. Thanks to Lemma 4.12, and since $\beta \le \alpha_0$, there exists $C > 0$ such that $|u|_{1,\beta} \le C$, provided that $|\nabla u|_{\infty,\partial\Omega} \le 1-\epsilon$. Therefore, in order to conclude, it is sufficient to check this boundary estimate for the gradient.

Let $q_0 \in \partial\Omega$ such that $|\nabla u(q_0)| = |\nabla u|_{\infty,\partial\Omega}$. If $\nabla u(q_0) = 0$, it follows that $\nabla u = 0$ on $\partial\Omega$ and hence there is nothing to prove. Therefore, let us assume that $\nabla u(q_0) \ne 0$.

Since $(\Omega, \widehat{H})$ is admissible, for any $t \in [0,1]$ there exist $\varphi_1, \varphi_2 \in C^2(\overline{\Omega})$ satisfying (i)–(iv) of Definition 4.4 at q_0. Hence, taking into account of the choice of ϵ and Remark 4.2, we have

$$\widehat{Q}_\epsilon^t(\varphi_1) \ge \widehat{Q}_\epsilon^t(u) \ge \widehat{Q}_\epsilon^t(\varphi_2) \quad \text{in } \Omega,$$

and $\varphi_1 \le u \le \varphi_2$ on $\partial\Omega$. Let us write

$$\widehat{Q}_\epsilon^t(u) = \sum_{i,j=1}^n a_\epsilon^{ij} u_{ij} + \hat{b}_{\epsilon,t}(q, u, \nabla u),$$

where

$$\hat{b}_{\epsilon,t}(q,u,\nabla u) := -nt(1-\eta_\epsilon^2|\nabla u|^2) + nt(1-\eta_\epsilon^2|\nabla u|^2)^{3/2} e^u \widehat{H}(e^u q).$$

Notice that thanks to assumption (ii) and Remark 4.9, it follows that for any fixed $q \in \Omega$ the map $z \mapsto e^z \widehat{H}(e^z q)$ is not increasing.

Thanks to Lemma 4.1 and Remark 4.3, under the hyperbolic stereographic projection $F : \mathbb{H}^n \to \mathbb{B}^n$, the operator $\widehat{Q}_\epsilon^t$ is transformed into a uniformly elliptic operator of the form

$$\widetilde{Q}_\epsilon^t \tilde{u} = \sum_{i,j=1}^n \tilde{a}_\epsilon^{ij}(y, \nabla_0 \tilde{u}) \tilde{u}_{ij} + \tilde{b}_{\epsilon,t}(y, \tilde{u}, \nabla_0 \tilde{u}),$$

where $y \in F(\Omega)$, $\nabla_0 \tilde{u}$ is the euclidean gradient, and $\tilde{u}_{ij}$ are the second partial derivatives of $\tilde{u} = u \circ F^{-1}$. In view of Remark 4.3 and assumption (ii) the principal part $\tilde{a}_\epsilon^{ij}(y,p)$ does not depend on z, and for each $(y,p) \in F(\Omega) \times \mathbb{R}^n$ the map $z \mapsto \tilde{b}_{\epsilon,t}(y,z,p)$ is nonincreasing. Hence the comparison principle applies, see [Gilbarg and Trudinger 1977, Theorem 10.1], and thus setting $\tilde{\varphi}_i := \varphi_i \circ F^{-1}$, for $i = 1,2$, from $\widetilde{Q}_\epsilon^t(\tilde{\varphi}_1) \ge \widetilde{Q}_\epsilon^t(\tilde{u}) \ge \widetilde{Q}_\epsilon^t(\tilde{\varphi}_2)$ in $F(\Omega)$, and $\tilde{\varphi}_1 \le \tilde{u} \le \tilde{\varphi}_2$ on $\partial F(\Omega)$, it follows that $\tilde{\varphi}_1 \le \tilde{u} \le \tilde{\varphi}_2$ in $F(\Omega)$. Therefore we obtain

$$\varphi_1 \le u \le \varphi_2 \quad \text{in } \Omega. \tag{5-8}$$

We observe that since $u = 0$ on $\partial\Omega$, we know $\nabla u(q_0)$ is orthogonal to $T_{q_0}\partial\Omega$, where $T_{q_0}\partial\Omega$ is the tangent space at q_0 for $\partial\Omega$, and we have the orthogonal decomposition $\mathrm{Span}\{\nabla u(q_0)\} \oplus T_{q_0}\partial\Omega = T_{q_0}\mathbb{H}^n$.

Let us set $\hat{w} := \nabla u(q_0)/|\nabla u(q_0)|$ and consider a curve $\psi : (-\delta, \delta) \to \mathbb{H}^n$ such that $\psi(0) = q_0$, $\psi(s) \in \Omega$ for $s \in (0, \delta)$ and $\psi'(0) = \hat{w}$ if $\hat{w}$ points towards the interior of Ω (otherwise we take $\psi'(0) = -\hat{w}$). Since Ω has a smooth boundary we can always find a curve satisfying these properties. From (5-8), and since $u(q_0) = \varphi_1(q_0) = \varphi_2(q_0) = 0$, we deduce that for all sufficiently small $h > 0$

$$\frac{\varphi_1(\psi(h)) - \varphi_1(\psi(0))}{h} \le \frac{u(\psi(h)) - u(\psi(0))}{h} \le \frac{\varphi_2(\psi(h)) - \varphi_2(\psi(0))}{h}. \tag{5-9}$$

Passing to the limit as $h \to 0^+$ we get

$$d\varphi_1(q_0)[\hat{w}] \le du(q_0)[\hat{w}] \le d\varphi_2(q_0)[\hat{w}]$$

(if $\hat{w}$ points in the opposite direction, (5-9) holds but with the reversed inequalities). Thus, it follows that

$$|du(q_0)[\hat{w}]| \le \max\{|d\varphi_1(q_0)[\hat{w}]|, |d\varphi_2(q_0)[\hat{w}]|\}.$$

Since $\mathbb{H}^n$ a spacelike hypersurface, for any $q \in \mathbb{H}^n$ the Cauchy–Schwarz inequality holds in $T_q\mathbb{H}^n$ for $\langle\cdot,\cdot\rangle_{T_q\mathbb{H}^n}$ (we point out that, in general, the Cauchy–Schwarz inequality does not hold in $\mathbb{L}^{n+1}$; see [López 2013]). In particular $|d\varphi_i(q_0)[\hat{w}]| = |\langle\nabla\varphi_i(q_0), \hat{w}\rangle| \le |\varphi_i(q_0)|\,|\hat{w}| = |\varphi_i(q_0)|$.

Hence, by the previous discussion and by Definition 4.4 we have

$$|\nabla u(q_0)| = |\langle\nabla u(q_0), \hat{w}\rangle| = |du(q_0)[\hat{w}]| \le \max\{|d\varphi_1[\hat{w}], d\varphi_2[\hat{w}]\} \le 1 - \theta.$$

Finally, since $\epsilon < \theta$ we get that $|\nabla u|$ satisfies the desired boundary estimate, and thus from the initial discussion, the proof of Step 6 is complete.

Step 7: Existence of a solution of problem (4-2). Thanks to Steps 5 and 6 it follows that the operator $T_\epsilon : C_0^{1,\beta}(\overline{\Omega}) \to C_0^{1,\beta}(\overline{\Omega})$ satisfies the hypotheses of the Leray–Schauder theorem, see [Gilbarg and Trudinger 1977, Theorem 11.3], and thus there exists $u \in C_0^{1,\beta}(\overline{\Omega})$ which solves $u = T_\epsilon u$. Hence, $u \in C_0^{2,\beta}(\overline{\Omega})$, and by the definition of $\widehat{H}$ and Lemma 4.11, u is a solution of problem (4-2). The proof of Step 7 is complete.

Step 8: Uniqueness.

For the uniqueness of the solution it is sufficient to argue as in [Caldiroli and Gullino 2013, Section 2.3]. For the sake of completeness we give a sketch of the proof.

Let us fix $\epsilon \in (0, 1)$ and let $u_1, u_2 \in C^2(\Omega) \cap C^0(\overline{\Omega})$ be two solutions of problem (4-2) such that the corresponding radial graphs are contained in $\mathcal{C}_{\overline{\Omega}}(r_1, r_2)$. If $u_1 \ne u_2$ then there exists $\bar{q} \in \Omega$ such that $u_1(\bar{q}) \ne u_2(\bar{q})$. Without loss of generality we can assume that $u_1(\bar{q}) < u_2(\bar{q})$. Then there exists $\mu > 0$ such that $u_1(q) + \mu \geqslant u_2(q)$ for every $q \in \Omega$ and $u_1(q_0) + \mu = u_2(q_0)$ at some $q_0 \in \Omega$. Set $\bar{u}_1 := u_1 + \mu$ and observe that $\bar{u}_1$ satisfies

$$\sum_{i,j=1}^{n} ((1 - \eta_\epsilon^2|\nabla\bar{u}_1|^2)\delta_{ij} + \eta_\epsilon^2\bar{u}_{1i}\bar{u}_{1j})\bar{u}_{1ij} \leqslant n(1 - |\eta_\epsilon^2\nabla\bar{u}_1|^2)(1 - \sqrt{1 - \eta_\epsilon^2|\nabla\bar{u}_1|^2}\,e^{\bar{u}_1}\widehat{H}(e^{\bar{u}_1}q)) \quad \text{in } \Omega$$

because of (ii) and $\mu > 0$. Notice that the radial graph defined by $\bar{u}_1$ stays over (in the radial direction) that one corresponding to u_2 and they intersect at the point $X_0 = q_0e^{u_2(q_0)}$. Now, in order to conclude, it sufficient to compare $\bar{u}_1$ and u_2 by means of the Hopf maximum principle. To this end we use the version stated in [Pucci and Serrin 2004, Theorem 2.3] for the operator

$$Q_\epsilon(u) = \sum_{i,j=1}^n ((1-\eta_\epsilon^2|\nabla u|^2)\delta_{ij} + u_i u_j)u_{ij} - n(1-\eta_\epsilon^2|\nabla u|^2)(1-\sqrt{1-\eta_\epsilon^2|\nabla u|^2}e^u \widehat{H}(e^u q)).$$

It is easy to see that, up to passing to hyperbolic stereographic coordinates, the assumptions of [Pucci and Serrin 2004, Theorem 2.3] are fulfilled, and applying the theorem as in [Caldiroli and Gullino 2013], we deduce that $\bar{u}_1 = u_2$ in Ω. But this gives a contradiction since $\bar{u}_1|_{\partial\Omega} = \mu > 0 = u_2|_{\partial\Omega}$. Hence it must be that $u_1 = u_2$ and we are done. □

6. An interior estimate for the gradient

In this section we prove an estimate for the gradient when the maximum point of its modulus lies in the interior of the domain Ω. We begin with a preliminary elementary result of linear algebra.

Lemma 6.1. *Let $A = (a_{ij})$, $B = (b_{ij}) \in \mathcal{M}_n(\mathbb{R})$ be two symmetric matrices. Assume that A is positive semidefinite and B is negative semidefinite. Then*

$$\sum_{i,j=1}^n a_{ij} b_{ij} \leq 0.$$

Proof. Since A, B are symmetric we have $\sum_{i,j=1}^n a_{ij} b_{ij} = \operatorname{trace}(AB)$, and there exist two invertible matrices P, Q such that $P^{-1}AP = D_A$ and $Q^{-1}BQ = D_B$ are diagonal. Thanks to the assumptions we have that D_A has nonnegative elements on the diagonal, while D_B has nonpositive elements on the diagonal. Therefore, since the trace is invariant under similitude, and diagonal matrices commute in the product, we have

$$\operatorname{trace}(AB) = \operatorname{trace}(P^{-1}APP^{-1}BP) = \operatorname{trace}(D_A P^{-1}BP) = \operatorname{trace}(P^{-1}D_A BP) = \operatorname{trace}(D_A B).$$

Now, by the same argument we get

$$\operatorname{trace}(D_A B) = \operatorname{trace}(Q^{-1}D_A Q Q^{-1}BQ) = \operatorname{trace}(D_A Q^{-1}QD_B) = \operatorname{trace}(D_A D_B).$$

Therefore, $\sum_{i,j=1}^n a_{ij} b_{ij} = \operatorname{trace}(AB) = \operatorname{trace}(D_A D_B) \leq 0$, and the proof is complete. □

Definition 6.2. Let $H \in C^1(\mathcal{C}_{\overline{\Omega}})$, and let $\nabla_0 H$ be gradient of H in $\mathbb{R}^{n+1}$ with respect to the flat metric. We define the (euclidean) tangential component of $\nabla_0 H$ on $T_{x/|x|}\mathbb{H}^n$ as the vector

$$\nabla_0^T H(x) := \nabla_0 H(x) - (\nabla_0 H(x), \hat{r}(x))_{n+1} \hat{r}(x), \quad x \in \mathcal{C}_{\overline{\Omega}},$$

where

$$\hat{r}(x) := \frac{(x_1, \ldots, x_n, -x_{n+1})}{\|x\|_{n+1}}.$$

Remark 6.3. We point out that by definition $\nabla_0^T H(x) = \nabla_0 H(x) - \langle \nabla_0 H(x), x/\|x\|_{n+1}\rangle \hat{r}(x)$, and if $v \in \mathbb{R}^{n+1}$ is such that $(\hat{r}(x), v)_{n+1} = 0$ then $\langle x/|x|, v\rangle = 0$, and vice versa. In particular $(\hat{r}(x), v)_{n+1} = 0$ for any $v \in T_{x/|x|}\mathbb{H}^n$, $x \in \mathcal{C}_{\overline{\Omega}}$.

In the sequel we will make use also of the following formulas for the second and third covariant derivatives of a smooth function u defined over Ω, see [Yau 1975, Section 2]:

$$\sum_{j=1}^{n} u_{ij}\omega^j = du_i - \sum_{j=1}^{n} u_j \omega_{ji}, \tag{6-1}$$

$$\sum_{k=1}^{n} u_{ijk}\omega^k = du_{ij} - \sum_{k=1}^{n} u_{kj}\omega_{ki} - \sum_{k=1}^{n} u_{ik}\omega_{kj}. \tag{6-2}$$

Proposition 6.4. *Let Ω be a bounded domain of $\mathbb{H}^n$, let $H \in C^1(\mathcal{C}_{\overline{\Omega}})$, $\epsilon \in (0,1)$, and let $u \in C^3(\overline{\Omega})$ be a solution of*

$$\sum_{i,j=1}^{n} ((1-\eta_\epsilon^2|\nabla u|^2)\delta_{ij} + \eta_\epsilon^2 u_i u_j)u_{ij} = n(1-\eta_\epsilon^2|\nabla u|^2)(1-\sqrt{1-\eta_\epsilon^2|\nabla u|^2}\,e^u H(e^u q)). \tag{6-3}$$

Then, if the maximum point q_0 of $|\nabla u|$ lies in the interior of Ω, we have $|\nabla u(q_0)| = 0$ or

$$\Big[-(n-1) - n(1-\eta_\epsilon^2|\nabla u(q_0)|^2)^{1/2} e^{u(q_0)} \frac{\partial}{\partial\lambda}(\lambda H(\lambda q))\Big|_{\lambda=e^{u(q_0)}}\Big]|\nabla u(q_0)| \\ - n^{3/2}(1-\eta_\epsilon^2|\nabla u(q_0)|^2)^{1/2} e^{2u(q_0)} \|\nabla_0^T H(e^{u(q_0)}q_0)\|_{n+1} \le 0, \tag{6-4}$$

where $\nabla_0^T H$ is the (euclidean) tangential component of $(\nabla_0 H)(e^{u(q_0)}q_0)$ on $T_{q_0}\mathbb{H}^n$.

Proof. We will prove a more general version of (6-4). Let us fix a smooth positive function $f : \mathbb{R} \to \mathbb{R}^+$ and consider the auxiliary function $\varphi := f(2Cu)|\nabla u|^2$, where $C \in \mathbb{R}$ is a fixed constant. In order to simplify the notation we set $v := |\nabla u|^2$; hence $\varphi = f(2Cu)v$. Assume that φ has a maximum point at some q_0 lying in the interior of Ω. Hence $\nabla\varphi(q_0) = 0$ and the Hessian $(\varphi_{ij}(q_0))$ is negative semidefinite.

By direct computation we have $v_i = 2\sum_{j=1}^n u_j u_{ji}$ and from $\nabla\varphi(q_0) = 0$ we get

$$\sum_{h=1}^{n} f(2Cu)u_h u_{hi} + f'(2Cu)Cvu_i = 0 \quad \text{for all } i = 1, \dots, n, \tag{6-5}$$

which implies

$$\sum_{i,h=1}^{n} u_i u_{ih} u_h = -C\frac{f'}{f}v^2, \tag{6-6}$$

where, f, f' stand, respectively, for $f(2Cu)$, $f'(2Cu)$. By a simple computation, from (6-6), we get

$$\sum_{i,h,k=1}^{n} u_i u_{ih} u_{hk} u_k = \left(\frac{f'}{f}\right)^2 C^2 v^3. \tag{6-7}$$

Let us set

$$a_\epsilon^{ij} := ((1-\eta_\epsilon^2|\nabla u|^2)\delta_{ij} + \eta_\epsilon^2 u_i u_j),$$
$$b_\epsilon := n(1-\eta_\epsilon^2|\nabla u|^2)(1-\sqrt{1-\eta_\epsilon^2|\nabla u|^2}\,e^u H(e^u q)).$$

Since (a_ϵ^{ij}) is a positive definite symmetric matrix and $(\varphi_{ij}(q_0))$ is symmetric negative semidefinite, from Lemma 6.1 it follows that

$$\sum_{i,j=1}^{n} a_\epsilon^{ij}\varphi_{ij}(q_0) \le 0. \tag{6-8}$$

In order to get an estimate for $v = |\nabla u|^2$, the idea is to use (6-8). To this end we compute explicitly $\varphi_{ij}(q_0)$. Recalling that

$$\varphi_i = 2\biggl(\sum_{h=1}^{n} f(2Cu)u_h u_{hi} + f'(2Cu)Cvu_i\biggr),$$

and using (6-5), (6-1) we have

$$\sum_{j=1}^{n}\varphi_{ij}(q_0)\omega^j = 2\sum_{j=1}^{n}\biggl[\sum_{h=1}^{n}(2Cf'u_ju_hu_{hi}+fu_{hj}u_{hi} + fu_{hij}u_h)+2C^2f''u_ju_iv+Cf'u_{ij}v+\sum_{h=1}^{n}2Cf'u_iu_hu_{hj}\biggr]\omega^j,$$

and thus from (6-8) we infer that

$$\begin{aligned}2\biggl[\sum_{i,h=1}^{n}\bigl(4Cf'(1-\eta_\epsilon^2v)u_iu_hu_{hi}+f(1-\eta_\epsilon^2v)u_{hi}^2+f(1-\eta_\epsilon^2v)u_{hii}u_h\bigr)\\ +2C^2f''(1-\eta_\epsilon^2v)v^2+\sum_{i=1}^{n}Cf'(1-\eta_\epsilon^2v)u_{ii}v\\ +\sum_{i,h=1}^{n}2Cf'\eta_\epsilon^2u_iu_hu_{hi}+\sum_{i,j,h=1}^{n}(f\eta_\epsilon^2u_iu_ju_{hj}u_{hi}+f\eta_\epsilon^2u_iu_ju_hu_{hij})\\ +2C^2f''\eta v^3+\sum_{i,j=1}^{n}Cf'\eta_\epsilon^2u_iu_ju_{ij}v+\sum_{j,h=1}^{n}2Cf'\eta_\epsilon^2u_ju_hu_{hj}v\biggr]\le 0.\end{aligned} \tag{6-9}$$

Now we estimate and rewrite the terms involving the second and third covariant derivatives. We may choose a coordinate frame at q_0 satisfying $\delta_{1i}v^{1/2} = u_i$. If $v(q_0) = 0$, then $\max_{\overline{\Omega}}\varphi = \varphi(q_0) = 0$ and the thesis follows immediately. Otherwise in these coordinates, from (6-5), it follows that

$$u_{11} = -\frac{f'}{f}Cv, \tag{6-10}$$

which implies

$$\sum_{i,j=1}^{n} u_{ij}u_{ij} \ge \left(\frac{f'}{f}\right)^2 C^2v^2. \tag{6-11}$$

Since u is a solution of (6-3), computing at q_0 in these coordinates we have

$$\sum_{i=1}^{n}(1-\eta_\epsilon^2v)u_{ii} + \eta_\epsilon^2vu_{11} = b_\epsilon,$$

and from (6-10) we obtain

$$\sum_{i=1}^{n} u_{ii} = \left(b_\epsilon + C\frac{f'}{f}\eta_\epsilon^2 v^2\right)(1-\eta_\epsilon^2 v)^{-1}. \tag{6-12}$$

Recalling that $\eta_\epsilon(|\nabla u|) = \eta_\epsilon(v^{1/2})$, by direct computation we infer

$$\nabla_k(\eta_\epsilon^2 v) = \sum_{h=1}^{n}\left(2\eta_\epsilon\eta_\epsilon' v^{1/2} u_h u_{hk} + 2\eta_\epsilon^2 u_h u_{hk}\right), \tag{6-13}$$

where it is understood that η_ϵ' stands for $\eta_\epsilon'(v^{1/2})$. By differentiating (6-3), taking into account (6-1),(6-2) and (6-13), after some standard computations we deduce that

$$\sum_{k=1}^{n}\Bigg[-\sum_{i,h=1}^{n}(2\eta_\epsilon\eta_\epsilon' v^{1/2}u_h u_{hk} + 2\eta_\epsilon^2 u_h u_{hk})u_{ii} + \sum_{i,j,h=1}^{n} 2\eta_\epsilon\eta_\epsilon' v^{-1/2} u_i u_j u_h u_{hk} u_{ij} \\ + \sum_{i,j=1}^{n}\left(\eta_\epsilon^2 u_j u_{ik} u_{ij} + \eta_\epsilon^2 u_i u_{jk} u_{ij} + ((1-\eta_\epsilon^2 v)\delta_{ij} + \eta_\epsilon^2 u_i u_j)u_{ijk}\right)\Bigg]\omega^k = \sum_{k=1}^{n}(b_\epsilon)_k\omega^k.$$

Now, contracting the equation with u_k, we get

$$\sum_{i,h,k=1}^{n}(-2\eta_\epsilon\eta_\epsilon' v^{1/2} u_h u_k u_{hk} u_{ii} - 2\eta_\epsilon^2 u_h u_k u_{hk} u_{ii}) \\ + \sum_{i,j,h,k=1}^{n} 2\eta_\epsilon\eta_\epsilon' v^{-1/2} u_i u_j u_k u_h u_{hk} u_{ij} \\ + \sum_{i,j,k=1}^{n}(\eta_\epsilon^2 u_j u_k u_{ik} u_{ij} + \eta_\epsilon^2 u_i u_k u_{jk} u_{ij}) \\ + \sum_{i,k=1}^{n}(1-\eta_\epsilon^2 v)u_k u_{iik} + \sum_{i,j,k=1}^{n}\eta_\epsilon^2 u_i u_j u_k u_{ijk} = \sum_{k=1}^{n}(b_\epsilon)_k u_k. \tag{6-14}$$

Since the Ricci curvature of the hyperbolic space is $R_{ij} = -(n-1)\delta_{ij}$, the Ricci formula, see formula (2.11) in [Yau 1975], gives

$$\sum_{k=1}^{n} u_k u_{kii} = \sum_{k=1}^{n} u_k u_{iik} - (n-1)v. \tag{6-15}$$

Hence, using (6-15), and taking into account (6-6), (6-7),(6-12), we rewrite (6-14) as

$$\sum_{i,k=1}^{n}(1-\eta_\epsilon^2 v)u_k u_{kii} + \sum_{i,j,k=1}^{n}\eta_\epsilon^2 u_i u_j u_k u_{ijk} \\ = -(n-1)v(1-\eta_\epsilon^2 v) - 2\eta_\epsilon\eta_\epsilon' C\frac{f'}{f}v^{5/2}(1-\eta_\epsilon^2 v)^{-1}\left(b_\epsilon + C\frac{f'}{f}\eta_\epsilon^2 v^2\right) \\ -2\eta_\epsilon^2 C v^2(1-\eta_\epsilon^2 v)^{-1}(b_\epsilon + C\eta_\epsilon^2 v^2) \\ -2\eta_\epsilon\eta_\epsilon' C^2\left(\frac{f'}{f}\right)^2 v^{7/2} - 2\eta_\epsilon^2 C^2\left(\frac{f'}{f}\right)^2 v^3 + \sum_{k=1}^{n}(b_\epsilon)_k u_k. \tag{6-16}$$

Now, from (6-9) and by using (6-6), (6-7), (6-11), (6-12) (6-16), we deduce that

$$
\begin{aligned}
-4C^2(1-\eta_\epsilon^2 v)\frac{(f')^2}{f}v^2&+C^2(1-\eta_\epsilon^2 v)\frac{(f')^2}{f}v^2\\
&+f\Big[-(n-1)v(1-\eta_\epsilon^2 v)-2\eta_\epsilon\eta_\epsilon' C\frac{f'}{f}v^{5/2}(1-\eta_\epsilon^2 v)^{-1}\Big(b_\epsilon+C\frac{f'}{f}\eta_\epsilon^2 v^2\Big)\\
&\qquad-2C\eta_\epsilon^2 v^2(1-\eta_\epsilon^2 v)^{-1}\Big(b_\epsilon+C\frac{f'}{f}\eta_\epsilon^2 v^2\Big)\\
&\qquad\qquad-2C^2\eta_\epsilon\eta_\epsilon'\Big(\frac{f'}{f}\Big)^2 v^{7/2}-2C^2\Big(\frac{f'}{f}\Big)^2 v^3+\sum_{k=1}^n (b_\epsilon)_k u_k\Big]\\
&\quad+2C^2 f''(1-\eta_\epsilon^2 v)v^2+Cf'v\Big(b_\epsilon+C\frac{f'}{f}\eta_\epsilon^2 v^2\Big)\\
&\qquad-2C^2\frac{(f')^2}{f}\eta_\epsilon^2 v^2+C^2\frac{(f')^2}{f}\eta_\epsilon^2 v^3+2C^2 f''\eta_\epsilon^2 v^3\\
&\qquad\qquad-C^2\frac{(f')^2}{f}\eta_\epsilon^2 v^3-2C^2\frac{(f')^2}{f}\eta_\epsilon^2 v^3\le 0. \qquad (6\text{-}17)
\end{aligned}
$$

Now we compute the term $\sum_{k=1}^n (b_\epsilon)_k u_k$. To this end let us observe that

$$
\begin{aligned}
&(e^u H(e^u q))_k(q_0)\\
&\quad= e^{u(q_0)}u_k H(e^{u(q_0)}q_0)+e^{2u(q_0)}u_k(q_0)(\nabla_0 H)(e^{u(q_0)}q_0)\cdot q_0+e^{2u(q_0)}(\nabla_0 H)(e^{u(q_0)}q_0)\cdot e_k(q_0),
\end{aligned}
$$

where $\cdot$ denotes the standard euclidean product of $\mathbb{R}^{n+1}$ and $\nabla_0 H$ is the gradient of H with respect to the flat metric in $\mathbb{R}^{n+1}$. Then, after some computations and taking into account (6-13), we get

$$
\begin{aligned}
\sum_{k=1}^n (b_\epsilon)_k u_k(q_0) &= 2nC\frac{f'}{f}\eta_\epsilon\eta_\epsilon' v^{5/2}(1-\sqrt{1-\eta_\epsilon^2 v}\; e^{u(q_0)}H(e^{u(q_0)}q_0))\\
&\quad+2nC\frac{f'}{f}(\eta_\epsilon^2 v)v(1-\sqrt{1-\eta_\epsilon^2 v}\; e^{u(q_0)}H(e^{u(q_0)}q_0))\\
&\quad-nC\frac{f'}{f}\eta_\epsilon\eta_\epsilon'\sqrt{1-\eta_\epsilon^2 v}(e^{u(q_0)}H(e^{u(q_0)}q_0))v^2\\
&\quad-nC\frac{f'}{f}(\eta_\epsilon^2 v)(1-\eta_\epsilon^2 v)(e^{u(q_0)}H(e^{u(q_0)}q_0))v\\
&\quad-n(1-\eta_\epsilon^2 v)^{3/2}e^{u(q_0)}\big(e^{u(q_0)}H(e^{u(q_0)}q_0)+e^{2u(q_0)}\nabla_0 H(e^{u(q_0)}q_0)\cdot q_0\big)v\\
&\quad-n(1-\eta_\epsilon^2 v)^{3/2}e^{2u(q_0)}\nabla_0 H(e^{u(q_0)}q_0)\cdot\nabla u, \qquad (6\text{-}18)
\end{aligned}
$$

Computing $(\partial/\partial\lambda)(\lambda H(\lambda q_0))|_{\lambda=e^{u(q_0)}}$ and taking into account Remark 6.3, we rewrite the last two terms of (6-18) as

$$
-n(1-\eta_\epsilon^2 v)^{3/2}e^{u(q_0)}\frac{\partial}{\partial\lambda}(\lambda H(\lambda q_0))\Big|_{\lambda=e^{u(q_0)}}v-n(1-\eta_\epsilon^2 v)^{3/2}e^{2u(q_0)}\nabla_0^T H(e^{u(q_0)}q_0)\cdot\nabla u, \qquad (6\text{-}19)
$$

where $\nabla_0^T H$ is the euclidean tangential component of $(\nabla_0 H)$ on $T_{q_0}\mathbb{H}^n$. Hence, from (6-17), (6-18), (6-19), regrouping and simplifying terms, we get

$$\begin{aligned}
&-2C^2(f')^2[1-\eta_\epsilon^2+\eta_\epsilon'(\eta_\epsilon v^{1/2})]v^3\\
&\quad -C^2[(f')^2(3-4\eta_\epsilon^2 v+(\eta_\epsilon^2 v))+2f'f(\eta_\epsilon^2 v)^2-2f''f(1-\eta_\epsilon^2 v)]v^2\\
&\quad\quad -nC(1-\eta_\epsilon^2 v)^{3/2}f'f\eta_\epsilon'(\eta_\epsilon v^{1/2})e^{u(q_0)}H(e^{u(q_0)}q_0)v^{3/2}\\
&\quad\quad\quad +\Big[-(n-1)f^2(1-\eta_\epsilon^2 v)^2-C^2(f')^2f(\eta_\epsilon^2 v)(1-\eta_\epsilon^2 v)\\
&\quad\quad\quad\quad -nCff'(1-\eta_\epsilon^2 v)^2(\eta_\epsilon^2 v)e^{u(q_0)}H(e^{u(q_0)}q_0)\\
&\quad\quad\quad\quad\quad -nf^2(1-\eta_\epsilon^2 v)^{5/2}e^{u(q_0)}\frac{\partial}{\partial\lambda}(\lambda H(\lambda q))\Big|_{\lambda=e^{u(q_0)}}\Big]v\\
&\quad\quad\quad -nf^2(1-\eta_\epsilon^2 v)^{5/2}e^{2u(q_0)}\nabla_0^T H(e^{u(q_0)}q_0)\cdot\nabla u\\
&\quad\quad\quad\quad +nCf'f(1-\eta_\epsilon^2 v)^2(1-\sqrt{1-\eta_\epsilon^2 v}\;e^{u(q_0)}H(e^{u(q_0)}q_0))\le 0,
\end{aligned}\tag{6-20}$$

and the proof of the general inequality is complete. Now we prove (6-4). Taking $f\equiv 1$, and dividing (6-20) by $(1-\eta_\epsilon^2 v)^2$, we get

$$\begin{aligned}
&\Big[-(n-1)-n(1-\eta_\epsilon^2 v)^{1/2}e^{u(q_0)}\frac{\partial}{\partial\lambda}(\lambda H(\lambda q))\Big|_{\lambda=e^{u(q_0)}}\Big]v\\
&\qquad\qquad -n(1-\eta_\epsilon^2 v)^{1/2}e^{2u(q_0)}\nabla_0^T H(e^{u(q_0)}q_0)\cdot\nabla u\;\le 0.
\end{aligned}\tag{6-21}$$

Assume that $v(q_0)\neq 0$ (otherwise $v\equiv 0$ and there is nothing to prove). To conclude the proof it remains to estimate the term $\nabla_0^T H(e^{u(q_0)}q_0)\cdot\nabla u$. To this end, recalling the notation used in the proof of Lemma 4.1, we define $\tilde h\in\mathbb{R}^n$ as the vector whose i-th component is

$$\tilde h_i := \nabla_0^T H(e^{u(q_0)}q_0)\cdot\frac{\partial_i}{\|\partial_i\|_{n+1}},$$

$i=1,\dots,n$, where $\partial_i=(\partial\phi/\partial y_i)(F(q_0))$. Then, by construction and the Cauchy–Schwarz inequality we have

$$\|\tilde h\|_n^2=\sum_{i=1}^n\left(\nabla_0^T H(e^{u(q_0)}q_0)\cdot\frac{\partial_i}{\|\partial_i\|_{n+1}}\right)^2\le n\|\nabla_0^T H(e^{u(q_0)}q_0)\|_{n+1}^2.\tag{6-22}$$

Now, exploiting (4-6) we have

$$\nabla_0^T H(e^{u(q_0)}q_0)\cdot\nabla u=\lambda^{-2}\sum_{i=1}^n\frac{\partial\tilde u}{\partial y_i}\nabla_0^T H(e^{u(q_0)}q_0)\cdot\partial_i=\lambda^{-1}\sum_{i=1}^n\frac{\partial\tilde u}{\partial y_i}\tilde h_i=\lambda^{-1}(\nabla_0\tilde u,\tilde h)_n,$$

and thus, from (4-7), (6-22) we deduce that

$$|\nabla_0^T H(e^{u(q_0)}q_0)\cdot\nabla u|=\lambda^{-1}|(\nabla_0\tilde u,\tilde h)_n|\le\lambda^{-1}\|\nabla_0\tilde u\|_n\|\tilde h\|_n\le\sqrt{n}|\nabla u|\|\nabla_0^T H(e^{u(q_0)}q_0)\|_{n+1}.\tag{6-23}$$

Finally, combining (6-21), (6-23) and dividing by $v^{1/2}$, we obtain (6-4). □

Remark 6.5. Applying the gradient estimate (6-4) to the solutions of $Q_\epsilon^t(u)=0$, we obtain

$$\Big[-(n-1)-nt(1-\eta_\epsilon^2|\nabla u(q_0)|^2)^{1/2}e^{u(q_0)}\frac{\partial}{\partial\lambda}(\lambda H(\lambda q))\Big|_{\lambda=e^{u(q_0)}}\Big]|\nabla u(q_0)| \\ -n^{3/2}t(1-\eta_\epsilon^2|\nabla u(q_0)|^2)^{1/2}e^{2u(q_0)}\|\nabla_0^T H(e^{u(q_0)}q_0)\|_{n+1}\le 0. \quad (6\text{-}24)$$

Hence, it is not possible, by using only this strategy, to get a uniform bound with respect to t for $|\nabla u|_\infty$ as in [Treibergs and Wei 1983]. In fact here we deal with functions defined on a manifold with negative Ricci curvature, and thus in (6-24) we have a term $-(n-1)$, while in [loc. cit.], for the sphere, this term has the opposite sign. We also point out that this trouble does not depend on the choice of the auxiliary function in the proof of Proposition 6.4, as shown by (6-20), where the leading term v^3 has a negative coefficient.

7. Proof of Theorem 1.5

Proof of Theorem 1.5. We first observe that by definition (Ω, H) is admissible with constant θ, and thus in the proof of Theorem 5.1 we can take $\bar\epsilon=\theta$. Therefore, for any $\epsilon\in(0,\theta)$, there exists a solution u_ϵ of the regularized problem (4-2). Let us choose $\epsilon\in(0,\theta)$ sufficiently close to θ so that

$$\frac{\partial}{\partial\lambda}(\lambda H(\lambda q)) < -\frac{1}{r_1(\epsilon-\epsilon^2/4)^{1/2}} \quad \text{for all } q\in\overline\Omega,\ \lambda\in[r_1,r_2], \quad (7\text{-}1)$$

$$\|\nabla_0^T H(X)\|_{n+1} < \frac{1-\epsilon}{n^{3/2}r_2^2}, \qquad X\in\mathcal{C}_{\overline\Omega}(r_1,r_2), \quad (7\text{-}2)$$

and let u be the solution of the regularized problem (4-2).

Let $q_0\in\overline\Omega$ be the maximum point of $|\nabla u|$, and set $v=|\nabla u(q_0)|^2$. There are only two possibilities: $v<(1-\epsilon)^2$ or $v\ge(1-\epsilon)^2$. In the first case there is nothing to prove; in fact, by definition of η_ϵ we have that u is a solution of problem (1-1) and we are done. Therefore let us assume that $v\ge(1-\epsilon)^2$. We point out that in this case q_0 cannot belong to $\partial\Omega$ because by Step 6 of the proof of Theorem 5.1 and since $\epsilon<\theta$, we have

$$\sup_{\partial\Omega}|\nabla u|^2\le(1-\theta)^2<(1-\epsilon)^2.$$

Hence $q_0\in\Omega$. We also observe that $u\in C^{3,\beta}(\overline\Omega)$, for some $\beta\in(0,\alpha]$. In fact, by Theorem 5.1 we know that $u\in C_0^{2,\beta}(\overline\Omega)$. Thanks to Lemma 4.11 we know that $\Sigma(u)$ is contained in $\mathcal{C}_{\overline\Omega}(r_1,r_2)$ and since $H\in C^{1,\alpha}(\mathcal{C}_{\overline\Omega}(r_1,r_2))$, $\partial\Omega\in C^{3,\alpha}$, by standard regularity results, see [Gilbarg and Trudinger 1977], we get $u\in C^{3,\beta}(\overline\Omega)$. Therefore, we can apply Proposition 6.4 and recalling that by definition

$$1-\eta_\epsilon^2 v=1-\eta_\epsilon^2(|\nabla u(q_0)|)|\nabla u(q_0)|^2,$$

we have

$$\Big[-(n-1)-n(1-\eta_\epsilon^2 v)^{1/2}e^{u(q_0)}\frac{\partial}{\partial\lambda}(\lambda H(\lambda q))\Big|_{\lambda=e^{u(q_0)}}\Big]v^{1/2} \\ -n^{3/2}(1-\eta_\epsilon^2 v)^{1/2}e^{2u(q_0)}\|\nabla_0^T H(e^{u(q_0)}q_0)\|_{n+1}\le 0, \quad (7\text{-}3)$$

but on the other hand, since we are assuming that $v \geq (1-\epsilon)^2$, by the definition of η_ϵ we have $1 \geq 1-\eta_\epsilon^2 v \geq 1-\left(1-\frac{1}{2}\epsilon\right)^2$, and in view of (7-1), (7-2) we have

$$\begin{aligned}\Big[-(n-1)-n(1-\eta_\epsilon^2 v)^{1/2}e^{u(q_0)}\frac{\partial}{\partial\lambda}(\lambda H(\lambda q))\Big|_{\lambda=e^{u(q_0)}}\Big]v^{1/2}\\ -n^{3/2}(1-\eta_\epsilon^2 v)^{1/2}e^{2u(q_0)}\|\nabla^T H(e^{u(q_0)}q_0)\|_{n+1}\\ > \Big[-(n-1)+n\frac{e^{u(q_0)}}{r_1}\frac{(1-\eta_\epsilon^2 v)^{1/2}}{\left(\epsilon-\frac{1}{4}\epsilon^2\right)^{1/2}}\Big]v^{1/2}-\frac{e^{2u(q_0)}}{r_2^2}(1-\epsilon)\\ \geq \Big[-(n-1)+n\frac{\left(\epsilon-\frac{1}{4}\epsilon^2\right)^{1/2}}{\left(\epsilon-\frac{1}{4}\epsilon^2\right)^{1/2}}\Big]v^{1/2}-(1-\epsilon)\\ \geq (1-\epsilon)-(1-\epsilon)=0\end{aligned}$$

and thus we contradict (7-3). Therefore the only possibility is $v<(1-\epsilon)^2$, and by the definition of η_ϵ this means that u is a solution of problem (1-1). Moreover, as proved in Theorem 5.1, such a solution is the unique solution whose associated radial graph is contained in $\mathcal{C}_{\overline{\Omega}}(r_1,r_2)$, and this completes the proof. □

8. A finer gradient estimate

In this section we prove an a priori estimate for the gradient of the solutions of

$$\begin{cases}-\operatorname{div}_{\mathbb{H}^n}(\nabla u/\sqrt{1-|\nabla u|^2})+nt/\sqrt{1-|\nabla u|^2}=nte^u H(e^u q) & \text{in } \Omega,\\ |\nabla u|<1 & \text{in } \Omega,\\ u=0 & \text{on } \partial\Omega,\end{cases} \tag{8-1}$$

where $t\in[0,1]$. As in Section 4 we introduce the function $\nu=1/\sqrt{1-|\nabla u|^2}$.

Proposition 8.1. *Let Ω be a bounded domain of $\mathbb{H}^n$, let $H\in C^1(\mathcal{C}_{\overline{\Omega}})$, let $r_1,r_2\in\mathbb{R}$ be such that $r_1\neq r_2$, $0<r_1\leq 1\leq r_2$, and let $\nu_0>0$ be a positive number. Then, there exists a constant $C=C(r_1,r_2,\nu_0,\Omega,H)>0$ such that for any $t\in[0,1]$, for any solution $u\in C^3(\overline{\Omega})$ of* (8-1) *satisfying $\log r_1\leq u\leq \log r_2$ and $\sup_{\partial\Omega}\nu\leq\nu_0$, we have*

$$\sup_{\Omega}\nu\leq C.$$

Proof. Let $u\in C^3(\overline{\Omega})$ be a solution of (8-1) satisfying $\log r_1\leq u\leq\log r_2$ and $\sup_{\partial\Omega}\nu\leq\nu_0$. Clearly $\nu\in C^0(\overline{\Omega})$ and we can introduce the differential operator $P_u: C^1(\overline{\Omega})\to C^0(\overline{\Omega})$ defined by

$$P_u w := \nu\sum_{k=1}^n u_k w_k,$$

where u_k, w_k are the covariant derivatives with respect to a orthonormal frame field. Applying P_u to both sides of the equation in (8-1) and arguing as in Proposition 6.4 we deduce that ν satisfies the equation

$$\begin{aligned}\sum_{i,j=1}^n\nabla_i(\nu^{-2}f_{ij}\nu_j)+\nu^{-2}|\nabla\nu|^2+|\langle\nabla u,\nabla\nu\rangle|^2\\ +\nu\sum_{i,j,k=1}^n f_{ij}u_{jk}u_{ik}+\sum_{i,j=1}^n\nu^2R_{ij}u_iu_j+nt\nu\langle\nabla u,\nabla\nu\rangle=\nu nt\sum_{k=1}^n u_k\nabla_k(e^uH(e^uq)),\end{aligned} \tag{8-2}$$

where $f_{ij} := \nu\delta_{ij} + \nu^3 u_i u_j$ and $R_{ij} = -(n-1)\delta_{ij}$ is the Ricci curvature tensor of $\mathbb{H}^n$, $i, j = 1, \dots, n$. This relation resembles that appearing in [Gerhardt 1983, (4.8)], and it can be proved by direct computation taking into account of the identities $\nu^{-2} = 1 - |\nabla u|^2$, $\nu_i = \nu^3 \sum_{l=1}^n u_l u_{li}$, $\langle \nabla u, \nabla \nu\rangle = \nu^3 \sum_{l,m=1}^n u_l u_m u_{lm}$, $|\nabla \nu|^2 = \nu^6 \sum_{i=1}^n \big(\sum_{l=1}^n u_l u_{li}\big)^2$, and [Yau 1975, (2.6)]. In order to estimate the terms appearing in (8-2) we first observe that

$$\sum_{i,j,k=1}^n f_{ij} u_{jk} u_{ik} = \sum_{i,j,k=1}^n (\nu\delta_{ij} + \nu^3 u_i u_j) u_{jk} u_{ik}$$
$$= \nu \sum_{i,k=1}^n u_{ik}^2 + \nu^3 \sum_{k=1}^n \bigg(\sum_{i=1}^n u_i u_{ik}\bigg)^2 \geq \nu |D^2 u|^2, \tag{8-3}$$

where $|D^2 u|^2 := \sum_{i,k=1}^n u_{ik}^2$ is the square of the matrix norm of the Hessian. For the term $nt\langle \nabla u, \nabla \nu\rangle$, we write the equation in (8-1) in nondivergence form as

$$-\nu \Delta u - \langle \nabla u, \nabla \nu\rangle + nt\nu = nte^u H(e^u q). \tag{8-4}$$

Then, multiplying each side by $nt\nu$, recalling that $\nu \geq 1$, $e^u H(e^u q)$ is uniformly bounded with respect to t, and using the inequality $|\Delta u| \leq \sqrt{n}|D^2 u|$ we deduce that

$$|nt\nu\langle \nabla u, \nabla \nu\rangle| \leq c_1 \nu^2 (1 + |\Delta u|) \leq c_2 \nu^2 (1 + |D^2 u|)$$

for some constants $c_1, c_2 > 0$ depending on n, r_1, r_2 and $\|H\|_{\infty, C_{\overline{\Omega}}}$, but not on t. From now on c_3, c_4, etc. will denote positive constants which do not depend on t. Now, if $|D^2 u| < c_2(1 + \sqrt{1 + 1/c_2})$, we get $|nt\nu\langle \nabla u, \nabla \nu\rangle| \leq c_3 \nu^2$, where c_3 depends just on c_2, and thus $nt\nu\langle \nabla u, \nabla \nu\rangle \geq -c_3 \nu^2$. On the other hand, if $|D^2 u| \geq c_2(1 + \sqrt{1 + 1/c_2})$, by an elementary computation we infer that

$$-c_2 \nu^2 (1 + |D^2 u|) + \tfrac{1}{2}\nu^2 |D^2 u|^2 \geq 0.$$

Hence, in view of (8-3) and the previous inequalities we obtain

$$\sum_{i,j,k=1}^n f_{ij} u_{jk} u_{ik} + nt\nu\langle \nabla u, \nabla \nu\rangle \geq -c_4 \nu^2 + \tfrac{1}{2}\nu^2 |D^2 u|^2. \tag{8-5}$$

Therefore, from (8-2), (8-5) we have

$$-\sum_{i,j=1}^n \nabla_i (\nu^{-2} f_{ij} \nu_j) + |\langle \nabla u, \nabla \nu\rangle|^2 + \tfrac{1}{2}\nu^2 |D^2 u|^2 \leq c_5 \nu^2 + \nu nt \sum_{k=1}^n u_k \nabla_k (e^u H(e^u q)). \tag{8-6}$$

Now, writing (8-4) as $-\nu\Delta u - \langle \nabla u, \nabla \nu\rangle = nte^u H(e^u q) - nt\nu$ and squaring, by using elementary inequalities we get

$$\nu^2 |\Delta u|^2 - 2\nu |\Delta u| |\langle \nabla u, \nabla \nu\rangle| + |\langle \nabla u, \nabla \nu\rangle|^2 \leq 2n^2 e^{2u} H^2(e^u q) + 2n^2 \nu^2. \tag{8-7}$$

Multiplying (8-4) by ν, and using $|\Delta u| \leq \sqrt{n}|D^2 u|$, we deduce that

$$\nu |\langle \nabla u, \nabla \nu\rangle| \leq c_5 \nu^2 (1 + |D^2 u|).$$

Hence, from this, using again $|\Delta u| \le \sqrt{n}|D^2u|$, and (8-7) we obtain

$$-n\nu^2|D^2u|^2 - 2\sqrt{n}c_5\nu^2(1+|D^2u|) + |\langle \nabla u, \nabla \nu\rangle|^2 \le 2n^2e^{2u}H^2(e^uq) + 2n^2\nu^2,$$

and thus by elementary computations we deduce that

$$-c_6\nu^2|D^2u|^2 + |\langle \nabla u, \nabla \nu\rangle|^2 \le 2n^2e^{2u}H^2(e^uq) + c_7\nu^2. \tag{8-8}$$

Therefore, dividing (8-8) by $C := 2c_6+1$ and summing with (8-6) we deduce

$$-\sum_{i,j=1}^n \nabla_i(\nu^{-2}f_{ij}\nu_j) + (1+2c_*)|\langle \nabla u, \nabla \nu\rangle|^2 + c_*\nu^2|D^2u|^2 \\ \le c_8\nu^2 + c_8e^{2u}H^2(e^uq) + \nu nt\sum_{k=1}^n u_k\nabla_k(e^uH(e^uq)), \tag{8-9}$$

where $c_* = \frac{1}{2} - c_6/(2c_6+1) > 0$ does not depend on t. From (8-9), by arguing as in [Gerhardt 1983, Theorem 4.1], we can conclude the proof. In fact, using Stampacchia's truncation method (for the details see the Appendix in [loc. cit.]), multiplying (8-9) with

$$\psi_l := \nu\max\{\nu - l, 0\}, \quad l \ge \nu_0,$$

and integrating by parts we deduce

$$\sup_\Omega \nu \le \nu_0 + c_9(1+|\nu|^3_{2n,\Omega}), \tag{8-10}$$

where $c_9 > 0$ is a constant depending on n, Ω, r_1, r_2 and $\|H\|_{\infty,\mathcal{C}_{\overline{\Omega}}}$ but not on t, and $|\cdot|_{p,\Omega}$ denotes the standard L^p-norm. Therefore, in order to conclude the proof it suffices to prove a uniform estimate for the L^{2n}-norm of ν with respect to the parameter t. To this end, recalling that $\nu \ge 1$, and that $e^uH(e^u)$ is uniformly bounded by a constant depending only on $r_1, r_2, \|H\|_{\infty,\mathcal{C}_{\overline{\Omega}}}$, we can rewrite the right-hand side of (8-9) in a simpler way:

$$-\sum_{i,j=1}^n \nabla_i(\nu^{-2}f_{ij}\nu_j) + (1+2c_*)|\langle \nabla u, \nabla \nu\rangle|^2 + c_*\nu^2|D^2u|^2 \\ \le c_{10}\nu^2 + \nu nt\sum_{k=1}^n u_k\nabla_k(e^uH(e^uq)). \tag{8-11}$$

Now, let $p \ge 2$ be any fixed real number, let $\lambda > 0$ be a real number to be chosen later and multiply (8-11) by

$$\rho_l := \nu_l^p e^{\lambda u},$$

where $\nu_l := \max\{\nu - l, 0\}$ and l is any fixed number such that $l \ge \nu_0$. Since $\nu_l^p e^{\lambda u} \in H_0^{1,q}(\Omega)$, for any $q \in [1, +\infty)$, we can integrate by parts and thus we obtain

$$p\sum_{i,j=1}^n \int_\Omega \nu^{-2}f_{ij}\nu_j\nu_i\nu_l^{p-1}e^{\lambda u} + \lambda\sum_{i,j=1}^n \int_\Omega \nu^{-2}f_{ij}\nu_ju_i\nu_l^pe^{\lambda u} \\ + (1+2c_*)\int_\Omega |\langle \nabla u, \nabla \nu\rangle|^2\nu_l^pe^{\lambda u} + c_*\int_\Omega \nu^2|D^2u|^2\nu_l^pe^{\lambda u} \\ \le c_{11}\int_\Omega \nu^2\nu_l^pe^{\lambda u} + c_{11}(p+1)\int_\Omega \nu\nu_l^{p-1}|\langle \nabla u, \nabla \nu\rangle|e^{\lambda u} + c_{11}\lambda\int_\Omega \nu\nu_l^p|D^2u|e^{\lambda u}. \tag{8-12}$$

Now let us observe that

$$\sum_{i,j=1}^{n} f_{ij}\nu_i\nu_j = \sum_{i,j=1}^{n} (\delta_{ij}\nu + \nu^3 u_i u_j)\nu_i\nu_j = \nu|\nabla\nu|^2 + \nu^3|\langle\nabla u, \nabla\nu\rangle|^2 \geq \nu^3|\langle\nabla u, \nabla\nu\rangle|^2. \tag{8-13}$$

In addition, by direct computation we have

$$\begin{aligned}\lambda\sum_{i,j=1}^{n}\int_\Omega \nu^{-2} f_{ij}\nu_j u_i \nu_l^p e^{\lambda u} &= \lambda\sum_{i,j=1}^{n}\int_\Omega (1-|\nabla u|^2)\nu(\delta_{ij}\nu + \nu^3 u_i u_j)\nu_j u_i \nu_l^p e^{\lambda u}\\ &= \lambda\int_\Omega \langle\nabla u, \nabla\nu\rangle\nu\nu_l^p e^{\lambda u}.\end{aligned} \tag{8-14}$$

Furthermore, fixing a large constant C_1 and splitting the domains of the integrals into two parts $|\langle\nabla u, \nabla\nu\rangle| \leq C_1$ and $|\langle\nabla u, \nabla\nu\rangle| > C_1$, by elementary computations it follows that for a suitable large constant $c_{12} > 0$ it holds that

$$\begin{aligned}c_*\int_\Omega \nu_l^p|\langle\nabla u, \nabla\nu\rangle|^2 e^{\lambda u} - c_{11}(p+1)\int_\Omega \nu\nu_l^{p-1}|\langle\nabla u, \nabla\nu\rangle| e^{\lambda u}&\\ \geq -c_{12}\int_\Omega \nu_l^p e^{\lambda u} - c_{12}\int_\Omega \nu\nu_l^{p-1}e^{\lambda u}.&\end{aligned} \tag{8-15}$$

Again by elementary considerations we obtain the further estimate

$$c_*\int_\Omega \nu^2|D^2u|^2\nu_l^p e^{\lambda u} - c_{11}\lambda\int_\Omega \nu|D^2u|\nu_l^p e^{\lambda u} \geq -c_{13}\int_\Omega \nu_l^p e^{\lambda u}. \tag{8-16}$$

Indeed, since it is always possible to find a constant $c_{11} > 0$ such that $c_* x^2 - c_{11}\lambda x + c_{13} > 0$ for all $x \geq 0$, then, taking $x = \nu|D^2u|$ we obtain the desired inequality. Therefore, from (8-12), and using the estimates (8-13)–(8-16), we deduce that

$$\begin{aligned}(1+p+c_*)&\int_\Omega |\langle\nabla u, \nabla\nu\rangle|^2\nu_l^p e^{\lambda u}\\ &\leq \lambda\int_\Omega |\langle\nabla u, \nabla\nu\rangle|\nu\nu_l^p e^{\lambda u} + c_{14}\int_\Omega \nu^2\nu_l^p e^{\lambda u} + \underbrace{c_{14}\int_\Omega \nu_l^p e^{\lambda u} + c_{14}\int_\Omega \nu\nu_l^{p-1}e^{\lambda u}}_{(I)}.\end{aligned} \tag{8-17}$$

Observe that (I) contains only powers of the form $\nu^a\nu_l^b$, with $a, b \geq 0$ such that $a+b \leq p+1$. From now on we will denote by I_1, I_2, etc. terms which are finite sums of integrals of the form $c\int_\Omega \nu^a\nu_l^b e^{\lambda u}$, where $a+b \leq p+1$, $a, b \geq 0$ and c is a constant which does not depend on t. The strategy to conclude the proof is to obtain an estimate of the kind

$$\int_\Omega \nu^2\nu_l^p e^{\lambda u} \leq I. \tag{8-18}$$

To this aim, from (8-17), dividing each side by $(p+1+c_*)$ and using the elementary inequality $xy \leq \frac{1}{2}x^2 + \frac{1}{2}y^2$, we obtain that

$$\int_\Omega |\langle\nabla u, \nabla\nu\rangle|^2\nu_l^p e^{\lambda u} \leq \frac{\lambda^2}{(1+p+c_*)^2}\int_\Omega \nu^2\nu_l^p e^{\lambda u} + \frac{2c_{12}}{1+p+c_*}\int_\Omega \nu^2\nu_l^p e^{\lambda u} + I_1. \tag{8-19}$$

Now, multiplying (8-4) by $\varphi = \nu \nu_l^p e^{\lambda u}$, integrating by parts, taking into account that $\nabla\varphi = \nabla\nu \nu_l^p e^{\lambda u} + p\nu \nabla\nu \nu_l^{p-1} e^{\lambda u} + \lambda \nu \nu_l^p \nabla u$, and $p \geq 2$, we get

$$\lambda \int_\Omega \nu^2 \nu_l^p |\nabla u|^2 e^{\lambda u} \leq c_{15} \int_\Omega \nu^2 \nu_l^p e^{\lambda u} + (p+1) \int_\Omega \nu^2 \nu_l^{p-1} |\langle \nabla u, \nabla \nu\rangle| e^{\lambda u} + I_2. \tag{8-20}$$

Now, choosing $\lambda > c_{15}$ and recalling that $\nu^{-2} = 1 - |\nabla u|^2$, from (8-20) we obtain

$$\lambda \int_\Omega \nu^2 \nu_l^p e^{\lambda u} \leq \frac{(p+1)^2}{(\lambda - c_{15})^2} \int_\Omega \nu_l^p |\langle \nabla u, \nabla \nu\rangle|^2 e^{\lambda u} + I_3. \tag{8-21}$$

From the combination of (8-19) and (8-21), for a large λ such that

$$\frac{\lambda^2 (p+1)^2}{(p+1+c_*)^2 (\lambda - c_{15})^2} + \frac{2c_{12}(p+1)^2}{(p+1+c_*)(\lambda - c_{15})^2} < 1$$

it follows that

$$\int_\Omega \nu_l^p |\langle \nabla u, \nabla \nu\rangle|^2 e^{\lambda u} \leq I_4,$$

and then, from this and (8-21), we conclude that

$$\int_\Omega \nu^2 \nu_l^p e^{\lambda u} \leq I_5,$$

which gives the desired inequality (8-18). Therefore, from (8-18) and the arbitrariness of p we deduce that $|\nu|_{2n,\Omega}$ is uniformly bounded in t and thus from (8-10) we deduce the thesis. □

9. Proofs of Theorems 1.3 and 1.4

The proofs of Theorem 1.3 and Theorem 1.4 are identical except for a small part and thus we give a unified proof in which at some point we distinguish between the two cases.

Proof. Let α, r_1, r_2, Ω and H be as in the statement of the theorem. Recalling the definition of the operators $\mathcal{Q}^t$, $\widehat{\mathcal{Q}}^t$ (see (4-13), (4-16)), by the same proof as that of Lemma 4.11 it follows that, for any $t \in [0,1]$, if $u \in C_0^2(\overline{\Omega})$ is a solution of $\widehat{\mathcal{Q}}^t(u) = 0$ and satisfies $|\nabla u|_{\infty,\Omega} < 1$ then

$$\log r_1 \leq u(q) \leq \log r_2 \quad \text{for any } q \in \overline{\Omega}. \tag{9-1}$$

Hence, by the definition of $\mathcal{Q}^t$, we have also a uniform bound with respect to t on the L^∞ norm of the solutions of $\mathcal{Q}^t(u) = 0$. In order to get a uniform bound on the gradient we use Proposition 8.1. To this end, in the case of Theorem 1.3 since Ω satisfies a uniform exterior geodesic condition and $H > 0$, thanks to Proposition 4.7 we have that (Ω, H) is admissible, and by arguing as in Step 6 of the proof of Theorem 5.1 we obtain that there exists $\theta \in (0,1)$ such that for any $t \in [0,1]$, if $u \in C_0^2(\overline{\Omega})$ is a solution of $\mathcal{Q}^t(u) = 0$ and satisfies $|\nabla u|_{\infty,\Omega} < 1$, then

$$|\nabla u(q)| \leq 1 - \theta \quad \text{for any } q \in \partial\Omega.$$

Indeed, if $|\nabla u|_{\infty,\Omega} < 1$ and $u \in C^1(\overline{\Omega})$, then, by the same proof as that of Lemma 4.1 we get that $\mathcal{Q}^t$ is uniformly elliptic in Ω (when passing to hyperbolic stereographic coordinates) and thus, thanks to the hypotheses on H, we can apply [Gilbarg and Trudinger 1977, Theorem 10.1] and argue as in Step 6 of the proof of Theorem 5.1. In the case of Theorem 1.4, the proof of this fact is identical and we use directly the hypothesis that (Ω, H) is admissible without invoking Proposition 4.7.

Since Ω is of class $C^{3,\alpha}$, $H \in C^{1,\alpha}(\mathcal{C}_{\overline{\Omega}}(r_1, r_2))$ and thanks to (9-1), by standard elliptic regularity theory, see [Gilbarg and Trudinger 1977], any solution $u \in C_0^{2,\alpha}(\overline{\Omega})$ of $\mathcal{Q}^t(u) = 0$ such that $|\nabla u| < 1$ in $\overline{\Omega}$ turns out to be of class $C^{3,\alpha}(\overline{\Omega})$. Hence, setting $\nu_0 := 1/\sqrt{1-\theta^2}$, by Proposition 8.1, it follows that there exists $\theta_* \in (0,1)$, depending only on $n, r_1, r_2, \nu_0, \Omega, H$ but not on t, such that for any solution $u \in C^3(\overline{\Omega})$ of $\mathcal{Q}^t(u) = 0$ satisfying $|\nabla u| < 1$ in Ω it holds that

$$|\nabla u(q)| \le 1 - \theta_* \quad \text{for any } q \in \overline{\Omega}. \tag{9-2}$$

Let us fix $\delta > 0$ sufficiently small so that $1 - \theta_* + \delta < 1$ and consider the set

$$U := \{w \in C_0^{1,\alpha}(\overline{\Omega}) : |\nabla w|_{\infty,\Omega} \le 1 - \theta_* + \delta\}.$$

Clearly U is a convex and closed subset of $C_0^{1,\alpha}(\overline{\Omega})$. We define the map $T : [0,1] \times U \to C_0^{1,\alpha}(\overline{\Omega})$, $T(t, w) := u$, where u is the unique solution of

$$\begin{cases} \sum_{i,j=1}^{n} ((1-|\nabla w|^2)\delta_{ij} + w_i w_j) u_{ij} = nt(1-|\nabla w|^2)(1-\sqrt{1-|\nabla w|^2} e^w \hat{H}(e^w q)) & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

We observe that T is well-defined. Indeed, for a fixed $w \in U$, considering the linear operator $L_w u := \sum_{i,j=1}^{n} ((1-|\nabla w|^2)\delta_{ij} + w_i w_j) u_{ij}$, and arguing as in Step 2 of the proof of Theorem 5.1, we see that $L_{w,\epsilon} : C_0^{2,\alpha}(\overline{\Omega}) \to C^{0,\alpha}(\overline{\Omega})$ is a bijection. Hence

$$T(w) = t L_w^{-1}\big(n(1-|\nabla w|^2)(1-\sqrt{1-|\nabla w|^2} e^w \hat{H}(e^w q))\big)$$

is defined and we are done.

It is easy to verify that T is continuous and, arguing as in the proof of Step 5 of Theorem 5.1, we have that $T([0,1] \times U)$ is a relatively compact subset of $C_0^{1,\alpha}(\overline{\Omega})$. Moreover 0 lies in the interior of U and $T(0 \times \partial U) \subset U$. To conclude the proof it suffices to prove that if $(t, u) \in [0,1] \times U$ satisfies $T(t, u) = u$ then $u \notin \partial U$. Indeed, if $T(t, u) = u$ then $u \in C_0^{2,\alpha}(\overline{\Omega})$ is a solution of $\hat{\mathcal{Q}}^t(u) = 0$ and thus from (9-1) we have $\mathcal{Q}^t(u) = 0$. Then, since $u \in U$ we have $|\nabla u|_{\infty,\Omega} \le 1 - \theta^* + \delta < 1$ and thus $\mathcal{Q}^t$ is uniformly elliptic. Therefore by elliptic regularity theory $u \in C^{3,\alpha}(\overline{\Omega})$ and thanks to (9-2) it follows that $|\nabla u|_{\infty,\Omega} \le 1 - \theta_* < 1 - \theta_* + \delta$, thus u cannot belong to ∂U and we are done.

Finally, from Theorem 2.8 we conclude that there exists $\bar{u} \in U$ which solves $T(1, \bar{u}) = \bar{u}$; i.e., $\bar{u}$ is a solution of (1-1). For the uniqueness it suffices to argue as in Step 8 of the proof of Theorem 5.1. □

References

[Bartnik and Simon 1982] R. Bartnik and L. Simon, "Spacelike hypersurfaces with prescribed boundary values and mean curvature", *Comm. Math. Phys.* **87**:1 (1982), 131–152. MR Zbl

[Bayard 2003] P. Bayard, "Dirichlet problem for space-like hypersurfaces with prescribed scalar curvature in $\mathbb{R}^{n,1}$", *Calc. Var. Partial Differential Equations* **18**:1 (2003), 1–30. MR Zbl

[Bayard and Delanoë 2009] P. Bayard and P. Delanoë, "Entire spacelike radial graphs in the Minkowski space, asymptotic to the light-cone, with prescribed scalar curvature", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26**:3 (2009), 903–915. MR Zbl

[Bonheure and Iacopetti 2019] D. Bonheure and A. Iacopetti, "On the regularity of the minimizer of the electrostatic Born–Infeld energy", *Arch. Rational Mech. Anal.* **232**:2 (2019), 697–725. MR Zbl

[Bonheure et al. 2016] D. Bonheure, P. d'Avenia, and A. Pomponio, "On the electrostatic Born–Infeld equation with extended charges", *Comm. Math. Phys.* **346**:3 (2016), 877–906. MR Zbl

[Born and Infeld 1934] M. Born and L. Infeld, "Foundations of the new field theory", *Proc. Roy. Soc. Lond. A* **144**:852 (1934), 425–451. Zbl

[Caldiroli and Gullino 2013] P. Caldiroli and G. Gullino, "Radial graphs over domains of $\mathbb{S}^n$ with prescribed mean curvature", *J. Fixed Point Theory Appl.* **13**:1 (2013), 151–161. MR Zbl

[Caldiroli and Iacopetti 2016] P. Caldiroli and A. Iacopetti, "Existence of stable H-surfaces in cones and their representation as radial graphs", *Calc. Var. Partial Differential Equations* **55**:6 (2016), art. id. 131. MR Zbl

[Cheng and Yau 1976] S. Y. Cheng and S. T. Yau, "Maximal space-like hypersurfaces in the Lorentz–Minkowski spaces", *Ann. of Math.* (2) **104**:3 (1976), 407–419. MR Zbl

[Gerhardt 1983] C. Gerhardt, "H-surfaces in Lorentzian manifolds", *Comm. Math. Phys.* **89**:4 (1983), 523–553. MR Zbl

[Gilbarg and Trudinger 1977] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Grundlehren der Math. Wissenschaften **224**, Springer, 1977. MR Zbl

[Hélein and Wood 2008] F. Hélein and J. C. Wood, "Harmonic maps", pp. 417–491 in *Handbook of global analysis*, edited by D. Krupka and D. Saunders, Elsevier, Amsterdam, 2008. MR Zbl

[Lee 1997] J. M. Lee, *Riemannian manifolds: an introduction to curvature*, Graduate Texts in Math. **176**, Springer, 1997. MR Zbl

[de Lira 2002] J. H. S. de Lira, "Radial graphs with constant mean curvature in the hyperbolic space", *Geom. Dedicata* **93** (2002), 11–23. MR Zbl

[López 2003] R. López, "A note on radial graphs with constant mean curvature", *Manuscripta Math.* **110**:1 (2003), 45–54. MR Zbl

[López 2013] R. López, *Constant mean curvature surfaces with boundary*, Springer, 2013. MR Zbl

[López 2014] R. López, "Differential geometry of curves and surfaces in Lorentz–Minkowski space", *Int. Electron. J. Geom.* **7**:1 (2014), 44–107. MR Zbl

[Papadopoulos 2005] A. Papadopoulos, *Metric spaces, convexity and nonpositive curvature*, IRMA Lect. Math. and Theor. Phys. **6**, Eur. Math. Soc., Zürich, 2005. MR Zbl

[Potter 1972] A. J. B. Potter, "An elementary version of the Leray–Schauder theorem", *J. London Math. Soc.* (2) **5** (1972), 414–416. MR Zbl

[Pucci and Serrin 2004] P. Pucci and J. Serrin, "The strong maximum principle revisited", *J. Differential Equations* **196**:1 (2004), 1–66. MR Zbl

[Radó 1932] T. Radó, "Contributions to the theory of minimal surfaces", *Acta Litt. Sci. Szeged* **6** (1932), 1–20. Zbl

[Serrin 1969] J. Serrin, "The problem of Dirichlet for quasilinear elliptic differential equations with many independent variables", *Philos. Trans. Roy. Soc. London Ser. A* **264** (1969), 413–496. MR Zbl

[Spivak 1975] M. Spivak, *A comprehensive introduction to differential geometry, IV*, Publish or Perish, Boston, 1975. MR Zbl

[Tausch 1981] E. Tausch, "The n-dimensional least area problem for boundaries on a convex cone", *Arch. Rational Mech. Anal.* **75**:4 (1981), 407–416. MR Zbl

[Treibergs 1982] A. E. Treibergs, "Entire spacelike hypersurfaces of constant mean curvature in Minkowski space", *Invent. Math.* **66**:1 (1982), 39–56. MR Zbl

[Treibergs and Wei 1983] A. E. Treibergs and S. W. Wei, "Embedded hyperspheres with prescribed mean curvature", *J. Differential Geom.* **18**:3 (1983), 513–521. MR Zbl

[Yau 1975] S. T. Yau, "Harmonic functions on complete Riemannian manifolds", *Comm. Pure Appl. Math.* **28** (1975), 201–228. MR Zbl

Received 5 Dec 2017. Revised 7 Sep 2018. Accepted 20 Nov 2018.

DENIS BONHEURE: denis.bonheure@ulb.ac.be
Département de Mathématique, Université Libre de Bruxelles, Campus de la Plaine, Bruxelles, Belgium

ALESSANDRO IACOPETTI: alessandro.iacopetti@ulb.ac.be
Département de Mathématique, Université Libre de Bruxelles, Campus de la Plaine, Bruxelles, Belgium

ANALYSIS AND PDE
Vol. 12, No. 7, 2019
dx.doi.org/10.2140/apde.2019.12.1843

SQUARE FUNCTION ESTIMATES, THE BMO DIRICHLET PROBLEM, AND ABSOLUTE CONTINUITY OF HARMONIC MEASURE ON LOWER-DIMENSIONAL SETS

SVITLANA MAYBORODA AND ZIHUI ZHAO

In the recent work G. David, J. Feneuil, and the first author have launched a program devoted to an analogue of harmonic measure for lower-dimensional sets. A relevant class of partial differential equations, analogous to the class of elliptic PDEs in the classical context, is given by linear degenerate equations with the degeneracy suitably depending on the distance to the boundary.

The present paper continues this line of research and focuses on the criteria of quantitative absolute continuity of the newly defined harmonic measure with respect to the Hausdorff measure, $\omega \in A_\infty(\sigma)$, in terms of solvability of boundary value problems. The authors establish, in particular, square function estimates and solvability of the Dirichlet problem in BMO for domains with lower-dimensional boundaries under the underlying assumption $\omega \in A_\infty(\sigma)$. More generally, it is proved that in all domains with Ahlfors regular boundaries the BMO solvability of the Dirichlet problem is necessary and sufficient for the absolute continuity of the harmonic measure.

1. Introduction

The last decade has seen great advances in the understanding of the connections between analytic, geometric, and PDE properties of sets. One of the central questions in this quest pertains to the necessary and sufficient conditions on the geometry of the domain which guarantee absolute continuity of the harmonic measure ω with respect to the surface measure σ of the boundary. The interest to this problem begins with the classical 1916 F. and M. Riesz theorem [Riesz and Riesz 1920], which asserts that for a simply connected planar domain with a rectifiable boundary, the harmonic measure is absolutely

Mayboroda was partially supported by the NSF INSPIRE Award DMS-1344235, NSF CAREER Award DMS-1220089, DMS-1839077 and the Simons Foundation grant 563916, SM. Zhao was partially supported by NSF grant numbers DMS-1361823, DMS-1500098 and DMS-1664867. This material is based upon work supported by the National Science Foundation under grant number DMS-1440140 while the authors were in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2017 semester.
MSC2010: 35J25, 42B37, 31B35.
Keywords: harmonic measure, A_∞ Muckenhoupt weight, BMO solvability, degenerate elliptic operator.

continuous with respect to the boundary surface measure (see [Lavrentyev 1936] for a quantitative version). A local analogue of this result was established in [Bishop and Jones 1990], which also showed that absolute continuity may fail in the absence of some topological hypothesis, even for a rectifiable domain. The emerging philosophy is that the key geometric properties at play are smoothness (or to be precise, rectifiability) and connectedness of the domain. In higher dimensions, the latter is much trickier, and without any pertinent details we mention that the absolute continuity of the harmonic measure with respect to the boundary surface measure has been proved in Lipschitz graph domains [Dahlberg 1977], and later in the so-called chord-arc domains in [David and Jerison 1990; Semmes 1990], and more recent achievements in the field have progressively further weakened the underlying geometric hypotheses [Bennewitz and Lewis 2004; Badger 2012; Hofmann and Martell 2014; 2017; Azzam et al. 2017; Mourgoglou 2015; Akman et al. 2016; 2017; Azzam 2017], although the sharp assumptions, particularly in terms of connectedness, are not completely clear yet. Meanwhile in the converse direction, the necessary conditions for the absolute continuity of harmonic measure with respect to the Hausdorff measure of the boundary have been obtained in 1-sided chord-arc domains in [Hofmann et al. 2014] (see also [Azzam et al. 2017]), and later in more general domains in [Mourgoglou and Tolsa 2017; Hofmann et al. 2017a]. As a culmination of this line of work, it was shown without any topological background assumptions that rectifiability is necessary for absolute continuity of the harmonic measure in [Azzam et al. 2016b]. These results were extended to general elliptic operators and other manifestations of solvability of the Dirichlet boundary value problem in [Hofmann et al. 2016; 2017b; 2017c; Toro and Zhao 2017; Azzam and Mourgoglou 2017; Garnett et al. 2018; Azzam et al. 2016a] to mention only a few: the area is blossoming and we do not aim at a complete listing of the related literature.

All of these advances heavily rely on the properties of harmonic functions, and as such, do not apply to domains with lower-dimensional boundaries, for instance, a complement of a curve in $\mathbb{R}^3$. In fact, sets of higher codimension are not visible by classical Brownian travelers (that is, the probability to hit such a set is zero) and equivalently, by classical harmonic functions. Led by these considerations, G. David, J. Feneuil, and the first author [David et al. 2017] have recently launched a program devoted to a new type of degenerate elliptic PDEs, such that the corresponding elliptic measure (still referred to as harmonic measure in the course of this discussion) is not only nontrivial, but absolutely continuous with respect to the Hausdorff measure in favorable geometric circumstances. The goal of the present paper is to establish equivalence of absolute continuity of harmonic measure to the BMO solvability of the Dirichlet problem on arbitrary Ahlfors regular domains and in the general class of degenerate elliptic operators, and to prove a technical roadblock that is very important in many applications: the square function estimates for solutions. Let us discuss this in more detail.

We shall work in the general context of d-Ahlfors–David regular sets, which are roughly speaking, d-dimensional uniformly at all scales.

Definition 1.1. Let $\Gamma \subset \mathbb{R}^n$ be a closed set and $d \leq n$ be an integer. We say Γ is d-Ahlfors regular if there exists a constant $C_0 \geq 1$ such that for any $q \in \Gamma$ and $r > 0$

$$C_0^{-1} r^d \leq \mathcal{H}^d(B(q,r) \cap \Gamma) \leq C_0 r^d,$$

where $\mathcal{H}^d$ is the d-dimensional Hausdorff measure. We shall often denote $\mathcal{H}^d|_\Gamma$, that is, $\mathcal{H}^d$ restricted to the set Γ, by σ.

Let Γ be a d-Ahlfors regular set in $\mathbb{R}^n$ with $d < n-1$, and $\Omega = \mathbb{R}^n \setminus \Gamma$. Consider the degenerate elliptic operator $L = -\operatorname{div}(A(X)\nabla)$ with a real, symmetric $n \times n$ matrix $A(X)$ satisfying

$$A(X)\xi \cdot \zeta \le C_1 |\xi||\zeta| \delta(X)^{d-n+1} \quad \text{for } X \in \Omega \text{ and } \xi, \zeta \in \mathbb{R}^n, \tag{1.2}$$

$$A(X)\xi \cdot \xi \ge C_1^{-1} |\xi|^2 \delta(X)^{d-n+1} \quad \text{for } X \in \Omega \text{ and } \xi \in \mathbb{R}^n \tag{1.3}$$

for some $C_1 \ge 1$, where $\delta(X) = \operatorname{dist}(X, \Gamma)$. We say a function u in the Sobolev space $W_r(\Omega)$ (see the definition in (2.19)) is a weak solution to $Lu = 0$ if

$$\iint_\Omega A(X)\nabla u \cdot \nabla \varphi \, dX = 0 \quad \text{for any } \varphi \in C_0^\infty(\Omega).$$

The basic elliptic theory of such equations was developed in [David et al. 2017]. In particular, it was shown that the Dirichlet problem

$$\begin{cases} Lu = 0 & \text{in } \Omega, \\ u = f & \text{on } \Gamma \end{cases} \tag{D}$$

has a suitably interpreted weak solution for smooth compactly supported (and more general) f on Γ, that such a solution is locally bounded and Hölder continuous in the interior and at the boundary, and finally, that it can be written in terms of the corresponding harmonic measure, and the latter satisfies the usual doubling, nondegeneracy, and change-of-pole conditions. We refer the reader to Section 2 for details. For now, we only recall that the harmonic measure is a (family of) positive regular Borel measure(s) ω^X on Γ, $X \in \Omega$, such that, in particular, for any boundary function $f \in C_0^0(\Gamma)$ the solution to (D) can be written as

$$u(X) = \int_\Gamma f \, d\omega^X. \tag{1.4}$$

Definition 1.5. We say the harmonic measure ω is of class A_∞ with respect to the surface measure $\sigma = \mathcal{H}^d|_\Gamma$, or simply $\omega \in A_\infty(\sigma)$, if for any $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that for any surface ball Δ, any surface ball $\Delta' \subset \Delta$ and any Borel set $E \subset \Delta'$ we have

$$\frac{\sigma(E)}{\sigma(\Delta')} < \delta \quad \Longrightarrow \quad \frac{\omega^A(E)}{\omega^A(\Delta')} < \epsilon. \tag{1.6}$$

Here $A = A_\Delta$ is a corkscrew point for Δ (see Lemma 2.50 for the definition and existence of a corkscrew point).

We remark that while the aforementioned basic properties of harmonic measure (existence, doubling, nondegeneracy, change-of-poles etc.) hold in full generality of d-Ahlfors regular sets, $d < n-1$, the A_∞ property of the harmonic measure is much more delicate and is not expected on very rough domains. In particular, already on a planar domain with 1-dimensional boundary, rectifiability of the boundary is necessary for $\omega \in A_\infty(\sigma)$. On the other hand, it is not vacuous either, as the authors in [David et al. 2019] have proved that, for any $d < n-1$ and Γ a d-dimensional Lipschitz graph with a small Lipschitz constant, the harmonic measure is absolutely continuous with respect to the Hausdorff measure for the operator $L = -\operatorname{div}(D(X)^{-n+d+1}\nabla)$, where

$$D(X) = \left\{ \int_\Gamma |X-y|^{-d-\alpha}\, d\mathcal{H}^d(y) \right\}^{-\frac{1}{\alpha}}, \quad X \in \Omega, \tag{1.7}$$

for some constant $\alpha > 0$. It is easy to see that $D(X)$ is equivalent to the Euclidean distance $\operatorname{dist}(X, \Gamma)$ (and this would even stay true when Γ is an Ahlfors regular set) but not equal.

For any $q \in \Gamma$ and $r > 0$, we use $\Delta = \Delta(q, r)$ to denote the surface ball $B(q, r) \cap \Gamma$, and use $T(\Delta) := B(q, r) \cap \Omega$ to denote the "tent" above Δ. A function f defined on Γ is a BMO function if

$$\|f\|_{\mathrm{BMO}} := \sup_{\Delta \subset \Gamma} \left(⨍_\Delta |f - f_\Delta|^2\, d\sigma \right)^{\frac{1}{2}} < \infty. \tag{1.8}$$

Here f_Δ denotes the average $⨍_\Delta f\, d\sigma$.

Definition 1.9. We say that the Dirichlet problem (D) is solvable in BMO if for any boundary function $f \in C_0^0(\Gamma)$, the solution u to (D) given by (1.4) satisfies the condition that $|\nabla u|^2\, \delta(X)^{d-n+2}\, dX$ is a Carleson measure with norm bounded by a constant multiple of $\|f\|_{\mathrm{BMO}}^2$, that is,

$$\sup_{\Delta \subset \Gamma} \frac{1}{\sigma(\Delta)} \iint_{T(\Delta)} |\nabla u|^2\, \delta(X)^{d-n+2}\, dX \le C \|f\|_{\mathrm{BMO}}^2. \tag{1.10}$$

One of the main results of the present paper is as follows.

Theorem 1.11. *Let Γ be a d-Ahlfors regular set in $\mathbb{R}^n$ with $d < n-1$ and $\Omega = \mathbb{R}^n \setminus \Gamma$. Consider the operator $L = -\operatorname{div}(A(X)\nabla)$ with a real, symmetric $n \times n$ matrix $A(X)$ satisfying* (1.2) *and* (1.3)*. Then the harmonic measure ω is of class $A_\infty(\sigma)$ if and only if the Dirichlet problem* (D) *is BMO-solvable.*

In codimension 1 this has been proved in [Dindos et al. 2011] for Lipschitz domains and in [Zhao 2018] for uniform domains with Ahlfors regular boundaries. One of the main difficulties in our case is to prove an upper bound on the square function by the nontangential maximal function. The latter, in codimension 1, goes back to the work of Dahlberg, Jerison, and Kenig [Dahlberg et al. 1984] for Lipschitz domains, and their method can be extended to more general sets with the help of preliminary estimates proved in [Jerison and Kenig 1982]. This result, and even more so the method behind it, underpinned many later developments in the subject. To prove it, [Dahlberg et al. 1984] systematically uses the harmonic measures of the sawtooth domains to get a good-λ inequality. This technique is not available to us. The sawtooth domain is a domain inside Ω on top of a set $E \subset \partial\Omega = \Gamma$ that satisfies some desired properties, and, roughly speaking, allows one to exchange local results with global ones. In some sense, it is the use of the sawtooth domains which allows one to exploit the fact that at every scale the A_∞ condition only carries information on a big portion of a boundary ball, rather than the entire boundary ball — a crucial ingredient in this and many other arguments in the theory. In the case of lower-dimensional Γ, however, the boundary of a sawtooth domain may have arbitrarily small/large pieces of dimension d and, simultaneously, pieces of dimension $n-1$. For that reason, it is not automatically clear if one can make sense of the harmonic measure for the sawtooth domain or resolve the Dirichlet problem on the sawtooth domain. Instead we are bound to work with the Green's function of the entire Ω, and get a good-λ inequality by using various considerations akin to the comparison principle. Needless to say, the geometric arguments for lower-dimensional sets are also very different and the technical side of the present paper ends up surprisingly far from [Dahlberg et al. 1984; Dindos et al. 2011; Zhao 2018]. Moreover, since the theory of the lower-dimensional sets is still in its infancy, these technical geometric arguments, e.g., Lemma 3.24, are likely to be useful in many future works.

The formal results in this direction are as follows. For any $q \in \Gamma$ and $\alpha > 0$, we define the nontangential cone $\Gamma^\alpha(q)$ with vertex q and aperture α as

$$\Gamma^\alpha(q) = \{X \in \Omega : |X-q| < (1+\alpha)\delta(X)\}, \tag{1.12}$$

and a truncated cone as

$$\Gamma_r^\alpha(q) = \Gamma^\alpha(q) \cap B(q,r).$$

When there is no confusion we drop the superindex α and simply denote them by $\Gamma(q)$ and $\Gamma_r(q)$, respectively. We define the nontangential square function

$$Su(q) = \left(\iint_{\Gamma(q)} |\nabla u|^2 \, \delta(X)^{1-d} \, dm(X)\right)^{\frac{1}{2}} \tag{1.13}$$

and the truncated square function

$$S_r u(q) = \left(\iint_{\Gamma_r(Q)} |\nabla u|^2 \, \delta(X)^{1-d} \, dm(X)\right)^{\frac{1}{2}}. \tag{1.14}$$

We also define the nontangential maximal function and its truncated analogue

$$Nu(q) = \sup_{X \in \Gamma(q)} |u(X)|, \quad N_r u(q) = \sup_{X \in \Gamma_r(q)} |u(X)|. \tag{1.15}$$

Given apertures $0 < \alpha < \alpha_1 < \beta$, for simplicity we denote by Su, $S'u$ the square functions on nontangential cones of apertures α, α_1, respectively, and denote by Nu the nontangential maximal function of aperture β. We have:

Proposition 1.16 (good-λ inequality for ω). *Suppose Γ is a d-Ahlfors regular set in $\mathbb{R}^n$ with $d < n-1$, $\Omega = \mathbb{R}^n \setminus \Gamma$ and $\mathbb{D}$ is a collection of dyadic cubes for Γ; see Lemma 3.3 for the details. Let $u \in W_r(\Omega)$ be a nonnegative solution of $Lu = 0$ such that for some dyadic cube $Q \in \mathbb{D}$ and $\lambda > 0$ there exists $q_1 \in \Gamma$ with*

$$S'u(q_1) \le \lambda \qquad \text{and} \qquad |q_1 - q| \le C_2 \operatorname{diam} Q \quad \text{for all } q \in Q.$$

Then for any $X_Q \notin B(x_Q, 2C_3\ell(Q))$ and δ sufficiently small, we have

$$\omega^{X_Q}\big(\{q \in Q : Su(q) > 2\lambda,\ Nu(q) \le \delta\lambda\}\big) \le C\delta^2 \omega^{X_Q}(Q). \tag{1.17}$$

Here x_Q is the "center" of Q and $\ell(Q)$ is the "size" of Q; see Lemma 3.3. The constant $C > 0$ depends on the allowable parameters d, n, C_0, C_1, the apertures α, α_1, β, and the given constants C_2, C_3.

If, moreover, $\omega \in A_\infty(\sigma)$, then the good-$\lambda$ inequality for σ follows and we conclude that

$$\|Su\|_{L^p(\sigma)} \leq C\|Nu\|_{L^p(\sigma)} \tag{1.18}$$

for any $1 \leq p < \infty$ and any solution $u \in W_r(\Omega)$ to $Lu = 0$ such that the right-hand side is finite.

The paper is structured as follows. In Section 2 we first state some lemmas proved in [David et al. 2017] and prove some preliminary results based off these lemmas. In Section 3 we prove the above proposition after a careful analysis of the sawtooth domains, and moreover we prove the upper bound of the square function by the nontangential maximal function. This is an independent result and will also be used in Section 4, where we prove if the harmonic measure ω is of class $A_\infty(\sigma)$, the Dirichlet problem is BMO-solvable. We prove the converse in Section 4C; that is, BMO-solvability implies the harmonic measure ω is of class $A_\infty(\sigma)$.

2. Preliminaries

The ground work for harmonic measures associated to the (degenerate) elliptic operators L on sets of lower dimensions $d < n-1$ has been laid out in the work of David, Feneuil and Mayboroda [David et al. 2017], henceforth abbreviated [DFM17]. In this section we state some relevant preliminary results proven in that paper; we also prove a few lemmas that follow easily and are needed in later sections. For the convenience of readers familiar with this subject, we point out that the new lemmas we prove here are Lemmas 2.10, 2.43 and 2.59. Unless specified otherwise, the constants that appear in the following lemmas depend only on the allowable constants, namely the dimensions n, d, the Ahlfors regular constant C_0 and the ellipticity constant C_1.

We start with the following notation:

- For any $X \in \Omega$, we define $\delta(X) = \operatorname{dist}(X, \Gamma)$, the Euclidean distance from X to Γ, and the weight $w(X) = \delta(X)^{d-n+1}$.

- We define
$$\mathcal{A}(X) := \frac{1}{w(X)} A(X) = \delta(X)^{n-1-d} A(X).$$

By (1.2) and (1.3), $\mathcal{A}(X)$ is a uniformly elliptic matrix.

- We define a measure m on Borel sets in $\mathbb{R}^n$ by letting $m(E) = \iint_E w(X)\, dm(X)$. We may write $dm(X) = w(X)\, dX$. Since $0 < w < \infty$ a.e. in $\mathbb{R}^n$, the measure m and the Lebesgue measure are mutually absolutely continuous.

- For any $q \in \Gamma$ and $r > 0$, we use the notation $\Delta(q, r)$, or sometimes simply Δ, to denote the surface ball $B(q, r) \cap \Gamma$, and $T(\Delta)$ to denote the "tent" $B(q, r) \cap \Omega$ over Δ.

- We define the surface measure $\sigma = \mathcal{H}^d|_\Gamma$.

- If $B = B(X, r)$ is a ball and $\alpha > 0$ a constant, we use $\alpha B = B(X, \alpha r)$ to denote the concentric dilation of B. The same notation applies to surface balls $\alpha\Delta$.

Lemma 2.1 (Harnack chain condition [DFM17, Lemma 2.1]). *Let Γ be a d-Ahlfors regular set in $\mathbb{R}^n$ and $d < n-1$. Then there exists a constant $c \in (0,1)$, that depends only on d, n, C_0, such that for $\Lambda \geq 1$ and $X_1, X_2 \in \Omega$ such that $\delta(X_i) \geq s$ and $|X_1 - X_2| \leq \Lambda s$, we can find two points $Y_i \in B\big(X_i, \frac{1}{2}s\big)$ such that* $\operatorname{dist}([Y_1, Y_2], \Gamma) \geq c\Lambda^{-d/(n-1-d)}s$. *That is, there is a thick tube in Ω that connects the balls $B\big(X_i, \frac{1}{2}s\big)$.*

Remark 2.2. We have

$$|Y_1 - Y_2| \leq |Y_1 - X_1| + |X_1 - X_2| + |X_2 + Y_2| < 2\Lambda s. \tag{2.3}$$

Let $\tau = c\Lambda^{-d/(n-1-d)}s$ and $Z_1 = Y_1$. For $2 \leq j \leq N$ let Z_j be consecutive points on the line segment $[Y_1, Y_2]$ such that $|Z_j - Z_{j-1}| = \frac{1}{3}\tau$. Then

$$(N-1)\tfrac{1}{3}\tau \leq |Y_1 - Y_2| < N\tfrac{1}{3}\tau.$$

Combined with (2.3) we get that the integer N satisfies

$$N \sim \frac{|Y_1 - Y_2|}{\frac{1}{3}\tau} \lesssim \Lambda^{\frac{n-1}{n-1-d}}. \tag{2.4}$$

Let $B_0 = B\big(X_1, \frac{1}{2}s\big)$, $B_j = B\big(Z_j, \frac{1}{4}\tau\big)$ for $1 \leq j \leq N$ and $B_{N+1} = B\big(X_2, \frac{1}{2}s\big)$. Clearly $B_j \cap B_{j+1} \neq \varnothing$ for all $0 \leq j \leq N$. Moreover $\operatorname{dist}(B_0, \Gamma), \operatorname{dist}(B_{N+1}, \Gamma) \geq \frac{1}{2}s$ and for $1 \leq j \leq N$,

$$\operatorname{dist}(B_j, \Gamma) \geq \tfrac{3}{4}\tau = \tfrac{3}{4}c\Lambda^{-\frac{d}{n-1-d}}s, \tag{2.5}$$

$$\operatorname{dist}(B_j, \Gamma) \leq \min\{\delta(X_1), \delta(X_2)\} + \tfrac{1}{2}s + |Y_1 - Y_2| < \min\{\delta(X_1), \delta(X_2)\} + 3\Lambda s. \tag{2.6}$$

Lemma 2.7 (estimates on the weight [DFM17, Lemma 2.3]).

(i) *For any $\theta > 0$ there exists $C_\theta > 0$ such that for any $X \in \mathbb{R}^n$ and $r > 0$ satisfying $\delta(X) \geq (1+\theta)r$,*

$$C_\theta^{-1} r^n w(X) \leq m(B(X,r)) = \iint_{B(X,r)} w(z)\,dz \leq C r^n w(X). \tag{2.8}$$

(ii) *There exists $C > 0$ such that for any $q \in \Gamma$ and $r > 0$,*

$$C^{-1} r^{d+1} \leq m(B(q,r)) = \iint_{B(q,r)\cap\Omega} w(z)\,dz \leq C r^{d+1}. \tag{2.9}$$

From the above we deduce the following estimate, which will be needed later.

Lemma 2.10. *Let Γ be d-Ahlfors regular. For any $\alpha > -1$, we have*

$$\iint_{T(2\Delta)} \delta(X)^\alpha\,dm(X) \lesssim r^{d+1+\alpha}. \tag{2.11}$$

Proof. The proof is a simple use of Vitali covering. For $j = 0, 1, \ldots$ let

$$T_j = T(2\Delta) \cap \{x \in \Omega : 2^{-j}r \leq \delta(X) < 2^{-j+1}r\},$$
$$T_{>j} = T(2\Delta) \cap \{x \in \Omega : \delta(X) < 2^{-j+1}r\}.$$

Then

$$\iint_{T(2\Delta)} \delta(X)^\alpha \, dm(X) = \sum_{j=0}^{\infty} \iint_{T_j} \delta(X)^\alpha \, dm(X) \le \sum_{j=0}^{\infty} (2^{-j} r)^\alpha m(T_{>j}). \tag{2.12}$$

For every fixed j, we consider a covering of 4Δ by

$$\bigcup_{q \in 4\Delta} B\left(q, \frac{2^{-j+1} r}{5}\right),$$

from which one can extract a countable Vitali subcovering $4\Delta \subset \bigcup_k B(q_k, 2^{-j+1} r)$, where $q_k \in 4\Delta$ and the balls $B_k = B(q_k, 2^{-j+1} r/5)$ are pairwise disjoint. The fact that $q_k \in 4\Delta = \Delta(q_0, 4r)$ implies

$$B_k := B\left(q_k, \frac{2^{-j+1} r}{5}\right) \subset B\left(q_0, 4r + \frac{2^{-j+1} r}{5}\right).$$

And the pairwise disjointness of the B_k's implies that for every fixed j, there are only finitely many of them. In fact,

$$\sum_k \sigma(B_k) = \sigma\left(\bigcup_k B_k\right) \le \sigma\left(\Delta\left(q_0, 4r + \frac{2^{-j+1} r}{5}\right)\right) \lesssim \left(4r + \frac{2r}{5}\right)^d. \tag{2.13}$$

Note that $\sigma(B_k) \approx (2^{-j+1} r/5)^d$ independent of k. Let N_j be the number of B_k's; by (2.13)

$$N_j \cdot \left(\frac{2^{-j+1} r}{5}\right)^d \le \left(4r + \frac{2r}{5}\right)^d; \quad \text{thus } N_j \lesssim 2^{jd}. \tag{2.14}$$

For any $X \in T_{>j}$, let $q_X \in \Gamma$ be such that $|X - q_X| = \delta(X)$. Then

$$|q_X - q_0| \le |q_X - X| + |X - q_0| < 4r; \quad \text{i.e., } q_X \in 4\Delta. \tag{2.15}$$

Hence $q_X \in B(q_k, 2^{-j+1} r)$ for some k. Moreover $T_{>j} \subset \bigcup_k B(q_k, 2 \cdot 2^{-j+1} r)$. Therefore by (2.14) and (2.9),

$$m(T_{>j}) \le N_j \cdot \sup_k m(B(q_k, 2 \cdot 2^{-j+1} r)) \lesssim 2^{jd} (2^{-j} r)^{d+1} \sim 2^{-j} r^{d+1}.$$

Combined with (2.12) we get

$$\iint_{T(2\Delta)} \delta(X)^\alpha \, dm(X) \lesssim \sum_{j=0}^{\infty} (2^{-j} r)^\alpha \cdot 2^{-j} r^{d+1} = r^{d+1+\alpha} \sum_{j=0}^{\infty} 2^{-j(\alpha+1)} \lesssim r^{d+1+\alpha}.$$

The last sum is convergent because $\alpha + 1 > 0$. $\square$

Now we define the suitable function spaces. We denote by $C_0^0(\Gamma)$ the space of compactly supported continuous functions on Γ, that is, $f \in C_0^0(\Gamma)$ if f is defined and continuous on Γ, and there exists a surface ball Δ such that $\operatorname{supp} f \subset \Delta$. We consider the weighted Sobolev space

$$W = \dot{W}_w^{1,2}(\Omega) = \{u \in L^1_{\text{loc}}(\Omega) : \nabla u \in L^2(\Omega, dm)\} \tag{2.16}$$

and set $\|u\|_W = \left(\iint_\Omega |\nabla u(X)|^2\, dm(X)\right)^{\frac{1}{2}}$ for $u \in W$. In fact, it was proved in Lemma 3.3 of [DFM17] that since Γ is d-Ahlfors regular with $d < n-1$,

$$W = \{u \in L^1_{\text{loc}}(\mathbb{R}^n) : \nabla u \in L^2(\mathbb{R}^n, dm)\}. \tag{2.17}$$

We also define a local version of W as follows: let $E \subset \mathbb{R}^n$ be an open set, and define

$$W_r(E) = \{u \in L^1_{\text{loc}}(E) : \varphi u \in W \text{ for all } \varphi \in C_0^\infty(E)\}. \tag{2.18}$$

As observed in [DFM17],

$$W_r(E) = \{u \in L^1_{\text{loc}}(E) : \nabla u \in L^2_{\text{loc}}(E, dm)\}. \tag{2.19}$$

It is easy to see that if $E \subset F$ are open subsets of $\mathbb{R}^n$, then $W_r(F) \subset W_r(E)$. We set

$$H = \dot{H}^{\frac{1}{2}}(\Gamma) = \left\{ g \text{ a measurable function on } \Gamma : \int_\Gamma \int_\Gamma \frac{|g(x)-g(y)|^2}{|x-y|^{d+1}}\, d\sigma(x)\, d\sigma(y) < \infty \right\}. \tag{2.20}$$

The reader may recognize this is the homogeneous Sobolev space, a special case of the Besov spaces. The authors in [DFM17] were able to define a trace operator $T : W \to H$; see Theorem 3.13 (and Lemma 8.3 for a local version $T : W_r(E) \to L^1_{\text{loc}}(\Gamma \cap E)$) there.

Lemma 2.21 (interior Caccioppoli inequality [DFM17, Lemma 8.26]). *Let $E \subset \Omega$ be an open set, and let $u \in W_r(E)$ be a nonnegative solution in E. Then for any $\phi \in C_0^\infty(E)$,*

$$\iint_\Omega \phi^2 |\nabla u|^2\, dm \le C \iint_\Omega |\nabla \phi|^2 u^2\, dm, \tag{2.22}$$

where C depends only on n, d and C_1.

In particular, if B is a ball of radius r such that $2B \subset \Omega$ and $u \in W_r(2B)$ is a nonnegative subsolution in $2B$, then

$$\iint_B |\nabla u|^2\, dm \le C r^{-2} \iint_{2B} u^2\, dm. \tag{2.23}$$

Remark 2.24. Inequality (2.23) holds if we replace $2B$ by $(1+\tau)B$, $\tau > 0$, and in that case the constant C depends on the value of τ.

Lemma 2.25 (Harnack inequality [DFM17, Lemmas 8.42, 8.44]).

(1) *Let B be a ball such that $3B \subset \Omega$ and let $u \in W_r(3B)$ be a nonnegative solution in $3B$. Then*

$$\sup_B u \le C \inf_B u, \tag{2.26}$$

where C depends on n, d and C_1.

(2) *Let K be a compact set of Ω and $u \in W_r(\Omega)$ be a nonnegative solution in Ω. Then*

$$\sup_K u \le C_K \inf_K u, \tag{2.27}$$

where C_K depends only on n, d, C_0, C_1, $\operatorname{dist}(K, \Gamma)$ *and* $\operatorname{diam} K$.

Lemma 2.28 (boundary Caccioppoli inequality [DFM17, Lemma 8.47]). *Let $B \subset \mathbb{R}^n$ be a ball centered on Γ of radius r, and let $u \in W_r(2B)$ be a nonnegative subsolution in $2B \setminus \Gamma$ such that $Tu = 0$ a.e. on $2B$. Then for any $\phi \in C_0^\infty(2B)$,*

$$\iint_{2B} \phi^2 |\nabla u|^2 \, dm \le C \iint_{2B} |\nabla \phi|^2 u^2 \, dm, \tag{2.29}$$

where C depends on n, d and C_1. In particular (2.29) *implies*

$$\iint_{B} |\nabla u|^2 \, dm \le C r^{-2} \iint_{2B} u^2 \, dm. \tag{2.30}$$

Lemma 2.31 (boundary Moser estimate [DFM17, Lemma 8.71]). *Let $p > 0$. Let B be a ball centered on Γ and $u \in W_r(2B)$ be a nonnegative subsolution in $2B \setminus \Gamma$ such that $Tu = 0$ a.e. on $2B$. Then*

$$\sup_B u \le C_p \left(\frac{1}{m(2B)} \iint_{2B} u^p \, dm \right)^{\frac{1}{p}}. \tag{2.32}$$

Lemma 2.33 (boundary Hölder regularity [DFM17, Lemma 8.106]). *Let $B = B(q, r)$ be a ball centered on Γ and $u \in W_r(B)$ be a solution in B such that $Tu \equiv 0$ on B. There exists $\beta \in (0, 1]$ such that for any $0 < s < \frac{1}{2} r$,*

$$\operatorname*{osc}_{B(q,s)} u \le C \left(\frac{s}{r} \right)^{\beta} \left(\frac{1}{m(B)} \iint_B |u|^2 \, dm \right)^{\frac{1}{2}}. \tag{2.34}$$

We are interested in the solution(s) of the Dirichlet problem (D).

Lemma 2.35 (existence and uniqueness of solution [DFM17, Lemma 9.3]). *For any $f \in H$, there exists a unique $u \in W$ such that*

$$\begin{cases} Lu = 0 & \text{in } \Omega, \\ Tu = f & \text{a.e. on } \Gamma. \end{cases} \tag{2.36}$$

Moreover $\|u\|_W \le C \|f\|_H$.

Lemma 2.37 (properties of solutions for $f \in C_0^0(\Gamma)$ [DFM17, Lemma 9.23]). *There exists a bounded linear operator*

$$U : C_0^0(\Gamma) \to C(\mathbb{R}^n)$$

such that for every $f \in C_0^0(\Gamma)$

(i) *the restriction of Uf to Γ is f;*

(ii) $\sup_{\mathbb{R}^n} Uf = \sup_\Gamma f$ *and* $\inf_{\mathbb{R}^n} Uf = \inf_\Gamma f$;

(iii) *$Uf \in W_r(\Omega)$ and is a solution of L in Ω;*

(iv) *if B is a ball centered on Γ and $f \equiv 0$ on B, then Uf lies in $W_r(B)$;*

(v) *if $f \in C_0^0(\Gamma) \cap H$, then $Uf \in W$ and is a unique solution of* (2.36).

Remark 2.38. Since $Uf \in C(\mathbb{R}^n)$, its trace $T(Uf)$ is exactly f. We also remark that $C_0^0(\Gamma) \cap H$ is dense in $C_0^0(\Gamma)$, with the supremum norm.

Lemma 2.39 (harmonic measure [DFM17, Lemmas 9.30, 9.33]). *For any $X \in \Omega$, there exists a unique positive regular Borel measure ω^X on Γ such that*

$$Uf(X) = \int_\Gamma f \, d\omega^X \quad \textit{for any } f \in C_0^0(\Gamma). \tag{2.40}$$

Additionally, for any Borel set $E \subset \Gamma$,

$$\omega^X(E) = \sup\{\omega^X(K) : E \supset K,\ K \textit{ is compact}\} = \inf\{\omega^X(V) : E \subset V,\ V \textit{ is open}\}. \tag{2.41}$$

Moreover, $\omega^X(\Gamma) = 1$.

Lemma 2.42 [DFM17, Lemma 9.38]. *Let $E \subset \Gamma$ be a Borel set and define the function u_E on Ω by $u_E(X) = \omega^X(E)$. Then:*

(i) *If there exists $X \in \Omega$ such that $u_E(X) = 0$, then $u_E \equiv 0$.*

(ii) *The function u_E lies in $W_r(\Omega)$ and is a solution in Ω.*

(iii) *If $B \subset \mathbb{R}^n$ is a ball such that $E \cap B = \varnothing$, then $u_E \in W_r(B)$ and $Tu_E = 0$ on $B \cap \Gamma$.*

For now we are only able to write down the solution to (D) if the boundary function f is in $C_0^0(\Gamma)$; see Lemma 2.37. With the help of the harmonic measure, we prove the following lemma:

Lemma 2.43. *For any function $f \in C_0^0(\Gamma)$ and any Borel set $E \subset \Gamma$, the function*

$$u(X) := \int_E f \, d\omega^X \tag{2.44}$$

defined on Ω satisfies the following:

(1) *It is continuous in Ω.*

(2) *It is a solution of $Lu = 0$ in Ω and lies in $W_r(\Omega)$.*

(3) *If $B \subset \mathbb{R}^n$ is an open ball such that $E \cap B = \varnothing$, then u is continuous in $B \cap \Omega$, u can be continuously extended to zero on $B \cap \Gamma$, and $u \in W_r(B)$.*

Remark 2.45. We note the following:

- Compared with Lemmas 2.39 and 2.37, this lemma says that $f\chi_E$ integrated against the harmonic measure gives rise to a continuous solution for any Borel set $E \subset \Gamma$.
- If the Borel set E is bounded, then the same properties hold for any bounded continuous function $f \in C_b(\Gamma)$.

Proof. Since the definition (2.44) is a linear integration, we may assume without loss of generality that f is nonnegative. Otherwise we just write $f = f_+ - f_-$, with $f_\pm \in C(\mathbb{R}^n)$. We first assume that E is an open set, and that $\omega^X(E) > 0$ for some $X \in \Omega$. By Lemma 2.42(i) it follows that $\omega^X(E) > 0$ for all $X \in \Omega$. Fix an arbitrary $X_0 \in \Omega$. Let K_j be an increasing sequence of compact sets in E such that $\omega^{X_0}(E \setminus K_j) < 1/j$. By Urysohn's lemma we can construct $g_j \in C_0^0(\Gamma)$ such that $\chi_{K_j} \le g_j \le \chi_E$, and without loss of generality we can choose the sequence g_j to be increasing. Note that $fg_j \in C_0^0(\Gamma)$, and hence by Lemma 2.37 we may define $u_j = U(fg_j) \in C^0(\Gamma)$. Then

$$0 \le u(X) - u_j(X) = \int f(\chi_E - g_j)\, d\omega^X \le \omega^X(E \setminus K_j)\|f\|_{L^\infty}.$$

By Lemmas 2.42 and 2.25, for any compact subset K in Ω containing X_0, we have

$$\omega^X(E \setminus K_j) \le C_K \omega^{X_0}(E \setminus K_j)$$

holds for every $X \in K$. Here the constant C_K only depends on $n, d, C_1, \operatorname{dist}(K, \Gamma)$ and $\operatorname{diam} K$, and in particular it is independent of j. Therefore

$$0 \le u(X) - u_j(X) \le \frac{C_K \|f\|_{L^\infty}}{j};$$

namely $\{u_j\}$ converges uniformly on compact sets of Ω to u, and thus u is continuous on Ω.

Let $\phi \in C_0^\infty(\Omega)$ be arbitrary; we claim that $\{u_j\}$ has a subsequence, which we relabel, such that

$$\nabla(\phi u_j) \rightharpoonup \nabla(\phi u) \quad \text{in } L^2(\Omega, w). \tag{2.46}$$

In particular $\nabla(\phi u) \in L^2(\Omega, w)$ for all $\phi \in C_0^\infty(\Omega)$, and thus $u \in W_r(\Omega)$. Indeed, by the interior Caccioppoli inequality (2.22), we have

$$\iint_\Omega |\nabla(\phi u_j)|^2\, dm \le 2 \iint_\Omega (|\nabla\phi|^2 u_j^2 + \phi^2 |\nabla u_j|^2)\, dm \le C \iint_\Omega |\nabla\phi|^2 u_j^2\, dm. \tag{2.47}$$

Recall that $u_j \to u$ uniformly on the compact set $\operatorname{supp}\phi$, and the right-hand side of (2.47) converges to $C \iint_\Omega |\nabla\phi|^2 u^2\, dm$. As a consequence the left-hand side of (2.47) is uniformly bounded in j. Therefore there is a subsequence (which we relabel) such that $\nabla(\phi u_j)$ converges weakly in $L^2(\Omega, w)$ to some function v. By the uniqueness of limit in the distributional sense, we conclude that $v = \nabla(\phi u)$, which finishes the proof of the claim (2.46).

Recall each u_j is a solution of L in Ω. Let $\varphi \in C_0^\infty(\Omega)$ be an arbitrary test function. We choose $\phi \in C_0^\infty(\Omega)$ such that $\phi \equiv 1$ on $\operatorname{supp}\varphi$. In particular $\nabla(\phi u) = \nabla u$ and $\nabla(\phi u_j) = \nabla u_j$ on $\operatorname{supp}\varphi$. Thus

$$\begin{aligned}\iint_\Omega A\nabla u \cdot \nabla\varphi\, dX &= \iint_\Omega \mathcal{A}\nabla u \cdot \nabla\varphi\, dm = \iint_\Omega \mathcal{A}\nabla(\phi u) \cdot \nabla\varphi\, dm \\ &= \lim_{j\to\infty} \iint_\Omega \mathcal{A}\nabla(\phi u_j) \cdot \nabla\varphi\, dm \\ &= \lim_{j\to\infty} \iint_\Omega \mathcal{A}\nabla u_j \cdot \nabla\varphi\, dm = \lim_{j\to\infty} \iint_\Omega A\nabla u_j \cdot \nabla\varphi\, dX = 0. \end{aligned} \tag{2.48}$$

If E is not an open set, the proof is similar, and we just need to approximate E from above by open sets. We omit the details here.

Going further, if $B \subset \mathbb{R}^n$ is an open ball such that $E \cap B = \varnothing$, we first prove that u can be continuously extended to zero on $\Gamma \cap B$. Take an arbitrary $q \in \Gamma \cap B$. Choose $r > 0$ sufficiently small so that $B(q, 2r) \subset B$. Consider a function $g \in C_0^\infty(\mathbb{R}^n)$ satisfying $\chi_{B(q,r)} \le g \le \chi_{B(q,2r)}$. If $f \in C_0^0(\Gamma)$, then

$f(1-g) \in C_0^0(\Gamma)$. If the Borel set E is bounded and f is only assumed to be bounded continuous, we let $\varphi \in C_0^\infty(\mathbb{R}^n)$ be a function such that $\varphi \equiv 1$ on a compact set containing E and $B(q, 2r)$. Then $f(1-g)\varphi \in C_0^0(\Gamma)$. Let

$$\tilde{u}(X) := U(f(1-g)\varphi) = \int_\Gamma f(1-g)\varphi \, d\omega^X.$$

(For simplicity we take $\varphi \equiv 1$ for case when $f \in C_0^0(\Gamma)$.) By the positivity of the harmonic measure and the fact that $E \subset \Gamma \setminus B(q, 2r)$, we deduce that $0 \le u(X) \le \tilde{u}(X)$ for all $X \in \Omega$. Recall by Lemma 2.37 that $\tilde{u} \in C(\mathbb{R}^n)$, and as $X \to q' \in B(q,r) \cap \Gamma$, we have $\tilde{u}(X) \to f(1-g)\varphi(q') = 0$. By the squeeze theorem, u can be continuously extended to zero on $B(q,r) \cap \Gamma$, and the resulting function, still denoted as u, is continuous in $B(q,r)$.

Now we show that $u \in W_r(B)$. To this end, let $\phi \in C_0^\infty(B)$; it suffices to show that $\nabla(\phi u) \in L^2(B, w)$. From Lemma 2.37(iv), Remark 2.38 and the boundary Caccioppoli inequality (2.29), we have

$$\iint_B |\nabla(\phi u_j)|^2 \, dM \le 2 \iint_B (|\nabla \phi|^2 u_j^2 + \phi^2 |\nabla u_j|^2) \, dm \le C \iint_B |\nabla \phi|^2 u_j^2 \, dm. \tag{2.49}$$

Recall that $u_j \to u$ pointwise on $B \setminus \Gamma$. Since u is continuous on B, we know $u \in L^2(\operatorname{supp} \phi, w)$. Hence by the dominated convergence theorem the right-hand side of (2.49) converges to $C \iint_B |\nabla \phi|^2 u^2 \, dm$. As a consequence the left-hand side is uniformly bounded, and thus, passing to a subsequence, $\nabla(\phi u_j)$ converges weakly in $L^2(B, w)$ to some function v. By the uniqueness of the limit we deduce $v = \nabla(\phi u)$. In particular this implies $\nabla(\phi u) \in L^2(B, w)$. □

In summary, we can write down the solution of L using the harmonic measure for the following classes of boundary data: continuous and compactly supported functions $f \in C_0^0(\Gamma)$ (see Lemma 2.37), characteristic functions χ_E for Borel sets $E \subset \Gamma$ (see Lemma 2.42), their products $f\chi_E$ (see the above Lemma 2.43), or a linear combination of the above. For the third case, if the Borel set E is bounded, we only need to assume $f \in C_b(\Gamma)$.

Lemma 2.50 (corkscrew point [DFM17, Lemma 11.46]). *There exists $M > 1$ such that for any $q \in \Gamma$ and $r > 0$, there exists a point $A = A_r(q) \in \Omega$ such that*

$$|A - q| < r, \quad \delta(A) \ge \frac{r}{M}. \tag{2.51}$$

This point will be referred to as a corkscrew point hereafter.

Remark 2.52. Note that neither Lemma 2.1 nor Lemma 2.50 is automatically true if $d = n-1$. In fact in the case of codimension 1, people often work with domains that satisfy the Harnack chain condition in which there exists a corkscrew point at all scales, called uniform domains or 1-sided NTA domains in the literature.

Lemma 2.53 (boundary Harnack inequality [DFM17, Lemma 11.50]). *Let $q \in \Gamma$ and $r > 0$ be given, and let $A = A_r(q)$ be a corkscrew point as in Lemma 2.50. Let $u \in W_r(B(q, 2r))$ be a nonnegative solution of $Lu = 0$ in $B(q, 2r) \cap \Omega$ that is not identically zero such that $Tu \equiv 0$ on $\Delta(q, 2r)$. Then*

$$u(X) \le C u(A) \quad \text{for all } X \in B(q,r). \tag{2.54}$$

We also recall the following "classical" Poincaré inequality for Sobolev functions.

Lemma 2.55 (Poincaré inequality [DFM17, Lemma 4.13]). *Let Γ be a d-Ahlfors regular set in $\mathbb{R}^n$ with $d < n-1$. For any function $v \in W$, $X \in \mathbb{R}^n$ and $r > 0$, let $B = B(X,r)$; then*

$$\left(\frac{1}{m(B)}\iint_B |v(Y)-v_B|^2\,dm(Y)\right)^{\frac{1}{2}} \le Cr\left(\frac{1}{m(B)}\iint_B |\nabla v(Y)|^2\,dm(Y)\right)^{\frac{1}{2}}, \tag{2.56}$$

where v_B denotes the average $m(B)^{-1}\int_B v\,dm$.

Suppose $\Delta = B(q_0,r)\cap\Gamma$ is a surface ball. For any $q\in\Delta$ and any $j\in\mathbb{N}$, let

$$\Gamma_j(q) = \Gamma(q)\cap\big(B(q,2^{-j}r)\setminus B(q,2^{-j-1}r)\big) \tag{2.57}$$

be a stripe in the cone $\Gamma(q)$ at height $2^{-j}r$, and

$$\Gamma_{j\to j+m}(q) = \bigcup_{i=j}^{j+m}\Gamma_i(q) = \Gamma(q)\cap\big(B(q,2^{-j}r)\setminus B(q,2^{-(j+m)-1}r)\big) \tag{2.58}$$

be a union of $m+1$ stripes. With this notation we can prove a less conventional form of the Poincaré inequality, available for solutions with vanishing boundary values.

Lemma 2.59. *Suppose that $u \in W_r(\Omega)$ is a nonnegative solution of L, $Tu = 0$ on 3Δ and $u \in W_r(B(q_0,3r))$. There exist an aperture $\bar\alpha > \alpha$ and integers m_1, m_2 such that for all $q\in\Delta$*

$$\iint_{\Gamma_j^\alpha(q)} u^2\,dm(X) \le C(2^{-j}r)^2\iint_{\Gamma^{\bar\alpha}_{j-m_1\to j+m_2}(q)} |\nabla u|^2\,dm(X). \tag{2.60}$$

The constants $m_1, m_2, \bar\alpha$ and C only depend on n, d, α, C_0, C_1.

Proof. Let B be a ball compactly contained in Ω. Recall that $u\in W_r(\Omega)$; in particular, $\varphi u\in W$ for $\varphi\in C_0^\infty(\Omega)$ such that $\varphi\equiv 1$ on B. Applying the above Lemma 2.55 to φu and squaring both sides, we get

$$\iint_B |u(Y)-u_B|^2\,dm(Y) \le Cr_B^2\iint_B |\nabla u(Y)|^2\,dm(Y). \tag{2.61}$$

For $j\in\mathbb{N}$, let A_j denote a corkscrew point for $B(q,2^{-j}r)$, whose existence is guaranteed by Lemma 2.50. Let m be a large integer whose value is to be determined later. Take $X\in\Gamma_j^\alpha(q)$ and $X' = A_{j+m}$; then

$$\delta(X) > \frac{1}{1+\alpha}|X-q| \ge \frac{2^{-j-1}r}{1+\alpha}, \quad \delta(X') \ge \frac{2^{-(j+m)}r}{M}, \tag{2.62}$$
$$|X-X'| \le |X-q|+|q-X'| \le 2^{-j}r + 2^{-(j+m)}r \le 2^{1-j}r.$$

Applying Lemma 2.1 and Remark 2.2 to X, X' with $s = 2^{-(j+m)}r/M$ and $\Lambda = 2^{m+1}M$, we can find balls $B_0 = B\big(X,\tfrac12 s\big)$, $B_i = B\big(Z_i,\tfrac14\tau\big)$, with $\tau = c\Lambda^{-d/(n-1-d)}s$, and $B_{N+1} = B\big(X',\tfrac12 s\big)$ that form a

Harnack chain connecting X to X' and satisfy (2.4), (2.5) and (2.6). Hence by Lemma 2.3(i) of [DFM17] and (2.6), (2.5), we have

$$m(B_i) \geq C^{-1}\left(\frac{\tau}{4}\right)^n \operatorname{dist}(B_i,\Gamma)^{d-n+1} \gtrsim \tau^n(\Lambda s)^{d-n+1} \sim \Lambda^{1-n}\tau^{d+1}, \tag{2.63}$$

$$m(B_i) \leq C\left(\frac{\tau}{4}\right)^n \operatorname{dist}(B_i,\Gamma)^{d-n+1} \lesssim \tau^n \tau^{d-n+1} \sim \tau^{d+1} \tag{2.64}$$

for all $i = 0,\dots,N,N+1$. A simple computation shows $B_{i+1} \subset 3B_i$ for all $i = 1,\dots N-1$, and $B_1 \subset \frac{3}{2}B_0$, $B_N \subset \frac{3}{2}B_{N+1}$ if m is sufficiently large. Therefore for each $i = 1,\dots,N-1$,

$$\begin{aligned}
|u_{B_{i+1}} - u_{3B_i}|^2 &\leq \left(\frac{1}{m(B_{i+1})}\iint_{B_{i+1}} |u(X) - u_{3B_i}|\,dm(X)\right)^2 \\
&\leq \frac{1}{m(B_{i+1})}\iint_{3B_i} |u(X)-u_{3B_i}|^2\,dm(X) \\
&\lesssim \Lambda^{n-1}\tau^{1-d}\iint_{3B_i} |\nabla u(Y)|^2\,dm(Y) \quad \text{by (2.61), (2.63).}
\end{aligned} \tag{2.65}$$

Similarly

$$|u_{B_i} - u_{3B_i}|^2 \lesssim \Lambda^{n-1}\tau^{1-d}\iint_{3B_i} |\nabla u(Y)|^2\,dm(Y).$$

Hence

$$|u_{B_i} - u_{B_{i+1}}|^2 \leq C\Lambda^{n-1}\tau^{1-d}\iint_{3B_i} |\nabla u(Y)|^2\,dm(Y). \tag{2.66}$$

A similar argument shows that for the endpoint case $i = 0$ or $N+1$

$$\begin{aligned}
|u_{B_i} - u_{B_{i\pm 1}}|^2 &\lesssim \max\{s^{1-d}, \Lambda^{n-1}s^2\tau^{-1-d}\}\iint_{\frac{3}{2}B_i} |\nabla u(Y)|^2\,dm(Y) \\
&\sim \Lambda^{n-1}s^2\tau^{-1-d}\iint_{\frac{3}{2}B_i} |\nabla u(Y)|^2\,dm(Y).
\end{aligned} \tag{2.67}$$

The last line is justified since $\Lambda \gg 1$ implies $\tau \ll s$. Combining this observation, (2.66), (2.67) and (2.4), we get

$$\begin{aligned}
&\iint_{B_0} |u(X) - u_{B_{N+1}}|^2\,dm(X) \\
&\qquad \lesssim N\cdot\iint_{B_0} |u(X)-u_{B_0}|^2\,dm(X) + N\cdot m(B_0)\sum_{i=0}^{N} |u_{B_i} - u_{B_{i+1}}|^2 \\
&\qquad \lesssim N\Lambda^{n-1}s^2\left(\frac{s}{\tau}\right)^{d+1}\iint_{\frac{3}{2}B_0\cup(\bigcup_{i=1}^N 3B_i)\cup\frac{3}{2}B_{N+1}} |\nabla u(Y)|^2\,dm(Y) \\
&\qquad \leq C'\Lambda^{\frac{n-1+d(d+1)}{n-1-d}+n-1}s^2\iint_{\frac{3}{2}B_0\cup(\bigcup_{i=1}^N 3B_i)\cup\frac{3}{2}B_{N+1}} |\nabla u(Y)|^2\,dm(Y).
\end{aligned} \tag{2.68}$$

On the other hand, by the Harnack inequality

$$u(X) \leq C u(X') \quad \text{for all } X \in B_{N+1} = B\big(X', \tfrac{1}{2}s\big).$$

Recall that $X' = A_{j+m}$. For any $q \in \Delta$, by the assumption we know that $u \in W_r(B(q, 2r))$ vanishes on $\Delta(q, 2r)$. By boundary Hölder regularity (Lemma 2.33) and the boundary Harnack principle (Lemma 2.53) we have

$$u(X') \leq C 2^{-m\beta} u(A_j),$$

with a constant C independent of j and m. Thus

$$u_{B_{N+1}}^2 \lesssim u^2(X') \lesssim 2^{-2m\beta} u^2(A_j) \lesssim 2^{-2m\beta} \cdot \frac{1}{m(B_0)} \iint_{B_0} u^2 \, dm(X). \tag{2.69}$$

The last inequality holds because A_j is a corkscrew point and $B_0 = B\big(X, \frac{1}{2}s\big)$ for some $X \in \Gamma_j(q)$. Combining (2.69) and (2.68) we obtain

$$\begin{aligned} \iint_{B_0} u^2 \, dm(X) &\leq 2m(B_0)(u_{B_{N+1}})^2 + 2 \iint_B |u(x) - u_{B_{N+1}}|^2 \, dm(X) \\ &\leq A_1 2^{-2m\beta} \iint_{B_0} u^2 \, dm(X) \\ &\quad + A_2 \Lambda^{\frac{n-1+d(d+1)}{n-1-d} + n-1} s^2 \iint_{\frac{3}{2}B_0 \cup (\bigcup_{i=1}^N 3B_i) \cup \frac{3}{2}B_{N+1}} |\nabla u(Y)|^2 \, dm(Y). \end{aligned} \tag{2.70}$$

Choose m big enough such that

$$A_1 2^{-2m\beta} \leq \tfrac{1}{2}, \quad \text{as well as} \quad 2 \cdot \frac{2^{-m}}{M} \leq \frac{1}{2(1+\alpha)}; \tag{2.71}$$

then we can absorb the first term on the right-hand side of (2.70) to the left. Recall that $B_0 = B\big(X, \frac{1}{2}s\big)$ for X satisfying (2.62). The reason for the second assumption in (2.71) is to guarantee the enlarged ball $\frac{3}{2}B_0$ is compactly contained in Ω. Fix the value of m from now on; thus the value of $\Lambda = 2^{m+1}/M$ is also fixed. We get

$$\iint_{B_0} u^2 \, dm(X) \leq C s^2 \iint_{\frac{3}{2}B_0 \cup (\bigcup_{i=1}^N 3B_i) \cup \frac{3}{2}B_{N+1}} |\nabla u(y)|^2 \, dy, \tag{2.72}$$

where $s = 2^{-(j+m)} r/M$ and the constant C depends on d, n, C_0, C_1 (recall the values of the corkscrew constant M and the Harnack chain constant c only depend on d, n, C_0, C_1). Since $B_0 = B\big(X, \frac{1}{2}s\big)$ with center $X \in \Gamma_j^\alpha(q)$, it is a simple exercise to show that given the second assumption of (2.71), there exists an aperture $\alpha_1 > \alpha$ such that

$$\tfrac{3}{2} B_0 \subset \Gamma_{j-1 \to j+1}^{\alpha_1}(q). \tag{2.73}$$

A similar statement holds for $\frac{3}{2}B_{N+1}$. Moreover (2.5) and (2.6) imply that for $i = 1, \dots, N$, there exist an aperture $\alpha_2 > \alpha$ and an integer m_0 depending on the constants c, M from Lemmas 2.1 and 2.50

such that

$$3B_i \subset \Gamma^{\alpha_2}_{j-3\to j+m+m_0}(q). \tag{2.74}$$

Let $\bar{\alpha} = \max\{\alpha_1, \alpha_2\}$. Combining the above observations with (2.72) we get

$$\iint_{B_0} u^2\, dm(X) \le Cs^2 \iint_{\Gamma^{\bar{\alpha}}_{j-3\to j+m+m_0}(q)} |\nabla u(y)|^2\, dy. \tag{2.75}$$

Consider the covering

$$\Gamma^{\alpha}_j(q) \subset \bigcup_{X\in\Gamma^{\alpha}_j(q)} B\big(X, \tfrac{1}{10}s\big). \tag{2.76}$$

We can extract a finite Vitali subcovering $\{B^k = B\big(X_k, \frac{1}{2}s\big)\}_k$ such that

$$\Gamma^{\alpha}_j(q) \subset \bigcup_k B^k \tag{2.77}$$

and $\{\frac{1}{5}B^k = B\big(X_k, \frac{1}{10}s\big)\}_k$ is mutually disjoint. Moreover the number of balls B^k is uniformly bounded by a constant $C(n,m,M)$. Note that (2.75) holds for all such balls B^k in place of B_0; we deduce

$$\begin{aligned}\iint_{\Gamma^{\alpha}_j(q)} u^2\, dm(X) &\le \sum_k \iint_{B^k} u^2\, dm(X)\\ &\le CC(n,m,M)s^2 \iint_{\Gamma^{\bar{\alpha}}_{j-3\to j+m+m_0}(q)} |\nabla u(y)|^2\, dy.\end{aligned} \tag{2.78}$$

Since the value of m is fixed, we finish the proof of Lemma 2.59. □

Lemma 2.79 (nondegeneracy of harmonic measure [DFM17, Lemma 11.73]). *Let $\lambda > 1$ be given. There exists a constant $C_\lambda > 1$ such that for any $q \in \Gamma$, $r > 0$, and $A = A_r(q)$, where $A_r(q)$ is a corkscrew point from Lemma 2.50, we have*

$$\omega^X(B(q,r)\cap\Gamma) \ge C_\lambda^{-1} \quad \textit{for } X \in B(q, r/\lambda), \tag{2.80}$$

$$\omega^X(B(q,r)\cap\Gamma) \ge C_\lambda^{-1} \quad \textit{for } X \in B(A, \delta(A)/\lambda). \tag{2.81}$$

In [DFM17] the authors also prove the existence, uniqueness and properties of the Green's function, that is, formally, a function G defined on $\Omega\times\Omega$ such that for any $Y \in \Omega$

$$\begin{cases} LG(\cdot, Y) = \delta_Y & \text{in } \Omega,\\ G(\cdot, Y) = 0 & \text{on } \Gamma,\end{cases}$$

where δ_Y is the delta function.

Lemma 2.82 (estimates of Green's function [DFM17, Lemma 11.78]). *There exists a constant $C \ge 1$ such that for any $q \in \Gamma$, $r > 0$, $\Delta = B(q,r)\cap\Gamma$ and a corkscrew point $A = A_r(q)$ we have*

$$C^{-1}r^{d-1}G(X_0, A) \le \omega^{X_0}(\Delta) \le Cr^{d-1}G(X_0, A) \quad \textit{for } X_0 \in \Omega\setminus B(q, 2r). \tag{2.83}$$

Lemma 2.84 (doubling of harmonic measure [DFM17, Lemma 11.102]). *For $q \in \Gamma$ and $r > 0$, we have*

$$\omega^X(B(q,2r)\cap\Gamma) \le C\omega^X(B(q,r)\cap\Gamma) \tag{2.85}$$

for any $X \in \Omega \setminus B(q,4r)$.

Lemma 2.86 (change of poles [DFM17, Lemma 11.135]). *Let $q \in \Gamma$ and $r > 0$ be given, and let $A = A_r(q)$ be a corkscrew point as in Lemma 2.50. Let $E, F \subset \Delta(q,r)$ be two Borel subsets of Γ such that $\omega^A(E)$ and $\omega^A(F)$ are positive. Then*

$$\frac{\omega^X(E)}{\omega^X(F)} \sim \frac{\omega^A(E)}{\omega^A(F)} \quad \textit{for any } X \in \Omega \setminus B(q,2r). \tag{2.87}$$

In particular with the choice $F = \Delta(q,r)$,

$$\frac{\omega^X(E)}{\omega^X(\Delta(q,r))} \sim \omega^A(E) \quad \textit{for any } X \in \Omega \setminus B(q,2r). \tag{2.88}$$

Let us restate the definition of $\omega \in A_\infty(\sigma)$ and make a few remarks that will become useful later.

Definition 2.89. We say the harmonic measure ω is of class A_∞ with respect to the surface measure $\sigma = \mathcal{H}^d|_\Gamma$, or simply $\omega \in A_\infty(\sigma)$, if for any $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that for any surface ball Δ, any surface ball $\Delta' \subset \Delta$ and any Borel set $E \subset \Delta'$ we have

$$\frac{\sigma(E)}{\sigma(\Delta')} < \delta \quad \Longrightarrow \quad \frac{\omega^A(E)}{\omega^A(\Delta')} < \epsilon. \tag{2.90}$$

Here $A = A_\Delta$ is a corkscrew point for Δ (see Lemma 2.50).

Remark 2.91. (i) The reader may recall that the standard definition for A_∞ is that the harmonic measure with a fixed pole, i.e., ω^{X_0}, satisfies (2.90). For unbounded boundary Γ though, the standard definition needs to be replaced by its scale-invariant analogue, which is Definition 2.89. In fact since Γ is unbounded, it is impossible to have $\omega^{X_0} \in A_\infty(\sigma)$ with a fixed pole X_0; see the comments after Theorem 1.18 of [David et al. 2019].

(ii) The above definition is symmetric: Suppose $\omega \in A_\infty(\sigma)$. Then we also have $\sigma \in A_\infty(\omega)$ (in a scale-invariant sense); i.e., the smallness of $\omega^A(E)/\omega^A(\Delta')$ implies the smallness of $\sigma(E)/\sigma(\Delta')$.

(iii) In particular, the assumption (2.90) implies that $\omega^A \ll \sigma$ when restricted to Δ. We denote the Radon–Nikodym derivative by $k^A = d\omega^A/d\sigma$. Since both ω^A and σ are Radon measures, we have

$$k^A(q) = \lim_{\substack{\Delta'=\Delta(q,r)\\ r\to 0}} \frac{\omega^A(\Delta')}{\sigma(\Delta')} \quad \text{for } \sigma\text{-a.e. } q \in \Delta. \tag{2.92}$$

Moreover since σ is doubling, by standard harmonic analysis techniques (see [García-Cuerva and Rubio de Francia 1985] for example for the proof) (2.90) implies that k^A satisfies a reverse Hölder inequality: there are constants $r_0 > 1$, $C > 0$ such that for all $r \in (1, r_0)$,

$$\left(⨍_\Delta |k^A|^r \, d\sigma\right)^{\frac{1}{r}} \le C ⨍_\Delta k^A \, d\sigma. \tag{2.93}$$

The constants r_0 and C only depend on the constants characterizing the A_∞ property (2.90); in particular, they are independent of Δ and A.

Recall that one of our main goals is to prove Theorem 1.11, which states the equivalence between $\omega \in A_\infty(\sigma)$ and the BMO solvability of the Dirichlet problem. We make a few preliminary remarks.

Note that (Γ, σ) is a space of homogeneous type. By the John–Nirenberg inequality for spaces of homogeneous type, we may also use any L^p norm $(1 \le p < \infty)$ in the definition (1.8), and the resulting BMO norms are all equivalent. See [Coifman and Weiss 1977; John and Nirenberg 1961]. Also it is easy to see that if $f \in L^\infty(\Gamma)$, then f is a BMO function with $\|f\|_{\mathrm{BMO}} \le \sqrt{2}\|f\|_{L^\infty}$.

We observe that the Carleson measure norm of $|\nabla u|^2\,\delta(X)\,dm(X)$ is in some sense equivalent to the integral of the truncated square function. Suppose $\Delta = \Delta(q_0, r)$ is an arbitrary surface ball. For any $X \in T(\Delta)$, we define

$$\Delta^X = \{q \in \Gamma : X \in \Gamma(q)\}.$$

Let $q_X \in \Gamma$be a point such that $|X - q_X| = \delta(X)$. Then

$$\Delta(q_X, \alpha\delta(X)) \subset \Delta^X \subset \Delta(q_X, (\alpha+2)\delta(X)). \tag{2.94}$$

Since Γ is d-Ahlfors regular, (2.94) implies $\sigma(\Delta^X) \approx \delta(X)^d$. Thus

$$\begin{aligned}\iint_{T(\Delta)} |\nabla u|^2\,\delta(X)\,dm(X) &\approx \iint_{T(\Delta)} |\nabla u|^2\,\delta(X)^{1-d}\sigma(\Delta^X)\,dm(X)\\ &= \iint_{T(\Delta)} |\nabla u|^2\,\delta(X)^{1-d}\int_{\Delta^X} d\sigma(q)\,dm(X).\end{aligned} \tag{2.95}$$

Changing the order of integration, on one hand we get an upper bound

$$\begin{aligned}\iint_{T(\Delta)} |\nabla u|^2\,\delta(X)^{1-d}\int_{\Delta^X} d\sigma(q)\,dm(X) &\le \int_{|q-q_0|<(\alpha+2)r}\iint_{\Gamma_{(\alpha+1)r}(q)} |\nabla u|^2\,\delta(X)^{1-d}\,dm(X)\,d\sigma\\ &\le \int_{(\alpha+2)\Delta} |S_{(\alpha+1)r}u|^2\,d\sigma.\end{aligned} \tag{2.96}$$

On the other hand, we get a lower bound

$$\begin{aligned}\iint_{T(\Delta)} |\nabla u|^2\,\delta(X)^{1-d}\int_{\Delta^X} d\sigma(q)\,dm(X) &\ge \int_{|q-q_0|<\frac{1}{2}r}\iint_{\Gamma_{r/2}(q)} |\nabla u|^2\,\delta(X)^{1-d}\,dm(X)\,d\sigma\\ &\ge \int_{\frac{1}{2}\Delta} |S_{\frac{1}{2}r}u|^2\,d\sigma.\end{aligned} \tag{2.97}$$

Therefore for any $q_0 \in \Gamma$,

$$\sup_{\substack{\Delta=\Delta(q_0,s)\\ s>0}} \frac{1}{\sigma(\Delta)}\iint_{T(\Delta)} |\nabla u|^2\,\delta(X)\,dm(X) \approx \sup_{\substack{\Delta=\Delta(q_0,r)\\ r>0}} \frac{1}{\sigma(\Delta)}\int_\Delta |S_r u|^2\,d\sigma. \tag{2.98}$$

3. Bound of the square function by the nontangential maximal function

The goal of this section is to prove:

Theorem 3.1. *Let Γ be a d-Ahlfors regular set in $\mathbb{R}^n$ with an integer $d \le n-1$, and let ω be the harmonic measure of the domain $\Omega = \mathbb{R}^n \setminus \Gamma$. Suppose $\omega \in A_\infty(\sigma)$; then*

$$\|Su\|_{L^p(\sigma)} \le C\|Nu\|_{L^p(\sigma)} \tag{3.2}$$

for any $1 \le p < \infty$ and any solution $u \in W_r(\Omega)$ to $Lu = 0$ such that the right-hand side is finite. Here the constant $C > 0$ depends on the allowable parameters d, n, C_0, C_1, the aperture α and the A_∞ constant(s).

It suffices to prove (3.2) for nonnegative harmonic functions u, because otherwise we just split u as $u = u_+ - u_-$ and use the linearity of L and the triangle inequality. Before starting to prove the theorem we need to recall some notation and preliminary results.

Lemma 3.3 (dyadic cubes for Ahlfors regular sets [David and Semmes 1991; David and Semmes 1993; Christ 1990]). *Let $\Gamma \subset \mathbb{R}^n$ be a d-Ahlfors regular set. Then there exist constants $a_0, A_1, \gamma > 0$, depending only on d, n and C_0, such that for each $k \in \mathbb{Z}$ there is a collection of Borel sets ("dyadic cubes")*

$$\mathbb{D}_k := \{Q_j^k \subset \Gamma : j \in \mathscr{J}_k\},$$

where $\mathscr{J}_k$ denotes some index set depending on k, satisfying the following properties:

(i) $\Gamma = \bigcup_{j\in\mathscr{J}_k} Q_j^k$ *for each* $k \in \mathbb{Z}$.

(ii) *If $m \ge k$ then either $Q_i^m \subset Q_j^k$ or $Q_i^m \cap Q_j^k = \varnothing$.*

(iii) *For each pair (j,k) and each $m < k$, there is a unique $i \in \mathscr{J}_m$ such that $Q_j^k \subset Q_i^m$.*

(iv) $\operatorname{diam} Q_j^k \le A_1 2^{-k}$.

(v) *Each Q_j^k contains some surface ball $\Delta(x_j^k, a_0 2^{-k}) := B(x_j^k, a_0 2^{-k}) \cap \Gamma$.*

(vi) $\mathcal{H}^d(\{q \in Q_j^k : \operatorname{dist}(q, \Gamma \setminus Q_j^k) \le \rho 2^{-k}\}) \le A_1 \rho^\gamma \mathcal{H}^d(Q_j^k)$ *for all (j,k) and all $\rho \in (0, a_0)$.*

We shall denote by $\mathbb{D} = \mathbb{D}(\Gamma)$ the collection of all relevant Q_j^k; i.e.,

$$\mathbb{D} = \bigcup_k \mathbb{D}_k.$$

Remark 3.4. (1) For a dyadic cube $Q \in \mathbb{D}$, we let $k(Q)$ denote the "dyadic generation" to which Q belongs; i.e., we set $k(Q) = k$ if $Q \in \mathbb{D}_k$. We also set its "length" to be $\ell(Q) = 2^{-k(Q)}$. Thus $\ell(Q) = 2^{-k(Q)} \sim \operatorname{diam} Q$.

(2) Properties (iv) and (v) imply that for each cube $Q \in \mathbb{D}$, there is a point $x_Q \in \Gamma$ such that

$$\Delta(x_Q, r_Q) \subset Q \subset \Delta(x_Q, C_2 r_Q), \tag{3.5}$$

where $r_Q = a_0 2^{-k(Q)} \sim \operatorname{diam} Q$ and $C_2 = A_1/a_0$.

Now we define sawtooth domains following the definitions of Hofmann and Martell; see for example [Hofmann and Martell 2014; Hofmann et al. 2016; 2017c]. Since Ω is an open set, it has a Whitney decomposition, that is, a collection of closed "Whitney" boxes in Ω, denoted by $\mathcal{W} = \mathcal{W}(\Omega)$, which form a covering of Ω with pairwise nonoverlapping interiors and satisfy

$$4 \operatorname{diam} I \le \operatorname{dist}(4I, \Gamma) \le \operatorname{dist}(I, \Gamma) \le 40 \operatorname{diam} I \quad \text{for any } I \in \mathcal{W}, \tag{3.6}$$

and also

$$\tfrac{1}{4} \operatorname{diam} I_1 \le \operatorname{diam} I_2 \le 4 \operatorname{diam} I_1 \tag{3.7}$$

whenever I_1 and I_2 in $\mathcal{W}$ touch. (See [Stein 1970] for reference.) Let X_I denote the center of I and $\ell(I)$ the side length of I; then $\operatorname{diam} I \sim \ell(I)$. We also write $k(I) = k$ if $\ell(I) = 2^{-k}$.

Let $\mathbb{D}$ be a collection of dyadic cubes for the Ahlfors regular set Γ, as in Lemma 3.3. For any dyadic cube $Q \in \mathbb{D}$, pick two parameters $\eta \ll 1$ and $K \gg 1$, and define

$$\mathcal{W}_Q^0 := \{I \in \mathcal{W} : \eta^{\frac{1}{4}} \ell(Q) \le \ell(I) \le K^{\frac{1}{2}} \ell(Q),\ \operatorname{dist}(I, Q) \le K^{\frac{1}{2}} \ell(Q)\}. \tag{3.8}$$

Let X_Q denote a corkscrew point for the surface ball $\Delta(x_Q, \frac{1}{2} r_Q)$. We can guarantee that X_Q is in some $I \in \mathcal{W}_Q^0$ provided we choose η small enough and K large enough. For each $I \in \mathcal{W}_Q^0$, by Lemma 2.1 and the discussions after that, there is a Harnack chain connecting X_I to X_Q; we call it $\mathcal{H}_I$. By the definition of $\mathcal{W}_Q^0$ we may construct this Harnack chain so that it consists of a bounded number of balls (depending on the values of η, K), and stays a distance at least $c\eta^{\frac{n-1}{4(n-1-d)}} \ell(Q)$ away from Γ; see (2.5). We let $\mathcal{W}_Q$ denote the set of all $J \in \mathcal{W}$ which meet at least one of the Harnack chains $\mathcal{H}_I$, with $I \in \mathcal{W}_Q^0$; i.e.,

$$\mathcal{W}_Q := \{J \in \mathcal{W} : \text{there exists } I \in \mathcal{W}_Q^0 \text{ for which } \mathcal{H}_I \cap J \ne \varnothing\}. \tag{3.9}$$

Clearly $\mathcal{W}_Q^0 \subset \mathcal{W}_Q$. Additionally, it follows from the construction of the augmented collections $\mathcal{W}_Q$ and the properties of the Harnack chains (in particular (2.5) and (2.6)) that there are uniform constants c and C such that

$$\begin{gathered} c\eta^{\frac{n-1}{4(n-1-d)}} \ell(Q) \le \ell(I) \le CK^{\frac{1}{2}} \ell(Q), \\ \operatorname{dist}(I, Q) \le CK^{\frac{1}{2}} \ell(Q) \end{gathered} \tag{3.10}$$

for any $I \in \mathcal{W}_Q$. In particular once η, K are fixed, for any $Q \in \mathbb{D}$ the cardinality of $\mathcal{W}_Q$ is uniformly bounded, which we denote by N_0.

Next we choose a small parameter $\theta \in (0, 1)$ so that for any $I \in \mathcal{W}$ the concentric dilation $I^* = (1+\theta)I$ still satisfies the Whitney property

$$\operatorname{diam} I \sim \operatorname{diam} I^* \sim \operatorname{dist}(I^*, \Gamma) \sim \operatorname{dist}(I, \Gamma). \tag{3.11}$$

Moreover by taking θ small enough we can guarantee that $\operatorname{dist}(I^*, J^*) \sim \operatorname{dist}(I, J)$ for every $I, J \in \mathcal{W}$, that I^* meets J^* if and only if ∂I meets ∂J and that $\frac{1}{2} J \cap I^* = \varnothing$ for any distinct $I, J \in \mathcal{W}$. In what follows we will need to work with further dilations $I^{**} = (1+2\theta)I$ or $I^{***} = (1+4\theta)I$ etc. (We may need to take θ even smaller to make sure the above properties also hold for I^{**}, I^{***} etc.) Given an arbitrary $Q \in \mathbb{D}$, we may define an associated Whitney region U_Q, U_Q^* as

$$U_Q := \bigcup_{I \in \mathcal{W}_Q} I^*, \quad U_Q^* := \bigcup_{I \in \mathcal{W}_Q} I^{**}. \tag{3.12}$$

Let $\mathbb{D}_Q = \{Q' \in \mathbb{D} : Q' \subset Q\}$. For any $Q \in \mathbb{D}$ and any family $\mathcal{F} = \{Q_j\}$ of disjoint cubes in $\mathbb{D}_Q \setminus \{Q\}$, we define the local discretized sawtooth relative to $\mathcal{F}$ by

$$\mathbb{D}_{\mathcal{F},Q} := \mathbb{D}_Q \setminus \bigcup_{Q_j \in \mathcal{F}} \mathbb{D}_{Q_j}. \tag{3.13}$$

We also define the local sawtooth domain relative to $\mathcal{F}$ by

$$\Omega_{\mathcal{F},Q} := \operatorname{int}\Big(\bigcup_{Q' \in \mathbb{D}_{\mathcal{F},Q}} U_{Q'} \Big), \quad \Omega^*_{\mathcal{F},Q} := \operatorname{int}\Big(\bigcup_{Q' \in \mathbb{D}_{\mathcal{F},Q}} U^*_{Q'} \Big). \tag{3.14}$$

For convenience we set

$$\mathcal{W}_{\mathcal{F},Q} := \bigcup_{Q' \in \mathbb{D}_{\mathcal{F},Q}} \mathcal{W}_{Q'} \tag{3.15}$$

so that in particular we may write

$$\Omega_{\mathcal{F},Q} = \operatorname{int}\Big(\bigcup_{I \in \mathcal{W}_{\mathcal{F},Q}} I^* \Big), \quad \Omega^*_{\mathcal{F},Q} = \operatorname{int}\Big(\bigcup_{I \in \mathcal{W}_{\mathcal{F},Q}} I^{**} \Big). \tag{3.16}$$

We will need further fattened sawtooth domains $\Omega^{**}_{\mathcal{F},Q}$ etc. whose definitions follow the same lines as above. We remark that by (3.10), there is a constant C_3 depending on K, θ such that

$$\Omega_{\mathcal{F},Q} \subset B(x_Q, C_3 \ell(Q)) \cap \Omega \tag{3.17}$$

for any $Q \in \mathbb{D}$ and collection of maximal cubes $\mathcal{F}$, where x_Q is the "center" of Q as in (3.5).

Finally, to work with sawtooth domains, it is more natural to use a discrete dyadic version of the approach region rather than the standard nontangential cone defined in (1.12): for every $q \in \Gamma$, we define the dyadic nontangential cones as

$$\Gamma_d(q) = \bigcup_{Q \in \mathbb{D}: Q \ni q} U_Q, \quad \widehat{\Gamma}_d(q) = \bigcup_{Q \in \mathbb{D}: Q \ni q} U_Q^{***}, \tag{3.18}$$

where we use $\widehat{\Gamma}_d$ to denote a cone with bigger "aperture" or fattened region; we also define the local dyadic nontangential cones as

$$\Gamma_d^Q(q) = \bigcup_{Q' \in \mathbb{D}_Q : Q' \ni q} U_{Q'}, \quad \widehat{\Gamma}_d^Q(q) = \bigcup_{Q' \in \mathbb{D}_Q : Q' \ni q} U_{Q'}^{***}. \tag{3.19}$$

We claim that given an aperture $\alpha > 0$, there exists K, in the definition (3.8), sufficiently large such that the standard nontangential cone $\Gamma^\alpha(q)$ satisfies $\Gamma^\alpha(q) \subset \Gamma_d(q)$ for all $q \in \Gamma$; and vice versa, for fixed values of η, K and the dilation constant θ, there exists $\alpha_1 > 0$ such that the dyadic cone $\Gamma_d(q)$

satisfies $\Gamma_d(q) \subset \Gamma^{\alpha_1}(q)$ for all $q \in \Gamma$. For any $X \in \Gamma^{\alpha}(q)$, let I be a Whitney box such that $X \in I^*$. By (3.6) we know $\ell(I) \sim \delta(X)$. Let Q be a cube containing q with length $\ell(Q) = \ell(I)$. Then

$$\operatorname{dist}(I, Q) \le |X - q| < (1+\alpha)\delta(X) \le C(1+\alpha)\ell(I) = C(1+\alpha)\ell(Q). \tag{3.20}$$

If K is sufficiently large so that $K^{\frac{1}{2}} \ge C(1+\alpha)$, then (3.20) and $\ell(I) = \ell(Q)$ implies that $I \in \mathcal{W}_Q^0$. By the definition (3.18) it follows that $X \in \Gamma_d(q)$. In particular, since $\Gamma^{\alpha}(q)$ is open, we also have $\Gamma^{\alpha}(q) \subset \operatorname{int} \Gamma_d(q)$. On the other hand, suppose $X \in \Gamma_d(q)$; by definition (3.18) X is contained in some $I^* = (1+\theta)I$ for a Whitney box $I \in \mathcal{W}_Q$ and dyadic cube Q containing q. Then by (3.10),

$$|X - q| \le \operatorname{diam} I^* + \operatorname{dist}(I, Q) + \operatorname{diam} Q \le C(K, \theta)\ell(Q),$$
$$\delta(X) \sim \ell(I) \ge C(\eta)\ell(Q).$$

Therefore there exists α_1 sufficiently large, depending on the values of η, K, θ, such that

$$|X - q| < (1+\alpha_1)\delta(X);$$

i.e., $X \in \Gamma^{\alpha_1}(q)$. We summarize that now we have

$$\Gamma^{\alpha}(q) \subset \operatorname{int} \Gamma_d(q) \subset \Gamma_d(q) \subset \Gamma^{\alpha_1}(q) \quad \text{for all } q \in \Gamma. \tag{3.21}$$

Clearly $\alpha_1 > \alpha$. Moreover, there exists $\beta > \alpha_1$ depending on η, K, θ such that the fattened dyadic nontangential cone $\widehat{\Gamma}_d(q)$ satisfies

$$\widehat{\Gamma}_d(q) \subset \Gamma^{\beta}(q) \quad \text{for all } q \in \Gamma. \tag{3.22}$$

From now on we fix the values of η, K, θ and $\beta > \alpha_1 > \alpha > 0$.

Let $F = Q \setminus \bigcup_{Q_j \in \mathcal{F}} Q_j$ and suppose it is not empty. We claim that

$$\operatorname{int}\biggl(\bigcup_{q \in F} \Gamma_d^Q(q)\biggr) \subset \Omega_{\mathcal{F},Q} \subset \overline{\Omega}_{\mathcal{F},Q} \subset \Omega_{\mathcal{F},Q}^{***} \subset \bigcup_{q \in F} \widehat{\Gamma}_d^Q(q). \tag{3.23}$$

In fact, for any $q \in F$, it is clear that q is in some $Q' \in \mathbb{D}_{\mathcal{F},Q}$; and by (3.14), the definition of $\Omega_{\mathcal{F},Q}$, we have the first inclusion. On the other hand any $X \in \Omega_{\mathcal{F},Q}^{***}$ belongs to some $U_{Q'}^{***}$ with $Q' \in \mathbb{D}_{\mathcal{F},Q}$, and thus $X \in \widehat{\Gamma}_d^Q(q)$ for arbitrary $q \in Q'$. By the definition of $\mathbb{D}_{\mathcal{F},Q}$, we know $Q' \cap F \neq \varnothing$, so by taking $q \in Q' \cap F$ we get $X \in \bigcup_{q \in F} \widehat{\Gamma}_d^Q(q)$.

For N sufficiently large, we augment the collection of maximal cubes $\mathcal{F}$ by adding all dyadic cubes in $\mathbb{D}$ of size smaller than or equal to $2^{-N}\ell(Q)$, and we denote by $\mathcal{F}^N$ a collection consisting of all maximal cubes of the above augmented collection. In particular $Q' \in \mathbb{D}_{\mathcal{F}^N,Q}$ if and only if $Q' \in \mathbb{D}_{\mathcal{F},Q}$ and $\ell(Q') > 2^{-N}\ell(Q)$. By doing this we guarantee that the sawtooth domain $\Omega_{\mathcal{F}^N,Q}$ is compactly contained in Ω (roughly speaking $\operatorname{dist}(\Omega_{\mathcal{F}^N,Q}, \Omega^c) \sim 2^{-N}\ell(Q)$). Similar to Lemma 4.44 of [Hofmann et al. 2017c], we can construct a smooth cutoff function of $\Omega_{\mathcal{F}^N,Q}$:

Lemma 3.24 (cut-off function of sawtooth domain). *There exists $\psi_N \in C_0^{\infty}(\mathbb{R}^n)$ such that*:

(i) $\chi_{\Omega_{\mathcal{F}^N,Q}^{*}} \lesssim \psi_N \le \chi_{\Omega_{\mathcal{F}^N,Q}^{**}}$.

(ii) $\sup_{X \in \Omega} |\nabla \psi_N(X)|\delta(X) \lesssim 1$.

(iii) *We abbreviate* $\mathcal{W}_{\mathcal{F}^N,Q}$ *as* $\mathcal{W}_N$ *and set* $\Sigma = \partial\Omega^*_{\mathcal{F}^N,Q}$,

$$\mathcal{W}_N^\Sigma = \{I \in \mathcal{W}_N : \text{there exists } J \in \mathcal{W} \setminus \mathcal{W}_N \text{ with } \partial I \cap \partial J \neq \varnothing\}.$$

Then

$$\nabla \psi_N \equiv 0 \quad \text{in} \bigcup_{I \in \mathcal{W}_N \setminus \mathcal{W}_N^\Sigma} I^{***}. \tag{3.25}$$

(iv) *For each* $I \in \mathcal{W}_N$, *let* Q_I *denote a cube in* $\mathbb{D}_{\mathcal{F}^N,Q}$ *such that* $I \in \mathcal{W}_{Q_I}$. *Suppose* ω *is the harmonic measure with pole* X_0 *and* X_0 *satisfies* $\operatorname{dist}(X_0, \Omega^{***}_{\mathcal{F}^N,Q}) \gtrsim \ell(Q)$. *Then*

$$\sum_{I \in \mathcal{W}_N^\Sigma} \omega(Q_I) \lesssim \omega(Q), \tag{3.26}$$

with a constant depending on η, K, a_0, C_1, d *and the Ahlfors regular constant of* Γ.

Remark 3.27. (1) We remark that the construction of ψ_N and the proof of its properties (i), (ii), (iii) are higher codimensional analogues of Lemma 4.44 of [Hofmann et al. 2017c]. However we prove (iv) instead of the second estimate in their (4.46) because we will need to prove a good-λ inequality for the harmonic measure, instead of the surface measure. Since harmonic measure could have much worse decay properties than the surface measure, not to mention that Γ and $\partial\Omega_{\mathcal{F}^N,Q}$ are objects of different dimensions, proving (iv) requires a different argument.

(2) Note that in (iv), the choice of Q_I may not be unique. Suppose both $Q_I, \widetilde{Q}_I$ are cubes in $\mathbb{D}_{\mathcal{F}^N,Q}$ such that $I \in \mathcal{W}_{Q_I}$ and $I \in \mathcal{W}_{\widetilde{Q}_I}$. By the construction of the $\mathcal{W}_Q$'s and in particular (3.10), we know

$$\ell(Q_I) \sim \ell(I) \sim \ell(\widetilde{Q}_I), \quad \operatorname{dist}(Q_I, \widetilde{Q}_I) \lesssim \ell(Q_I), \tag{3.28}$$

with constants depending on η, K. Since harmonic measure is doubling, we have

$$C_1\omega(Q_I) \le \omega(\widetilde{Q}_I) \le C_2\omega(Q_I), \tag{3.29}$$

with constants only depending on the doubling constant and η, K. That is to say, for different choices of Q_I the left-hand side of (3.26) differs at most by a constant multiple. But once we associate a cube Q_I to I, the choice will be fixed.

Proof. The proof of (i) is a modification of the proof from [Hofmann et al. 2017c] in higher codimensions. We recall that given I any closed dyadic cube in $\mathbb{R}^n$, we set $I^{**} = (1+2\theta)I$ and $I^{***} = (1+4\theta)I$. Let us introduce $\widetilde{I}^{**} = (1+3\theta)I$ so that

$$I^{**} \subsetneq \operatorname{int} \widetilde{I}^{**} \subsetneq \widetilde{I}^{**} \subset \operatorname{int} I^{***}. \tag{3.30}$$

Given $I_0 = \left[-\frac{1}{2}, \frac{1}{2}\right]^n \subset \mathbb{R}^n$, we fix $\phi_0 \in C_0^\infty(\mathbb{R}^n)$ such that $\chi_{I_0^{**}} \le \phi_0 \le \chi_{\widetilde{I}_0^{**}}$ and $|\nabla\phi_0| \lesssim 1$, with the implicit constant depending on θ. For every $I \in \mathcal{W}$ we set $\phi_I = \phi_0((\cdot - X_I)/\ell(I))$, where X_I is the center of I, so that $\phi_I \in C_0^\infty(\mathbb{R}^n)$, $\chi_{I^{**}} \le \phi_I \le \chi_{\widetilde{I}^{**}}$ and $|\nabla\phi_I| \lesssim 1/\ell(I)$. Let $\Phi(X) := \sum_{I\in\mathcal{W}} \phi_I(X)$ for every $X \in \Omega$. Since for each compact subset of Ω the previous sum has finitely many nonvanishing terms, we have $\Phi \in C^\infty_{\text{loc}}(\Omega)$. Also $0 \le \Phi(X) \lesssim C_\theta$ since the family $\{\widetilde{I}^{**}\}_{I\in\mathcal{W}}$ has bounded overlap.

Hence we can set $\Phi_I = \phi_I/\Phi$ and one can easily see that $\Phi_I \in C_0^\infty(\mathbb{R}^n)$, $C_\theta^{-1}\chi_{I^{**}} \le \Phi_I \le \chi_{\widetilde{I}^{**}}$ and $|\nabla\Phi_I| \lesssim 1/\ell(I)$. Recalling the definition of $\mathcal{W}_N = \mathcal{W}_{\mathcal{F}^N,Q}$ in (3.15), we set

$$\psi_N(X) = \sum_{I\in\mathcal{W}_N} \Phi_I(X) = \frac{\sum_{I\in\mathcal{W}_N}\phi_I(X)}{\sum_{I\in\mathcal{W}}\phi_I(X)}, \quad X \in \Omega. \tag{3.31}$$

We first note that the number of terms in the sum defining ψ_N is bounded depending on N. Indeed if $Q' \in \mathbb{D}_{\mathcal{F}^N,Q}$ then $Q' \in \mathbb{D}_Q$ and $2^{-N}\ell(Q) < \ell(Q') \le \ell(Q)$, which implies $\mathbb{D}_{\mathcal{F}^N,Q}$ has finite cardinality with bound depending only on the Ahlfors regular constant and N. Also by construction $\mathcal{W}_Q$ has cardinality depending only on the allowable parameters η, K. Hence $\#\mathcal{W}_N \le C_N < \infty$. This and the fact that $\Phi_I \in C_0^\infty(\mathbb{R}^n)$ for each I yield that $\psi_N \in C_0^\infty(\mathbb{R}^n)$. Moreover

$$\operatorname{supp}\psi_N \subset \bigcup_{I\in\mathcal{W}_N} \widetilde{I}^{**} = \bigcup_{Q'\in\mathbb{D}_{\mathcal{F}^N,Q}} \bigcup_{I\in\mathcal{W}_Q} \widetilde{I}^{**} \subset \operatorname{int}\Bigg(\bigcup_{Q'\in\mathbb{D}_{\mathcal{F}^N,Q}} U_{Q'}^{**}\Bigg) = \Omega_{\mathcal{F}^N,Q}^{**}. \tag{3.32}$$

This and the definition of ψ_N immediately give $\psi_N \le \chi_{\Omega_{\mathcal{F}^N,Q}^{**}}$. On the other hand, if $X \in \Omega_{\mathcal{F}^N,Q}^{*}$ then there exists $I \in \mathcal{W}_N$ such that $X \in I^{**}$, in which case we have $\psi_N(X) \ge \Phi_I(X) \ge C_\theta^{-1}$. This completes the proof of (i).

To obtain (ii) we note that for every $X \in \Omega$

$$|\nabla\psi_N(X)| \le \sum_{I\in\mathcal{W}_N} |\nabla\Phi_I(X)| \lesssim \sum_{I\in\mathcal{W}} \frac{1}{\ell(I)}\chi_{\widetilde{I}^{**}}(X) \lesssim \frac{1}{\delta(X)}, \tag{3.33}$$

where we have used that if $X \in \widetilde{I}^{**}$ then $\ell(I) \sim \delta(I)$ and also that the family $\{\widetilde{I}^{**}\}_{I\in\mathcal{W}}$ has bounded overlap.

Now we turn to (iii). Fix $I \in \mathcal{W}_N \setminus \mathcal{W}_N^\Sigma$ and $X \in I^{***}$, and set $\mathcal{W}_X = \{J \in \mathcal{W} : \phi_J(X) \ne 0\}$. We first note that $\mathcal{W}_X \subset \mathcal{W}_N$. Indeed if $\phi_J(X) \ne 0$ then $X \in \widetilde{J}^{**}$. Hence $X \in I^{***} \cap J^{***}$ and our choice of θ gives that ∂I meets ∂J; this in turn implies that $J \in \mathcal{W}_N$ since $I \notin \mathcal{W}_N^\Sigma$. All these imply

$$\psi_N(X) = \frac{\sum_{J\in\mathcal{W}_N}\phi_J(X)}{\sum_{J\in\mathcal{W}}\phi_J(X)} = \frac{\sum_{J\in\mathcal{W}_N\cap\mathcal{W}_X}\phi_J(X)}{\sum_{J\in\mathcal{W}\cap\mathcal{W}_X}\phi_J(X)} = \frac{\sum_{J\in\mathcal{W}_N\cap\mathcal{W}_X}\phi_J(X)}{\sum_{J\in\mathcal{W}_N\cap\mathcal{W}_X}\phi_J(X)} = 1. \tag{3.34}$$

Hence $\psi_N|_{I^{***}} \equiv 1$ for every $I \in \mathcal{W}_N \setminus \mathcal{W}_N^\Sigma$. This and the bounded overlap of the family $\{I^{***}\}_{I\in\mathcal{W}_N}$ immediately give that $\nabla\psi_N \equiv 0$ in $\bigcup_{I\in\mathcal{W}_N\setminus\mathcal{W}_N^\Sigma} I^{***}$.

Finally, it remains to prove the most difficult property, (iv). For any $I \in \mathcal{W}_N^\Sigma$, by definition there exists some $J_I \in \mathcal{W} \setminus \mathcal{W}_N$ such that $\partial I \cap \partial J_I \ne \varnothing$. Roughly speaking, this is to say that I is a Whitney box living in the "boundary" of $\Omega_{\mathcal{F}^N,Q}^*$. Thus picking any $Q_I' \in \mathbb{D}$ such that $\mathcal{W}_{Q_I'}$ contains J_I, we know $Q_I' \notin \mathbb{D}_{\mathcal{F}^N,Q}$, that is, either $Q_I' \in \mathbb{D}_{Q_j}$ for some $Q_j \in \mathcal{F}^N$, or $Q_I' \notin \mathbb{D}_Q$. We classify $I \in \mathcal{W}_N^\Sigma$ based on which category its associated cube Q_I' lives in: we define

$$\Sigma_j = \{I \in \mathcal{W}_N^\Sigma : Q_I' \in \mathbb{D}_{Q_j}\} \quad \text{for any } Q_j \in \mathcal{F}^N,$$
$$\Sigma_0 = \{I \in \mathcal{W}_N^\Sigma : Q_I' \notin \mathbb{D}_Q\}.$$

(Note that for each $I \in \mathcal{W}_N^\Sigma$, we associate it to a unique Q_I', even though the choice itself is not unique.) Recalling (3.7) we have $\ell(I) \sim \ell(J_I)$. Moreover by the definition of $\mathcal{W}_Q$ and (3.10),

$$\ell(Q_I') \sim \ell(J_I) \sim \ell(I) \sim \ell(Q_I), \tag{3.35}$$

$$\operatorname{dist}(Q_I, Q_I') \le \operatorname{dist}(Q_I, I) + \operatorname{dist}(I, J_I) + \operatorname{dist}(J_I, Q_I') \lesssim \ell(Q_I) + \ell(Q_I') \lesssim \ell(Q_I'). \tag{3.36}$$

By a similar argument to that in Remark 3.27(2) and the doubling property of harmonic measure, we have $\omega(Q_I) \sim \omega(Q_I')$ for any $I \in \mathcal{W}_N^\Sigma$, with a uniform constant depending on η, K. Therefore to prove (3.26) it suffices to show

$$\sum_{I \in \mathcal{W}_N^\Sigma} \omega(Q_I') \lesssim \omega(Q).$$

We claim that for any $Q_j \in \mathcal{F}^N$,

$$\sum_{I \in \Sigma_j} \omega(Q_I') \lesssim \omega(Q_j). \tag{3.37}$$

Recall that all such Q_I''s live in $\mathbb{D}_{Q_j}$. For each $k \in \mathbb{N}$ we define $\Sigma_j^k = \{I \in \Sigma_j : \ell(Q_I') = 2^{-k}\ell(Q_j)\}$. Since $Q_I \in \mathbb{D}_{\mathcal{F}^N, Q}$, $Q_j \in \mathcal{F}^N$, we always have $Q_j \cap Q_I = \varnothing$, so by (3.36)

$$\operatorname{dist}(Q_I', (Q_j)^c) \le \operatorname{dist}(Q_I', Q_I) \lesssim \ell(Q_I') = 2^{-k}\ell(Q_j). \tag{3.38}$$

That is, the smaller Q_I' is, the closer it is to the "boundary" of Q_j. The Q_I''s of different generations are very far from being disjoint; however we will sum up the $\omega(Q_I')$'s by swapping them for the harmonic measure of mutually disjoint cubes. By (3.38), for ρ sufficiently small there is an integer $k_1 = k_1(\rho)$ such that for any integer $k \ge k_1$

$$\bigcup_{k' \ge k} \bigcup_{I \in \Sigma_j^{k'}} Q_I' \subset \bigl\{q \in Q_j : \operatorname{dist}(q, (Q_j)^c) \le \tfrac{1}{2}\rho\ell(Q_j)\bigr\}. \tag{3.39}$$

In fact by choosing k_1 slightly bigger, we can even guarantee that for any integer $k \ge k_1$,

$$\bigcup_{k' \ge k} \bigcup_{I \in \Sigma_j^{k'}} Q_I' \subset \bigcup_{i \in \mathcal{I}_k} Q_j^i \subset \bigl\{q \in Q_j : \operatorname{dist}(q, (Q_j)^c) \le \tfrac{1}{2}\rho\ell(Q_j)\bigr\}, \tag{3.40}$$

where $\{Q_j^i\}_{i \in \mathcal{I}_k}$ is the collection of all dyadic cubes in $\mathbb{D}_{Q_j}$ of length $2^{-k}\ell(Q_j)$ such that $Q_j^i \subset \bigl\{q \in Q_j : \operatorname{dist}(q, (Q_j)^c) \le \frac{1}{2}\rho\ell(Q_j)\bigr\}$. By Lemma 3.3(v)–(vi) the index set $\mathcal{I}_k$ has finite cardinality and $\#\mathcal{I}_k \le C2^{kd}$. (A priori the set $\mathcal{I}_k$ could be empty, in which case (3.40) just means there is no Q_I' corresponding to any $I \in \bigcup_{k' \ge k} \Sigma_j^{k'}$. This case is easy to deal with.)

On the other hand by Lemma 3.3, as long as we fix $\rho \in (0, a_0)$ satisfying $A_1\rho^\gamma < 1$, the set $\{q \in Q_j : \operatorname{dist}(q, (Q_j)^c) > \frac{1}{2}\rho\ell(Q_j)\}$ is not empty; moreover, there is an integer k_2 sufficiently large such that for each $k \ge k_2$ we can find a cube $\widehat{Q}_j$ such that $\ell(\widehat{Q}_j) = 2^{-k}\ell(Q_j)$ and

$$\widehat{Q}_j \subset \bigl\{q \in Q_j : \operatorname{dist}(q, (Q_j)^c) > \tfrac{1}{2}\rho\ell(Q_j)\bigr\}. \tag{3.41}$$

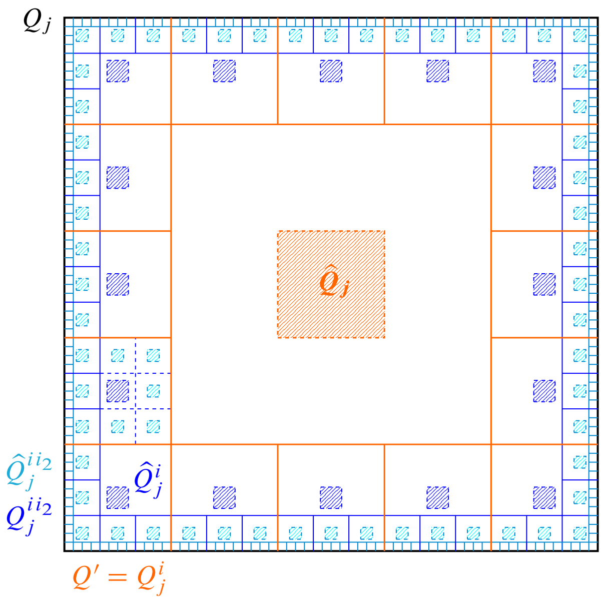

Figure 1. Illustration of the swap of cubes in iteration.

We may think of $\widehat{Q}_j$ as sitting in the "center" of Q_j, and all Q_I''s as being in a $\left(\frac{1}{2}\rho\right)$-boundary layer of Q_j. Let $k_0 = \max\{k_1, k_2\}$, and let N_1 denote the (maximal) number of Q_I''s with $\ell(Q_I') = 2^{-k_0}\ell(Q_j)$. By (3.39) and Lemma 3.3(vi), N_1 is uniformly bounded by a constant depending on a_0, A_1, ρ, k_0 and d. Moreover by the doubling property of ω, each such Q_I' satisfies

$$\omega(Q_I') \le \omega(Q_j) \le C(k_0)\omega(\widehat{Q}_j), \tag{3.42}$$

with the constant $C(k_0)$ depending on k_0 as well as the doubling constant of ω. Recall that for each Q_I', the number of all possible I's corresponding to it is uniformly bounded by $C(N_0)$. Therefore

$$\sum_{I \in \Sigma_j^{k_0}} \omega(Q_I') \le C(N_0) \sum_{Q_I' : \ell(Q_I') = 2^{-k_0}\ell(Q_j)} \omega(Q_I') \le C(N_0) N_1 C(k_0)\omega(\widehat{Q}_j). \tag{3.43}$$

Now for any $I \in \Sigma_j^k$ with $k = 1, \dots, k_0 - 1$, again by the doubling property of harmonic measure we have $\omega(Q_I') \le C(k_0)\omega(\widehat{Q}_j)$. By Lemma 3.3(iv)–(v), the total number of Q_I''s in $\mathbb{D}_{Q_j}$ such that $\ell(Q_I') = 2^{-k}\ell(Q_j)$, with $k = 1, \dots, k_0 - 1$, is uniformly bounded by a constant depending only on k_0, a_0, C_1, d and the Ahlfors regular constant of Γ. Thus the total number of I's in Σ_j^k with $k = 1, \dots, k_0 - 1$ is also uniformly bounded. Therefore combining with (3.43), we get

$$\sum_{k=1}^{k_0} \sum_{I \in \Sigma_j^k} \omega(Q_I') \lesssim \omega(\widehat{Q}_j). \tag{estimate k_0}$$

For future generations, we recall (3.40), which says all the Q_I''s corresponding to some $I \in \Sigma_j^k$, with $k \ge k_0$, are contained in $\bigcup_{i \in \mathcal{I}_{k_0}} Q_j^i$. The following proof is illustrated in the (idealized) Figure 1, where each label denotes the cube near it enclosed or shaded by the same color. Consider any cube

$Q' = Q_j^i$ for an arbitrary $i \in \mathcal{I}_{k_0}$. Applying the above argument to Q' in place of Q_j, we can find a cube $\widehat{Q}' = \widehat{Q}_j^i \in \mathbb{D}_{Q'}$ with length $\ell(\widehat{Q}') = 2^{-k_0}\ell(Q') = 2^{-2k_0}\ell(Q_j)$ sitting in the "center" of Q', in the sense that

$$\widehat{Q}_j^i \subset \{q \in Q' : \operatorname{dist}(q, (Q')^c) > \tfrac{1}{2}\rho\ell(Q')\}, \tag{3.44}$$

and all future generations satisfy

$$\bigcup_{k \geq 2k_0} \bigcup_{\substack{I \in \Sigma_j^k \\ Q_I' \in \mathbb{D}_{Q'}}} Q_I' \subset \bigcup_{i_2 \in \mathcal{I}_{k_0}} Q_j^{i i_2} \subset \{q \in Q' : \operatorname{dist}(q, (Q')^c) \leq \tfrac{1}{2}\rho\ell(Q')\}, \tag{3.45}$$

where $\{Q_j^{i i_2}\}_{i_2 \in \mathcal{I}_{k_0}}$ is the collection of all dyadic cubes of length $2^{-k_0}\ell(Q') = 2^{-2k_0}\ell(Q_j)$ that are completely contained in $\{q \in Q' = Q_j^i : \operatorname{dist}(q, (Q')^c) \leq \frac{1}{2}\rho\ell(Q')\}$. (The index set for i_2 may not be the same as the index set for i, but their cardinalities are uniformly bounded by $C2^{k_0 d}$, so we abuse the notation here and simply assume they are the same.) Moreover we can get an analogous estimate of (estimate k_0)

$$\sum_{k=k_0+1}^{2k_0} \sum_{\substack{I \in \Sigma_j^k \\ Q_I' \in \mathbb{D}_{Q'}}} \omega(Q_I') \lesssim \omega(\widehat{Q}_j^i). \tag{3.46}$$

Summing up (3.46) over all cubes $Q' \in \{Q_j^i\}_{i \in \mathcal{I}_{k_0}}$, recalling (3.40) we get

$$\sum_{k=k_0+1}^{2k_0} \sum_{I \in \Sigma_j^k} \omega(Q_I') \lesssim \sum_{i \in \mathcal{I}_{k_0}} \omega(\widehat{Q}_j^i). \tag{3.47}$$

Since $\{Q_j^i\}_{i \in \mathcal{I}_{k_0}}$ is a collection of cubes in the same generation, they are mutually disjoint, and their subcubes $\{\widehat{Q}_j^i\}_{i \in \mathcal{I}_{k_0}}$ are also mutually disjoint. Hence

$$\sum_{k=k_0+1}^{2k_0} \sum_{I \in \Sigma_j^k} \omega(Q_I') \lesssim \sum_{i \in \mathcal{I}_{k_0}} \omega(\widehat{Q}_j^i) = \omega\Big(\bigsqcup_{i \in \mathcal{I}_{k_0}} \widehat{Q}_j^i\Big). \tag{estimate $2k_0$}$$

Moreover, recalling the second inclusion of (3.40) and (3.41), each $\widehat{Q}_j^i$ is disjoint from $\widehat{Q}_j$, so we can add up (estimate k_0) and (estimate $2k_0$) with ease. We can repeat this argument iteratively: for any $l \in \mathbb{N}$ we apply the argument to cube $Q' = Q_j^{i_1 i_2 \cdots i_l}$, with $i_1, \ldots, i_l \in \mathcal{I}_{k_0}$, to get an analogous estimate of (3.46); then we sum up over the index sets and get

$$\sum_{k=lk_0+1}^{(l+1)k_0} \sum_{I \in \Sigma_j^k} \omega(Q_I') \lesssim \sum_{i_1, \ldots, i_l \in \mathcal{I}_{k_0}} \omega(\widehat{Q}_j^{i_1 \cdots i_l}) = \omega\Big(\bigsqcup_{i_1, \ldots, i_l \in \mathcal{I}_{k_0}} \widehat{Q}_j^{i_1 \cdots i_l}\Big). \tag{estimate $(l+1)k_0$}$$

Most significantly for us, for each $l \in \mathbb{N}$ the union of cubes on the right-hand side of (estimate $(l+1)k_0$) is disjoint from all the cubes from all previous summations. Therefore we conclude that

$$\sum_{k=1}^{\infty} \sum_{I \in \Sigma_j^k} \omega(Q_I') \lesssim \omega\bigg(\bigsqcup_{l \in \mathbb{N}} \Big(\bigsqcup_{i_1, \ldots, i_l \in \mathcal{I}_{k_0}} \widehat{Q}_j^{i_1 \cdots i_l}\Big)\bigg) \leq \omega(Q_j). \tag{3.48}$$

It is trivial to see $\sum_{I\in\Sigma_j^0}\omega(Q_I')\lesssim\omega(Q_j)$, so

$$\sum_{I\in\Sigma_j}\omega(Q_I')=\sum_{k\in\mathbb{N}}\sum_{I\in\Sigma_j^k}\omega(Q_I')\lesssim\omega(Q_j). \tag{3.49}$$

Since the maximal cubes Q_j in $\mathcal{F}^N$ are mutually disjoint and contained in Q, we have

$$\sum_{Q_j\in\mathcal{F}^N}\sum_{I\in\Sigma_j}\omega(Q_I')\lesssim\sum_{Q_j\in\mathcal{F}^N}\omega(Q_j)\le\omega(Q). \tag{3.50}$$

Now we consider $I\in\Sigma_0$, which by definition means $Q_I'\notin\mathbb{D}_Q$. Recalling (3.35) and (3.36), and that $\ell(I)\le C\ell(Q)$ for all $I\in\mathcal{W}_N=\mathcal{W}_{\mathcal{F}^N,Q}$, we have

$$\ell(Q_I')\sim\ell(I)\le C\ell(Q),\quad \operatorname{dist}(Q_I,Q_I')\lesssim\ell(Q_I')\le C\ell(Q). \tag{3.51}$$

In particular since $Q_I\in\mathbb{D}_Q$, we have

$$\operatorname{dist}(Q_I',Q)\le\operatorname{dist}(Q_I',Q_I)\lesssim\ell(Q_I')\le C\ell(Q). \tag{3.52}$$

If $\ell(Q_I')\ge\ell(Q)$, then

$$\ell(Q_I')\sim\ell(Q),\quad \operatorname{dist}(Q_I',Q)\le C\ell(Q). \tag{3.53}$$

There are finitely many such Q_I''s and by the doubling property of harmonic measure, $\omega(Q_I')\sim\omega(Q)$. If $\ell(Q_I')<\ell(Q)$, let $Q_0\in\mathbb{D}$ be the cube containing Q_I' with length $\ell(Q_0)=\ell(Q)$. By the assumption $Q_I'\notin\mathbb{D}_Q$, we know Q_0 is disjoint from Q. On the other hand (3.52) implies

$$\operatorname{dist}(Q_0,Q)\le\operatorname{dist}(Q_I',Q)\le C\ell(Q); \tag{3.54}$$

that is, Q_0 is a sibling (i.e., of the same generation) of Q in a $C\ell(Q)$-neighborhood of Q. There are finitely many such Q_0's. Moreover

$$\operatorname{dist}(Q_I',(Q_0)^c)\le\operatorname{dist}(Q_I',Q)\lesssim\ell(Q_I'). \tag{3.55}$$

So if $\ell(Q_I')\ll\ell(Q)$, we can guarantee that Q_I' lies in the $\left(\frac{1}{2}\rho\right)$-boundary layer of Q_0: $Q_I'\subset\left\{q\in Q_0:\operatorname{dist}(q,(Q_0)^c)\le\frac{1}{2}\rho\ell(Q_0)\right\}$. Applying the same argument to Q_0 in place of Q_j, we get

$$\sum_{\substack{I\in\Sigma_0\\ Q_I'\in\mathbb{D}_{Q_0}}}\omega(Q_I')\lesssim\omega(Q_0)\sim\omega(Q). \tag{3.56}$$

Summing up (3.56) over all (finitely many) Q_0's satisfying (3.54), we get

$$\sum_{I\in\Sigma_0}\omega(Q_I')\lesssim\omega(Q). \tag{3.57}$$

Finally we combine (3.50) and (3.57) and conclude that

$$\sum_{I\in\mathcal{W}_N^\Sigma}\omega(Q_I')=\sum_{Q_j\in\mathcal{F}^N}\sum_{I\in\Sigma_j}\omega(Q_I')+\sum_{I\in\Sigma_0}\omega(Q_I')\lesssim\omega(Q). \tag{3.58}$$

Therefore

$$\sum_{I\in\mathcal{W}_N^\Sigma}\omega(Q_I')\lesssim\omega(Q),\tag{3.59}$$

concluding the proof. □

Now that all the preparatory work has been done, we proceed to sketch the basic idea for the proof of Theorem 3.1. It is well known in harmonic analysis that the proof of $\|Su\|_{L^p(\sigma)}\le C\|Nu\|_{L^p(\sigma)}$ can be reduced to the proof of a certain good-λ inequality measured by σ. We first prove Proposition 1.16, which is a good-λ inequality measured by ω; then we use the assumption $\omega\in A_\infty(\sigma)$ to obtain the desired good-λ inequality for σ.

Recall that we use $Su, S'u, S''u$ to denote the square functions on standard nontangential cones of apertures α,α_1,β, respectively, and Nu to denote the nontangential maximal function on cones of aperture β, where $\beta>\alpha_1>\alpha$ are fixed apertures (see the discussion before Lemma 3.24). Also recall from (3.17) that for any collection $\mathcal{F}$ of dyadic cubes, the sawtooth domain $\Omega_{\mathcal{F},Q}$ satisfies $\Omega_{\mathcal{F},Q}\subset B(x_Q,C_3\ell(Q))\cap\Omega$. In fact, by choosing a slightly bigger constant C_3 we can also guarantee $\Omega^{***}_{\mathcal{F},Q}\subset B(x_Q,C_3\ell(Q))\cap\Omega$.

Proof of Proposition 1.16. For simplicity we set $\omega=\omega^{X_Q}$. Let $E=\{q\in Q: Su(q)>2\lambda,\ Nu(q)\le\delta\lambda\}$ and $F=\{q\in Q:Nu(q)\le\delta\lambda\}$. If F is empty, then the left-hand side of (1.17) is zero, and there is nothing to prove. So we assume $F\neq\varnothing$. Note that $Nu(q)$ is a continuous function, so $Q\setminus F=\{q\in Q:Nu(q)>\delta\lambda\}$ is relatively open in Q. We run a stopping-time procedure for the descendants of Q, and stop at $Q'\in\mathbb{D}_Q$ whenever $Nu(q)>\delta\lambda$ for all $q\in Q'$. We denote the collection of all maximal cubes by $\mathcal{F}_2=\{Q_j\}\subset\mathbb{D}_Q\setminus\{Q\}$. We claim that they form a partition:

$$Q\setminus F=\{q\in Q:Nu(q)>\delta\lambda\}=\bigcup_{Q_j\in\mathcal{F}_2}Q_j.\tag{3.60}$$

Clearly by construction $\bigcup_{Q_j\in\mathcal{F}_2}Q_j$ is contained in the set on the left. For any $q_0\in Q$ such that $Nu(q_0)>\delta\lambda$ (since the set $\{q\in\Gamma:Nu(q)>\delta\lambda\}$ is open), $Q\setminus F\neq Q$ and the cubes in $\mathbb{D}$ are nested, there exists a small cube $Q'\in\mathbb{D}_Q\setminus\{Q\}$ containing q_0 such that $Nu(q)>\delta\lambda$ for all $q\in Q'$. By the stopping-time procedure, either $Q'\in\mathcal{F}_2$, or Q' is contained in some cube $Q_j\in\mathcal{F}_2$. Hence $q_0\in Q'\subset\bigcup_{Q_j\in\mathcal{F}_2}Q_j$, and we prove the claim (3.60). Recall (3.23), which we rewrite here:

$$\operatorname{int}\Big(\bigcup_{q\in F}\Gamma_d^Q(q)\Big)\subset\Omega_{\mathcal{F}_2,Q}\subset\overline{\Omega}_{\mathcal{F}_2,Q}\subset\Omega^{***}_{\mathcal{F}_2,Q}\subset\bigcup_{q\in F}\widehat{\Gamma}_d^Q(q).\tag{3.61}$$

We claim that $|u(X)|\le\delta\lambda$ for all $X\in\Omega^{***}_{\mathcal{F}_2,Q}$. In fact, by (3.22) and (3.61) we know that every $X\in\Omega^{***}_{\mathcal{F}_2,Q}$ is contained in some $\widehat{\Gamma}_d^Q(q)\subset\Gamma^\beta(q)$ for some $q\in F$. Since $Nu(q)=\sup_{X\in\Gamma^\beta(q)}|u(X)|\le\delta\lambda$ for $q\in F$, we get $|u(X)|\le\delta\lambda$.

<u>Step 1</u>: Recall the assumption that $S'u(q_1)\le\lambda$ for some q_1 satisfying $|q_1-q|\le C_2\operatorname{diam}Q$ for all $q\in Q$. Set $r=\operatorname{diam}Q$. We claim that for any $\tau>0$ there exists $\delta>0$ sufficiently small such that the truncated square function $S_{\tau r}u(q)$ is greater than λ for any $q\in E$.

Fix $q \in E$. Recall that $Su(q) > 2\lambda$ for $q \in E$. We set $U = \Gamma^\alpha(q) \setminus B(q, \tau r)$; then we aim to show

$$\iint_U |\nabla u(X)|^2 \, \delta(X)^{1-d} \, dm(X) \leq 3\lambda^2. \tag{3.62}$$

Let $U_1 = \Gamma^\alpha(q) \setminus B(q, tr)$ for a constant $t > \tau$ to be chosen later, and $U_2 = \Gamma^\alpha(q) \cap (B(q, tr) \setminus B(q, \tau r))$. Then $U = U_1 \cup U_2$. A simple computation shows that

$$U_1 = \Gamma^\alpha(q) \setminus B(q, tr) \subset \Gamma^{\alpha_1}(q_1) \tag{3.63}$$

if the apertures satisfy

$$(1+\alpha)\left(1 + \frac{C_2}{t}\right) \leq 1 + \alpha_1,$$

that is, if t is sufficiently large such that

$$\alpha + \frac{C_2(1+\alpha)}{t} \leq \alpha_1.$$

Therefore

$$\begin{aligned} \iint_{U_1} |\nabla u(X)|^2 \, \delta(X)^{1-d} \, dm(X) &\leq \iint_{\Gamma^{\alpha_1}(q_1)} |\nabla u(X)|^2 \, \delta(X)^{1-d} \, dm(X) \\ &= S'u(q_1)^2 \leq \lambda^2. \end{aligned} \tag{3.64}$$

Let $\Gamma_j(q) = \Gamma^\alpha(q) \cap (B(q, 2^j \tau r) \setminus B(q, 2^{j-1}\tau r))$ for $j = 1, 2, \ldots$; then

$$U_2 \subset \bigcup_{j : 2^{j-1}\tau r < tr} \Gamma_j(q).$$

Each $\Gamma_j(q)$ can be covered by a finite union (depending on n) of balls $B_{j,k}$ with radius $r_{j,k} \sim_\alpha 2^j \tau r$. Let $B^*_{j,k}$ denote a slight fattening of $B_{j,k}$ such that we still have $B^*_{j,k} \subset \Gamma^\beta(q)$; then by Lemma 2.7(i), $m(B^*_{j,k}) \sim r_{j,k}^{d+1} \sim (2^j \tau r)^{d+1}$. Thus

$$\begin{aligned} \iint_{U_2} |\nabla u(X)|^2 \, \delta(X)^{1-d} \, dm(X) &= \sum_{2^{j-1}\tau r < tr} \iint_{\Gamma_j(q)} |\nabla u(X)|^2 \, \delta(X)^{1-d} \, dm(X) \\ &\sim_{\alpha,\beta} \sum_{2^{j-1}\tau r < tr} (2^j \tau r)^{1-d} \sum_{1 \leq k \leq C(n)} \iint_{B_{j,k}} |\nabla u(X)|^2 \, dm(X) \\ &\lesssim \sum_{\substack{2^{j-1}\tau r < tr \\ 1 \leq k \leq C(n)}} (2^j \tau r)^{-1-d} \iint_{B^*_{j,k}} |u(X)|^2 \, dm(X) \\ &\lesssim (\delta\lambda)^2 \sum_{2^{j-1}\tau r < tr} (2^j \tau r)^{-1-d} m(B^*_{j,k}) \\ &\lesssim (\delta\lambda)^2 \log_2\left(\frac{t}{\tau}\right) < 2\lambda^2, \end{aligned} \tag{3.65}$$

if δ is sufficiently small depending on the values of t, τ and α, β. Therefore (3.62) holds, and thus for any $q \in E$,

$$\begin{aligned} |S_{\tau r} u(q)|^2 &= \iint_{\Gamma^\alpha(q) \cap B(q, \tau r)} |\nabla u(X)|^2 \, \delta(X)^{1-d} \, dm(X) \\ &= \iint_{\Gamma^\alpha(q) \setminus U} |\nabla u(X)|^2 \delta(X)^{1-d} \, dm(X) \\ &> \lambda^2. \end{aligned} \tag{3.66}$$

<u>Step 2</u>: Combining (3.66) with $E \subset F$ we get

$$\lambda^2 \omega(E) \le \int_E |S_{\tau r} u(q)|^2 \, d\omega(q) \le \int_F \iint_{\Gamma^\alpha_{\tau r}(q)} |\nabla u(X)|^2 \, \delta(X)^{1-d} \, dm(X) d\omega(q). \tag{3.67}$$

By (3.21) we have

$$\Gamma^\alpha_{\tau r}(q) \subset \operatorname{int} \Gamma_d(q) \subset \Gamma_d(q) \tag{3.68}$$

for any $q \in Q$. In particular if X belongs to the left-hand side of (3.68), then $X \in U_{Q'}$ for some dyadic cube Q' containing q. Moreover

$$\delta(X) \le |X - q| < \tau r = \tau \operatorname{diam} Q \sim \tau \ell(Q). \tag{3.69}$$

By the definition of $U_{Q'}$ and (3.10), we have

$$\delta(X) \gtrsim c \eta^{\frac{n-1}{4(n-1-d)}} \ell(Q'). \tag{3.70}$$

By combining (3.69), (3.70) and choosing τ small enough depending on η, we can guarantee that $\ell(Q') < 2\ell(Q)$. Since $Q' \cap Q \ni q$, by property (ii) of Lemma 3.3 we know $Q' \in \mathbb{D}_Q$. Hence $\Gamma^\alpha_{\tau r}(q) \subset \Gamma^Q_d(q)$. Again since $\Gamma^\alpha_{\tau r}(q)$ is an open set, we also have $\Gamma^\alpha_{\tau r}(q) \subset \operatorname{int} \Gamma^Q_d(q)$. Therefore

$$\bigcup_{q \in F} \Gamma^\alpha_{\tau r}(q) \subset \bigcup_{q \in F} (\operatorname{int} \Gamma^Q_d(q)) \subset \operatorname{int}\Bigl(\bigcup_{q \in F} \Gamma^Q_d(q) \Bigr). \tag{3.71}$$

Applying Fubini's theorem to the right-hand side of (3.67), we conclude that it is bounded by

$$\iint_{\operatorname{int}(\bigcup_{p \in F} \Gamma^Q_d(p))} |\nabla u(X)|^2 \, \delta(X)^{1-d} \omega\bigl(\{q \in F : X \in \Gamma^Q_d(q)\}\bigr) \, dm(X). \tag{3.72}$$

For any $p \in F$ and any $X \in \Gamma^Q_d(p)$, we have $X \in I \in \mathcal{W}_{Q'}$ for a cube Q' in $\mathbb{D}_{\mathcal{F}_1, Q}$ containing p. Thus $|X - q| \sim \ell(Q') \sim \ell(I) \sim \delta(X)$. Since the family $\{I^*\}_{I \in \mathcal{W}}$ has bounded overlap and harmonic measure, ω has pole at X_Q, and we conclude by Lemma 2.82 that

$$\omega(\{q \in F : X \in \Gamma^Q_d(q)\}) \sim \omega\Bigl(\bigcup_{\substack{Q' \in \mathbb{D}_Q \\ \ell(Q') \sim \delta(X) \sim \operatorname{dist}(X, Q')}} Q' \Bigr) \sim G(X_Q, X) \delta(X)^{d-1}. \tag{3.73}$$

Combining (3.67), (3.72), (3.73) and (3.61) and using (1.3), we get

$$\begin{aligned}\lambda^2\omega(E) &\lesssim \iint_{\Omega_{\mathcal{F}_2,Q}} |\nabla u(X)|^2 G(X_Q,X)\,dm(X)\\ &= \iint_{\Omega_{\mathcal{F}_2,Q}} |\nabla u(X)|^2 G(X_Q,X) w(X)\,dX \lesssim \iint_{\Omega_{\mathcal{F}_2,Q}} A\nabla u\cdot\nabla u G\,dX. \qquad (3.74)\end{aligned}$$

Here we abbreviate $G(X)=G(X_Q,X)$ when there is no ambiguity as to what the pole is. Recall that $X_Q\notin B(x_Q,2C_3\ell(Q))$, and similar to (3.17) we may choose the dilation constant θ small enough so that $\overline{\Omega}^{***}_{\mathcal{F}_2,Q}\subset B\big(x_Q,\tfrac{3}{2}C_3\ell(Q)\big)$. They guarantee that $X_Q\notin\overline{\Omega}^{***}_{\mathcal{F}_2,Q}$, and moreover $\operatorname{dist}(X_Q,\overline{\Omega}^{***}_{\mathcal{F}_2,Q})\gtrsim\ell(Q)$. Hence $G(X)$ is harmonic in the fat sawtooth domain $\Omega^{***}_{\mathcal{F}_2,Q}$.

Step 3: Next we are going to prove

$$\iint_{\Omega_{\mathcal{F}_2,Q}} A\nabla u\cdot\nabla u G\,dX \lesssim (\delta\lambda)^2\omega(Q). \qquad (3.75)$$

Recalling the discussion before Lemma 3.24, we can augment $\mathcal{F}_2$ by adding all dyadic cubes of lengths less than or equal to $2^{-N}\ell(Q)$, and denote by $\mathcal{F}_2^N$ the collection of maximal cubes giving rise to the aforementioned augmented collection. We claim that

$$\iint_{\Omega_{\mathcal{F}_2^N,Q}} A\nabla u\cdot\nabla u G\,dX \lesssim (\delta\lambda)^2\omega(Q), \qquad (3.76)$$

with a constant independent of N. Thus by passing $N\to\infty$ we obtain (3.75).

Recall that in Lemma 3.24, we constructed a smooth cut-off function ψ_N such that $\chi_{\Omega^*_{\mathcal{F}_N,Q}}\lesssim\psi_N\le\chi_{\Omega^{**}_{\mathcal{F}_N,Q}}$. Hence

$$\iint_{\Omega_{\mathcal{F}_N,Q}} A\nabla u\cdot\nabla u G\,dX \le \iint_{\mathbb{R}^n} A\nabla u\cdot\nabla u G\psi_N\,dX. \qquad (3.77)$$

Since $u,G\in W_r(\Omega^{**}_{\mathcal{F}_N,Q})\cap L^\infty(\Omega^{**}_{\mathcal{F}_N,Q})$, we have $uG\psi_N, u^2\psi_N\in W_0^{1,2}(\Omega^{**}_{\mathcal{F}_N,Q})$. In particular they can be approximated by smooth functions in $C_0^\infty(\Omega^{**}_{\mathcal{F}_N,Q})\subset C_0^\infty(\Omega)$. In the sawtooth region $\Omega^{**}_{\mathcal{F}_N,Q}$ we have $-\operatorname{div}(A\nabla u)=-\operatorname{div}(A\nabla G)=0$; thus

$$\begin{aligned}\iint_{\mathbb{R}^n} A\nabla u\cdot\nabla u G\psi_N\,dX &= \iint_{\mathbb{R}^n} A\nabla u\cdot\nabla(uG\psi_N)-\frac{1}{2}A\nabla(u^2)\cdot\nabla(G\psi_N)\,dX\\ &= 0-\frac{1}{2}\iint_{\mathbb{R}^n} A\nabla(G\psi_N)\cdot\nabla(u^2)\,dX\\ &= -\frac{1}{2}\bigg(\iint_{\mathbb{R}^n}\psi_N A\nabla G\cdot\nabla(u^2)+GA\nabla\psi_N\cdot\nabla(u^2)\,dX\bigg)\\ &= -\frac{1}{2}\bigg(\iint_{\mathbb{R}^n} A\nabla G\cdot\nabla(u^2\psi_N)-u^2A\nabla G\cdot\nabla\psi_N+2uGA\nabla u\cdot\nabla\psi_N\,dX\bigg)\\ &= \frac{1}{2}\iint_{\mathbb{R}^n} u^2A\nabla G\cdot\nabla\psi_N\,dX-\iint_{\mathbb{R}^n} uGA\nabla u\cdot\nabla\psi_N\,dX =: \frac{1}{2}I-II, \qquad (3.78)\end{aligned}$$

where we use the symmetry of A and the equation $-\operatorname{div}(A\nabla u) = 0$ in the second equality, and $-\operatorname{div}(A\nabla G) = 0$ in the second-to-last equality. We first estimate the second term. By (3.25), the contribution to the integral II only comes from Whitney boxes $I \in \mathcal{W}_N^\Sigma$. Recall the harmonic function u is nonnegative and we use X_I to denote the center of Whitney box I. By Lemma 3.24(ii), the Hölder inequality, the estimate of the weight (2.8), the interior Caccioppoli inequality (2.23), the Harnack inequality (2.26) and (2.83), we have

$$\begin{aligned}
|II| &\le \sum_{I\in\mathcal{W}_N^\Sigma} \frac{u(X_I)G(X_I)}{\ell(I)} \iint_{I^{***}} |\nabla u|\,dm \\
&\le \sum_{I\in\mathcal{W}_N^\Sigma} \frac{u(X_I)G(X_I)}{\ell(I)} \cdot m(I^{***}) \bigg(\fint\!\!\fint_{I^{***}} |\nabla u|^2\,dm \bigg)^{\frac12} \\
&\lesssim \sum_{I\in\mathcal{W}_N^\Sigma} u(X_I)G(X_I)\ell(I)^{d-1} \bigg(\fint\!\!\fint_{I^{****}} |u|^2\,dm \bigg)^{\frac12} \\
&\lesssim \sum_{I\in\mathcal{W}_N^\Sigma} u(X_I)^2 G(X_I)\ell(I)^{d-1} \sim \sum_{I\in\mathcal{W}_N^\Sigma} u(X_I)^2\omega(Q_I),
\end{aligned} \tag{3.79}$$

where Q_I is defined as in Lemma 3.24(iv). Using the estimate $|u(X)| \le \delta\lambda$ for all $X \in \Omega^{***}_{\mathcal{F}_N,Q}$ and (3.26), we have

$$|II| \lesssim \sum_{I\in\mathcal{W}_N^\Sigma} u(X_I)^2\omega(Q_I) \lesssim (\delta\lambda)^2\omega(Q). \tag{3.80}$$

Similarly,

$$|I| \le \sum_{I\in\mathcal{W}_N^\Sigma} \frac{u(X_I)^2}{\ell(I)} \iint_{I^{***}} |\nabla G|\,dm \lesssim \sum_{I\in\mathcal{W}_N^\Sigma} u(X_I)^2\omega(Q_I) \lesssim (\delta\lambda)^2\omega(Q). \tag{3.81}$$

We finish the proof of (3.75) by combining (3.78), (3.80) and (3.81).

Finally we combine (3.67) and (3.75), and get

$$\lambda^2\omega(E) \lesssim (\delta\lambda)^2\omega(Q), \tag{3.82}$$

and thus

$$\omega(E) \le C\delta^2\omega(Q). \tag{3.83}$$

This finishes the proof of the good-λ inequality for ω. □

We will also need the following auxiliary fact:

Lemma 3.84. *For any apertures $0 < \alpha < \alpha'$ and any function $u \in W_r(\Omega)$, let Su and $\widetilde{S}u$ denote the square functions with apertures α and α' respectively. Suppose $\widetilde{S}u < \infty$ for σ-almost every $q \in \Gamma$; then the set $\{q \in \Gamma : Su(q) > \lambda\}$ is open for every $\lambda > 0$.*

The proof is similar in spirit to that of Lemma 4.6 in [Milakis et al. 2013].

Proof. If $q \in \Gamma$ is such that $S'u(q) > \lambda$, then there exists $\eta > 0$ so that

$$\iint_{\Gamma^\alpha(q)\setminus B(q,\eta)} |\nabla u|^2\, \delta(X)^{1-d}\, dm(X) > \left(\frac{Su(q)+\lambda}{2}\right)^2.$$

We claim that there exists $\epsilon > 0$ such that for any $p \in \Delta(q, \epsilon\eta)$ we have

$$\iint_{\Gamma^\alpha(p)\setminus B(p,\eta)} |\nabla u|^2 \delta(X)^{1-d}\, dm(X) > \lambda^2, \tag{3.85}$$

and therefore $Su(p) > \lambda$.

We observe that

$$\left|\iint_{\Gamma^\alpha(q)\setminus B(q,\eta)} |\nabla u|^2\, \delta(X)^{1-d}\, dm(X) - \iint_{\Gamma^\alpha(p)\setminus B(p,\eta)} |\nabla u|^2\, \delta(X)^{1-d}\, dm(X)\right| \le \iint_D |\nabla u|^2\, \delta(X)^{1-d}\, dm(X), \tag{3.86}$$

where $D = (\Gamma^\alpha(q)\setminus B(q,\eta))\triangle(\Gamma^\alpha(p)\setminus B(p,\eta))$ is the set difference. It suffices to show that the integral $\iint_D |\nabla u|^2\, \delta(X)^{1-d}\, dm(X)$ is sufficiently small, if we choose ϵ sufficiently small.

Suppose that $X \in \Gamma^\alpha(q)\setminus B(q,\eta)$; then $|X-q| < (1+\alpha)\delta(X)$ and $|X-q| \ge \eta$. Thus $\delta(X) > \eta/(1+\alpha)$. If moreover $X \notin \Gamma^\alpha(p)\setminus B(p,\eta)$ and $p \in B(q,\epsilon\eta)$, then $|X-q| > (1+\alpha)(1-\epsilon)\delta(X)$. By symmetry, we need to study sets of the form

$$V_q = \big\{X \in \Omega : |X-q| \ge \eta,\ (1+\alpha)(1-\epsilon)\delta(X) < |X-q| < (1+\alpha)\delta(X)\big\},$$
$$V_p = \big\{X \in \Omega : |X-p| \ge \eta,\ (1+\alpha)(1-\epsilon)\delta(X) < |X-p| < (1+\alpha)\delta(X)\big\}.$$

Without loss of generality we may assume $S'u(q) < \infty$. If not, by the assumption that $S'u < \infty$ almost everywhere, we can always find $q' \in \Delta\big(q, \tfrac{1}{2}\epsilon\eta\big)$ such that $S'u(q') < \infty$, and in particular $p \in \Delta(q,\epsilon\eta) \subset \Delta(q', 2\epsilon\eta)$. In this case we just replace q by q', and ϵ by 2ϵ. Moreover, if $\epsilon < \frac{1}{4}$, we have

$$V_q \cup V_p \subset V_\epsilon := \left\{X \in \Omega : |X-q| \ge \frac{\eta}{2},\ (1+\alpha)(1-\epsilon)^2\, \delta(X) < |X-q| < (1+\alpha)\frac{1-\epsilon}{1-2\epsilon}\delta(X)\right\}.$$

Note that for given $\alpha' > \alpha$, by choosing ϵ sufficiently small we can guarantee $(1+\alpha)(1-\epsilon)/(1-2\epsilon) \le 1+\alpha'$. Thus $V_\epsilon \subset \Gamma^{\alpha'}(q)\setminus B\big(q, \tfrac{1}{2}\eta\big) =: V_0$, and as ϵ tends to zero, the set V_ϵ decreases to an empty set. Moreover,

$$\iint_{V_0} |\nabla u|^2\, \delta(X)^{1-d}\, dm(X) \le \iint_{\Gamma^{\alpha'}(q)} |\nabla u|^2\, \delta(X)^{1-d}\, dm(X) = |S'u(q)|^2 < \infty;$$

hence by the continuity of measure from above, we deduce that

$$\iint_{V_\epsilon} |\nabla u|^2\, \delta(X)^{1-d}\, dm(X) \searrow 0.$$

In particular, by choosing ϵ sufficiently small, we can guarantee

$$\iint_D |\nabla u|^2\, \delta(X)^{1-d}\, dm(X) \le \iint_{V_\epsilon} |\nabla u|^2\, \delta(X)^{1-d}\, dm(X) < \left(\frac{Su(q)+\lambda}{2}\right)^2 - \lambda^2 \tag{3.87}$$

Combining (3.87) with (3.86), we conclude the proof of the claim (3.85). □

Now we set out to complete the proof of Theorem 3.1.

Proof of Theorem 3.1. We first prove the theorem assuming that $\|S'u\|_{L^p(\sigma)}$ is finite. Under this assumption, we have $\|S''u\|_{L^p(\sigma)} \sim \|S'u\|_{L^p(\sigma)} \sim \|Su\|_{L^p(\sigma)}$. For reference, see Proposition 4 of [Coifman et al. 1985]. (The stated proof in that paper is for the upper half-plane, but the argument goes through for Ahlfors regular sets of higher codimension.) Therefore by a standard argument, the proof of (3.2) can be reduced to the following good-λ inequality: for any $\epsilon > 0$ sufficiently small, we can find $\delta = \delta(\epsilon) > 0$ such that for all $\lambda > 0$,

$$\sigma\big(\{q \in \Gamma : Su(q) > 2\lambda,\ Nu(q) \le \delta\lambda\}\big) \le \epsilon\sigma\big(\{q \in \Gamma : S'u(q) > \lambda\}\big), \tag{3.88}$$

and $\delta \to 0$ as $\epsilon \to 0$. If $\{q \in \Gamma : S'u(q) > \lambda\}$ is empty, (3.88) is trivial, so we assume the set is not empty. We apply Lemma 3.84 with apertures $0 < \alpha_1 < \beta$. Since $\|S''u\|_{L^p(\sigma)} \sim \|S'u\|_{L^p(\sigma)} < \infty$, in particular $S''u(q) < \infty$ almost everywhere. Therefore $\{q \in \Gamma : S'u(q) > \lambda\}$ is open. We also remark that the set $\{q \in \Gamma : S'u(q) > \lambda\}$ has finite σ-measure, and moreover

$$\sigma\big(\{q \in \Gamma : S'u(q) > \lambda\}\big) \le \frac{1}{\lambda^p}\int_{S'u(q)>\lambda} |S'u|^p \, d\sigma \le \frac{\|S'u\|^p_{L^p(\sigma)}}{\lambda^p} < \infty. \tag{3.89}$$

In particular, for any dyadic cube $Q \in \mathbb{D}$ completely contained in $\{q \in \Gamma : S'u(q) > \lambda\}$

$$\ell(Q)^d \sim \sigma(Q) \le \sigma\big(\{q \in \Gamma : S'u(q) > \lambda\}\big) \le \frac{\|S'u\|_{L^p(\sigma)}}{\lambda^p}, \tag{3.90}$$

so its length has a uniform upper bound (albeit depending on the value of λ). Recall that $\ell(Q) \sim 2^{-k(Q)}$, and suppose $k_0 \in \mathbb{Z}$ is such that

$$2^{-k_0 d} \gtrsim \frac{\|S'u\|_{L^p(\sigma)}}{\lambda^p}, \tag{3.91}$$

with a sufficiently large implicit constant. Then by (3.90), any cube Q_0 in $\mathbb{D}_{k_0}$ cannot be completely contained in $\{q \in \Gamma : S'u(q) > \lambda\}$.

We run a stopping-time procedure as follows: for each $Q_0 \in \mathbb{D}_{k_0}$, we traverse all its descendants, and stop whenever we find a cube $Q \in \mathbb{D}_{Q_0}$ such that $S'u(q) > \lambda$ for all $q \in Q$. Let $\mathcal{F}_1 = \{Q_l\}$ be the collection of all stopping cubes in $\bigcup_{Q_0 \in \mathbb{D}_{k_0}} \mathbb{D}_{Q_0}$. Similar to the proof of (3.60), we can show that they form a partition:

$$\{q \in \Gamma : S'u(q) > \lambda\} = \bigcup_{Q_l \in \mathcal{F}_1} Q_l. \tag{3.92}$$

Note that the assumption $Su(q) > 2\lambda$ clearly implies $S'u(q) > \lambda$; namely

$$\{q \in \Gamma : Su(q) > 2\lambda\} \subset \{q \in \Gamma : S'u(q) > \lambda\} = \bigcup_{Q_l \in \mathcal{F}_1} Q_l.$$

Therefore to prove (3.88), it suffices to localize and show that

$$\sigma\big(\{q \in Q : Su(q) > 2\lambda, Nu(q) \le \delta\lambda\}\big) \le \epsilon\sigma(Q) \quad \text{for any } Q = Q_l \in \mathcal{F}_1. \tag{3.93}$$

Recall that by (3.5), every $Q \in \mathbb{D}$ is contained in a surface ball $\Delta(x_Q, C_2 r_Q)$. Let X'_Q denote a corkscrew point for $B(x_Q, C_2 r_Q)$. Recall Definition 2.89 of $\omega \in A_\infty(\sigma)$ and Remark 2.91(ii) right afterwards. Assuming $\omega \in A_\infty(\sigma)$, to prove (3.93) it suffices to show that

$$\omega^{X'_Q}\big(\{q \in Q : Su(q) > 2\lambda, Nu(q) \leq \delta\lambda\}\big) \leq C(\delta)\omega^{X'_Q}(Q), \tag{3.94}$$

with a constant $C(\delta)$ independent of Q and λ, and that $C(\delta) \to 0$ as $\delta \to 0$. Recall that for any collection $\mathcal{F}$ of dyadic cubes, there is a constant C_3 such that $\Omega^{***}_{\mathcal{F},Q} \subset B(x_Q, C_3\ell(Q)) \cap \Omega$. Let X_Q be a corkscrew point for $B(x_Q, 2C_3 M\ell(Q))$; then

$$|X_Q - x_Q| \geq \delta(X_Q) \geq 2C_3\ell(Q). \tag{3.95}$$

Thus $X_Q \notin B(x_Q, 2C_3\ell(Q))$, and in particular $X_Q \notin \overline{\Omega}^{***}_{\mathcal{F},Q}$. Moreover, there is a Harnack chain of finite length (depending only on M, C_2 and C_3) connecting X_Q to X'_Q; in particular $\omega^{X_Q}(E) \sim \omega^{X'_Q}(E)$ for any Borel set $E \subset Q$. Therefore the proof of (3.94) is equivalent to the proof of

$$\omega^{X_Q}\big(\{q \in Q : Su(q) > 2\lambda, Nu(q) \leq \delta\lambda\}\big) \leq C(\delta)\omega^{X_Q}(Q). \tag{3.96}$$

Recall that $Q = Q_l \in \mathcal{F}_1$ is a maximal cube with respect to the stopping criterion $\{S'u(q) > \lambda\}$. By maximality the parent of Q, denoted by $\widetilde{Q}$, contains at least one point $q_1 \notin \{q \in \Gamma : S'u(q) > \lambda\}$; that is, $S'u(q_1) \leq \lambda$. For any $q \in Q$ we have

$$|q_1 - q| \leq \operatorname{diam} \widetilde{Q} \leq A_1 2^{-k(\widetilde{Q})} = A_1 2^{-(k(Q)-1)} \leq \frac{A_1}{a_0} \operatorname{diam} Q. \tag{3.97}$$

Therefore for any maximal cube, we may use Proposition 1.16, with constant $C_2 = A_1/a_0$, to conclude the desired estimate (3.96).

All the above arguments show that if we know a priori $\|S'u\|_{L^p(\sigma)}$ is finite, we can prove $\|Su\|_{L^p(\sigma)} \lesssim \|Nu\|_{L^p(\sigma)}$. If we do not have this a priori information, then for κ sufficiently small we let

$$\mathbb{D}_\kappa = \{Q \in \mathbb{D} : \kappa \leq \ell(Q) \leq 1/\kappa\}, \tag{3.98}$$

$$\Omega_\kappa = \bigcup_{Q \in \mathbb{D}_\kappa} U_Q, \quad \Omega^*_\kappa = \bigcup_{Q \in \mathbb{D}_\kappa} U^*_Q, \quad \Omega^{**}_\kappa = \bigcup_{Q \in \mathbb{D}_\kappa} U^{**}_Q \quad \text{etc.}, \tag{3.99}$$

define the κ-approximate nontangential cones as

$$\Gamma^\alpha_\kappa(q) = \Gamma^\alpha(q) \cap \Omega_\kappa, \quad \Gamma^{\alpha_1}_\kappa = \Gamma^{\alpha_1}(q) \cap \Omega_\kappa, \quad \Gamma^\beta_\kappa(q) = \Gamma^\beta(q) \cap \Omega^{***}_\kappa,$$

and define the κ-approximate *dyadic* nontangential cones as

$$\Gamma_{d,\kappa}(q) = \Gamma_d(q) \cap \Omega_\kappa = \bigcup_{Q \in \mathbb{D}^\kappa : Q \ni q} U_Q, \quad \widehat{\Gamma}_{d,\kappa}(q) = \widehat{\Gamma}_d(q) \cap \Omega^{***}_\kappa.$$

In this regime we have the following inclusions analogous to (3.21) and (3.22):

$$\Gamma^\alpha_\kappa(q) \subset \Gamma_{d,\kappa}(q) \subset \Gamma^{\alpha_1}_\kappa(q), \quad \widehat{\Gamma}_{d,\kappa}(q) \subset \Gamma^\beta_\kappa(q). \tag{3.100}$$

Moreover, the κ-approximate local nontangential cones

$$\Gamma^Q_{d,\kappa}(q)=\Gamma^Q_d(q)\cap\Omega_\kappa=\bigcup_{Q'\in\mathbb{D}_Q\cap\mathbb{D}^\kappa:Q'\ni q}U_{Q'},\quad \widehat{\Gamma}^Q_{d,\kappa}(q)=\widehat{\Gamma}^Q_d\cap\Omega^{***}_\kappa$$

satisfy the following inclusions analogous to (3.23):

$$\bigcup_{q\in F}\Gamma^Q_{d,\kappa}(q)\subset\Omega_{\mathcal{F},Q}\cap\Omega_\kappa\subset\overline{\Omega_{\mathcal{F},Q}\cap\Omega_\kappa}\subset\Omega^{***}_{\mathcal{F},Q}\cap\Omega^{***}_\kappa\subset\bigcup_{q\in F}\widehat{\Gamma}^Q_{d,\kappa}(q)$$

for any dyadic cube Q and collection of maximal cubes $\Gamma\subset\mathbb{D}_Q\setminus\{Q\}$, under the assumption that $F=Q\setminus\bigcup_{Q_j\in\mathcal{F}}Q_j$ is not empty. We then define the κ-approximate square functions $S_\kappa u, S'_\kappa u$ and nontangential maximal function $N_\kappa u$ accordingly, as integrals defined on the κ-approximate nontangential cones instead of standard nontangential cones. Since $N_\kappa u(q)\le Nu(q)$ for all $q\in\Gamma$, we have $\|N_\kappa u\|_{L^p(\sigma)}\le\|Nu\|_{L^p(\sigma)}<\infty$. By the interior Caccioppoli inequality (2.23) and $\beta>\alpha_1>\alpha$, we have

$$S_\kappa u(q)\le S'_\kappa u(q)\lesssim C(\kappa)N_\kappa u(q),$$

and thus

$$\|S'_\kappa u\|_{L^p(\sigma)}\lesssim C(\kappa)\|N_\kappa u\|_{L^p(\sigma)}\le C(\kappa)\|Nu\|_{L^p(\sigma)}<\infty. \tag{3.101}$$

We cannot let κ go to zero in (3.101) since the upper bound on the right-hand side depends on κ (in fact $C(\kappa)\to\infty$ as $\kappa\to 0$). However, since $\|S'_\kappa u\|_{L^p(\sigma)}$ is finite, we can apply the previous arguments and prove that $\|S_\kappa u\|_{L^p(\sigma)}\lesssim\|N_\kappa u\|_{L^p(\sigma)}$, with a constant independent of κ. Hence

$$\|S_\kappa u\|_{L^p(\sigma)}\lesssim\|N_\kappa u\|_{L^p(\sigma)}\le C\|Nu\|_{L^p(\sigma)},$$

with a constant C independent of κ. Therefore we can safely let κ go to zero and conclude that

$$\|Su\|_{L^p(\sigma)}=\limsup_{\kappa\to 0}\|S_\kappa u\|_{L^p(\sigma)}\le C\|Nu\|_{L^p(\sigma)}.$$

This finishes the proof of Theorem 3.1. □

4. $\omega\in A_\infty(\sigma)$ is equivalent to BMO-solvability

4A. *From $\omega\in A_\infty(\sigma)$ to L^p-solvability.*

Theorem 4.1. *Assume $\omega\in A_\infty(\sigma)$. Then there exist some $p_0\in(1,\infty)$ such that the elliptic problem* (D) *is L^p-solvable for all $p\in(p_0,\infty)$, in the sense that there exists a universal constant $C>0$ such that for any $f\in C^0_0(\Gamma)$ and any Borel set $E\subset\Gamma$, the solution $u(X)=\int_E f\,d\omega^X$ satisfies the estimate $\|Nu\|_{L^p(\sigma)}\le C\|f\chi_E\|_{L^p(\sigma)}$.*

Remark 4.2. For a bounded set E, it suffices to assume that $f\in C_b(\Gamma)$.

Proof. We first treat the case when $E=\Gamma$. Let $q\in\Gamma$ and define for any $p>1$

$$\mathcal{M}_p f(q)=\sup_{\Delta\ni q}\left(⨍_\Delta|f|^p\,d\sigma\right)^{\frac{1}{p}}<\infty. \tag{4.3}$$

We claim

$$|u(X)| \le C\mathcal{M}_p f(q) \quad \text{for any } X \in \Gamma(q). \tag{4.4}$$

Hence $Nu(q) \le C\mathcal{M}_p f(q)$, and thus by the L^p-boundedness ($p > 1$) of Hardy–Littlewood maximal function (see [Coifman and Weiss 1977] for spaces of homogeneous type and [Stein 1993])

$$\|Nu\|_{L^p(\sigma)} \le C\|\mathcal{M} f\|_{L^p(\sigma)} \lesssim \|f\|_{L^p(\sigma)}.$$

In fact, let $X \in \Gamma(q)$ be fixed and $\Delta = \Delta(q, (1+\alpha)\delta(X))$. For $j \in \mathbb{N}$ let $\Delta_j = 2^j \Delta$, and set $\Delta_{-1} = \varnothing$. We have

$$u(X) = \int f\, d\omega^X = \sum_{j=0}^{\infty} \int_{\Delta_j \setminus \Delta_{j-1}} f\, d\omega^X. \tag{4.5}$$

For each $j \in \mathbb{N}$, let A_j denote a corkscrew point for Δ_j. Recall Definition 2.89 of $\omega \in A_\infty(\sigma)$ and the discussion after that, in particular (2.92) and (2.93). We have that for each j, the Radon–Nikodym derivative

$$k^{A_j}(q') = \frac{d\omega^{A_j}}{d\sigma}(q') = \lim_{\Delta' \to q'} \frac{\omega^{A_j}(\Delta')}{\sigma(\Delta')}$$

satisfies a reverse Hölder inequality

$$\left(\fint_{\Delta_j} |k^{A_j}|^r\, d\sigma \right)^{\frac{1}{r}} \le C \fint_{\Delta_j} k^{A_j}\, d\sigma \tag{4.6}$$

for all $r \in (1, r_0)$, with uniform constants $r_0 > 1$ and $C > 0$. For any $j \ge 2$ and any surface ball $\Delta' \subset \Delta_j \setminus \Delta_{j-1}$, by the Hölder regularity of solutions near the boundary (see Lemma 2.33), we have

$$\omega^X(\Delta') \lesssim 2^{-j\beta} \omega^{A_{j-2}}(\Delta') \sim 2^{-j\beta} \omega^{A_j}(\Delta'). \tag{4.7}$$

Hence for any $q' \in \Delta_j \setminus \Delta_{j-1}$,

$$k^X(q') = \lim_{\Delta' \to q'} \frac{\omega^X(\Delta')}{\sigma(\Delta')} = \lim_{\substack{\Delta' \to q' \\ \Delta' \subset \Delta_j \setminus \Delta_{j-1}}} \frac{\omega^X(\Delta')}{\sigma(\Delta')} \lesssim 2^{-j\beta} \lim_{\substack{\Delta' \ni q' \\ \Delta' \subset \Delta_j \setminus \Delta_{j-1}}} \frac{\omega^{A_j}(\Delta')}{\sigma(\Delta')} = 2^{-j\beta} k^{A_j}(q'). \tag{4.8}$$

Therefore by (4.6), (4.8), and Hölder inequality for conjugates $\frac{1}{p} + \frac{1}{r} = 1$ with $r \in (1, r_0)$, we obtain

$$\begin{aligned}
|u(X)| &\le \sum_{j=0}^{\infty} \int_{\Delta_j \setminus \Delta_{j-1}} |f k^X|\, d\sigma \lesssim \sum_{j=0}^{\infty} 2^{-j\beta} \int_{\Delta_j} |f| k^{A_j}\, d\sigma \\
&\le \sum_{j=0}^{\infty} 2^{-j\beta} \sigma(\Delta_j) \left(\fint_{\Delta_j} |f|^p\, d\sigma \right)^{\frac{1}{p}} \left(\fint_{\Delta_j} |k^{A_j}|^r\, d\sigma \right)^{\frac{1}{r}} \\
&\lesssim \sum_{j=0}^{\infty} 2^{-j\beta} \sigma(\Delta_j) \left(\fint_{\Delta_j} |f|^p\, d\sigma \right)^{\frac{1}{p}} \left(\fint_{\Delta_j} k^{A_j}\, d\sigma \right) \\
&\le \sum_{j=0}^{\infty} 2^{-j\beta} \mathcal{M}_p f(q) \omega^{A_j}(\Delta_j) \lesssim \mathcal{M}_p f(q).
\end{aligned} \tag{4.9}$$

Thus we have finished proving the claim (4.4) for any $p \in (p_0, \infty)$, where p_0 is the conjugate of r_0. Note that we never used the continuity or compact support of f, and replacing f by $f\chi_E$ we can repeat the same argument with no change. The assumption that E is bounded or f has compact support guarantees we still have a priori finite integrability in (4.3). $\square$

4B. ***Proof of the BMO-solvability.***

Theorem 4.10. *Assume that $\omega \in A_\infty(\sigma)$. For any $f \in C_0^0(\Gamma)$, let $u = Uf \in W_r(\Omega)$ be a solution to $Lu = 0$ given by Lemmas 2.37 and 2.39. Then $|\nabla u|^2\,\delta(X)\,dm(X)$ is a Carleson measure, and moreover*

$$\sup_{\Delta\subset\Gamma} \frac{1}{\sigma(\Delta)} \iint_{T(\Delta)} |\nabla u|^2\,\delta(X)\,dm(X) \le C\|f\|^2_{\mathrm{BMO}(\sigma)}. \tag{4.11}$$

Proof. Fix an arbitrary surface ball $\Delta = \Delta(q_0, r)$. Let $\alpha > 0$. Define the constant $c = \max\{\alpha+2, 12\}$ and let $\tilde\Delta = c\Delta = \Delta(q_0, cr)$ be a concentric dilation. We define the average $f_{\tilde\Delta} = \fint_{\tilde\Delta} f\,d\sigma$. Let

$$f_1 = (f - f_{\tilde\Delta})\chi_{\tilde\Delta}, \quad f_2 = (f - f_{\tilde\Delta})\chi_{\Gamma\setminus\tilde\Delta}, \quad f_3 = f_{\tilde\Delta},$$

and for any $X \in \Omega$ let

$$\begin{aligned} u_1(X) &= \int_\Gamma f_1\,d\omega^X = \int_{\tilde\Delta} (f - f_{\tilde\Delta})\,d\omega^X, \\ u_2(X) &= \int_\Gamma f_2\,d\omega^X = \int_{\Gamma\setminus\tilde\Delta} (f - f_{\tilde\Delta})\,d\omega^X = \int_{\Gamma\setminus\tilde\Delta} f\,d\omega^X - f_{\tilde\Delta}\omega^X(\Gamma\setminus\tilde\Delta), \\ u_3 &\equiv f_{\tilde\Delta}. \end{aligned}$$

By Lemmas 2.37, 2.39, 2.42 and 2.43, they are solutions to L, and u_1, u_2 can be continuously extended to $\Gamma\setminus\tilde\Delta$ and $\tilde\Delta$, respectively. Moreover

$$(u_1 + u_2 + u_3)(X) = \int_\Gamma f\,d\omega^X = Uf(X) = u(X).$$

Clearly the Carleson measure of the constant function u_3 is trivial.

Applying Theorem 4.1 to f_1 and u_1, we get $\|Nu_1\|_{L^p(\sigma)} \le C\|f_1\|_{L^p(\sigma)} < \infty$. Combined with Theorem 3.1,

$$\|Su_1\|_{L^p(\sigma)} \lesssim \|Nu_1\|_{L^p(\sigma)} \lesssim \|f_1\|_{L^p(\sigma)} = \left(\int_{\tilde\Delta} |f - f_{\tilde\Delta}|^p\,d\sigma\right)^{\frac{1}{p}} \tag{4.12}$$

for any $p \in (p_0, \infty)$. By (2.95) and (2.96)

$$\iint_{T(\Delta)} |\nabla u_1|^2\,\delta(X)\,dm(X) \le C\int_{(\alpha+2)\Delta} |S_{(\alpha+1)r}u_1|^2\,d\sigma.$$

Recall that $\tilde\Delta = c\Delta \supset (\alpha+2)\Delta$; thus

$$\begin{aligned} \iint_{T(\Delta)} |\nabla u_1|^2\,\delta(X)\,dm(X) &\le C\int_{\tilde\Delta} |S_{(\alpha+1)r}u_1|^2\,d\sigma \\ &\le C\sigma(\tilde\Delta)^{1-\frac{2}{p}}\left(\int_{\tilde\Delta} |Su_1|^p\,d\sigma\right)^{\frac{2}{p}} \\ &\le C\sigma(\tilde\Delta)^{1-\frac{2}{p}}\|Su_1\|^2_{L^p(\sigma)} \end{aligned} \tag{4.13}$$

for any $p > \max\{2, p_0\}$. Combining (4.13) and (4.12) we get

$$\iint_{T(\Delta)} |\nabla u_1|^2 \, \delta(X) \, dm(X) \le C\sigma(\Delta) \|f\|^2_{\mathrm{BMO}(\sigma)} < \infty. \tag{4.14}$$

Turning to the estimate for u_2, let $\{I_k\} \subset \mathcal{W}$ be a collection of dyadic Whitney boxes that intersect $T(\Delta)$ (recall the properties of Whitney decomposition $\mathcal{W}$ in (3.6)). On each Whitney box I_k, we have by the interior Caccioppoli inequality (2.23)

$$\begin{aligned} \iint_{I_k} |\nabla u_2|^2 \, \delta(X) \, dm(X) &\lesssim \ell(I_k) \iint_{I_k} |\nabla u_2|^2 \, dm(X) \\ &\lesssim \ell(I_k) \cdot \frac{1}{\ell(I_k)^2} \iint_{I_k^*} |u_2(X)|^2 \, dm(X) \\ &\lesssim \iint_{I_k^*} \frac{|u_2(X)|^2}{\delta(X)} \, dm(X). \end{aligned}$$

Recall $I_k^* = (1+\theta) I_k$ is the dilation of I_k satisfying (3.11). Then summing up we get

$$\begin{aligned} \iint_{T(\Delta)} |\nabla u_2|^2 \, \delta(X) \, dm(X) &\lesssim \sum_k \iint_{I_k^*} \frac{|u_2(X)|^2}{\delta(X)} \, dm(X) \\ &\lesssim \iint_{T(\frac{3}{2}\Delta)} \frac{|u_2(X)|^2}{\delta(X)} \, dm(X). \end{aligned} \tag{4.15}$$

In the last line we use the finite overlap of $\{I_k^*\}$, and the fact that by taking θ sufficiently small, we can ensure that $I_k^* \subset T\big(\frac{3}{2}\Delta\big)$ for all I_k intersects $T(\Delta)$. Recall that $\frac{3}{2}\Delta = \Delta\big(q_0, \frac{3}{2}r\big)$ and $T\big(\frac{3}{2}\Delta\big)$ denotes $B\big(q_0, \frac{3}{2}r\big) \cap \Omega$.

Let $f_2^{\pm}$ denote the positive and negative parts of f_2, and let $u_2^{\pm} = \int_{\Gamma \setminus \tilde{\Delta}} f_2^{\pm} \, d\omega^X \ge 0$. There is a technical issue that $f_2^{\pm} \notin C_0^0(\Gamma)$; however by splitting $u_2^{\pm}$ as

$$\begin{aligned} u_2^+(X) &= \int_{\{f \ge f_{\tilde{\Delta}}\} \setminus \tilde{\Delta}} f \, d\omega^X - f_{\tilde{\Delta}} \omega^X(\{f \ge f_{\tilde{\Delta}}\} \setminus \tilde{\Delta}), \\ u_2^-(X) &= -\int_{\{f < f_{\tilde{\Delta}}\} \setminus \tilde{\Delta}} f \, d\omega^X + f_{\tilde{\Delta}} \omega^X(\{f < f_{\tilde{\Delta}}\} \setminus \tilde{\Delta}), \end{aligned}$$

we can confirm by combining Lemmas 2.42 and 2.43 that $u_2^{\pm} \in W_r(\Omega)$ are indeed legitimate solutions of L, and they can be continuously extended to $\tilde{\Delta}$ by zero. By the linearity of integration, we have $u_2 = \int_\Gamma f_2 \, d\omega^X = u_2^+ - u_2^-$. Let $v(X) := u_2^+(X) + u_2^-(X)$; again by linearity we have

$$v(X) = \int_\Gamma |f_2| \, d\omega^X = \int_{\Gamma \setminus \tilde{\Delta}} |f - f_{\tilde{\Delta}}| \, d\omega^X. \tag{4.16}$$

Thus $|u_2(X)| \le v(X)$ for all $X \in \Omega$. Moreover by the properties of $u_2^{\pm}$, we know that $v \in W_r(\Omega)$ is a solution of L, that $Tv = 0$ on $\tilde{\Delta}$ and that $v \in W_r(B(q_0, cr))$. (Recall that $\tilde{\Delta} = c\Delta = B(q_0, cr) \cap \Gamma$.)

We claim

$$v(X) \le C\|f\|_{\mathrm{BMO}(\sigma)} \quad \text{for all } X \in T(6\Delta). \tag{4.17}$$

By the definition (4.16), the function v vanishes on $\tilde{\Delta}$. Note that $\tilde{\Delta} \supset 12\Delta$ by the choice of $\tilde{\Delta}$, $v \in W_r(B(q_0, 12r))$ is a nonnegative solution in $T(12\Delta)$ and $Tv \equiv 0$ on 12Δ. Let A be a corkscrew point for $T(12\Delta)$; by the boundary Harnack inequality (2.54)

$$v(X) \le Cv(A) \quad \text{for all } X \in T(6\Delta).$$

For any $j \in \mathbb{N}$, let A_j be a corkscrew point for the surface ball $2^j\tilde{\Delta}$. Similar to (4.9), we get

$$\begin{aligned} v(A) &\lesssim \sum_{j=1}^{\infty} 2^{-j\beta} \int_{2^j\tilde{\Delta}\setminus 2^{j-1}\tilde{\Delta}} |f - f_{\tilde{\Delta}}| k^{A_j}\, d\sigma \\ &\le \sum_{j=1}^{\infty} 2^{-j\beta} \left(\fint_{2^j\tilde{\Delta}} |f - f_{\tilde{\Delta}}|^p \, d\sigma \right)^{\frac{1}{p}} \left(\fint_{2^j\tilde{\Delta}} |k^{A_j}|^r \, d\sigma \right)^{\frac{1}{r}} \sigma(2^j\tilde{\Delta}) \\ &\lesssim \sum_{j=1}^{\infty} 2^{-j\beta} \|f\|_{\mathrm{BMO}(\sigma)} \omega^{A_j}(2^j\tilde{\Delta}) \\ &\lesssim \|f\|_{\mathrm{BMO}(\sigma)}. \end{aligned} \tag{4.18}$$

Here p is a conjugate to r. We conclude the proof of (4.17).

Next, we show a finer estimate based off (4.17), which is

$$v(X) \le C\left(\frac{\delta(X)}{r}\right)^{\beta} \|f\|_{\mathrm{BMO}(\sigma)} \quad \text{for all } X \in T\left(\tfrac{3}{2}\Delta\right), \tag{4.19}$$

where $\beta \in (0, 1]$ is the exponent from Lemma 2.33. To this end, for any $X \in T\left(\frac{3}{2}\Delta\right)$, let q_X be a boundary point such that $|X - q_X| = \delta(X)$. Note that

$$|X - q_X| = \delta(X) \le |X - q_0| < \tfrac{3}{2}r;$$

i.e., $X \in B\left(q_X, \frac{3}{2}r\right) \cap \Omega$. Note also

$$|q_X - q_0| \le |q_X - X| + |X - q_0| < \tfrac{3}{2}r + \tfrac{3}{2}r = 3r,$$

so $\overline{B(q_X, 3r)} \subset B(q_0, 6r)$. Since $\tilde{\Delta} \supset 6\Delta \supset \Delta(q_X, 3r)$, we have $v \in W_r(B(q_X, 3r))$ is a nonnegative solution in $B(q_X, 3r) \cap \Omega$ and $Tv \equiv 0$ on $\Delta(q_X, 3r)$. By the boundary Hölder regularity (2.34) and the first part of this lemma (4.17), we conclude

$$\begin{aligned} v(X) &\lesssim \left(\frac{|X - q_X|}{3r}\right)^{\beta} \left(\frac{1}{m(B(q_X, 3r))} \iint_{B(q_X,3r)\cap\Omega} |v|^2 \, dm \right)^{\frac{1}{2}} \\ &\lesssim \left(\frac{\delta(X)}{r}\right)^{\beta} \sup_{T(6\Delta)} v \lesssim \left(\frac{\delta(X)}{r}\right)^{\beta} \|f\|_{\mathrm{BMO}(\sigma)}. \end{aligned}$$

Combining (4.19) and (4.15), we get

$$\iint_{T(\Delta)} |\nabla u_2|^2\,\delta(X)\,dm(X) \lesssim \frac{\|f\|^2_{\mathrm{BMO}(\sigma)}}{r^{2\beta}}\left(\iint_{T(\frac{3}{2}\Delta)} \delta(X)^{2\beta-1}\,dm(X)\right). \tag{4.20}$$

Since $2\beta-1>-1$, we can use Lemma 2.10 with exponent $\alpha=2\beta-1$ to get

$$\iint_{T(\Delta)} |\nabla u_2|^2\,\delta(X)\,dm(X) \lesssim r^d\,\|f\|^2_{\mathrm{BMO}(\sigma)} \lesssim \sigma(\Delta)\|f\|^2_{\mathrm{BMO}(\sigma)}. \tag{4.21}$$

Combining (4.14) and (4.21) finishes the proof. □

4C. ***From BMO-solvability to $\omega \in A_\infty(\sigma)$.*** In this subsection, we prove the other half of Theorem 1.11:

Theorem 4.22. *Assume that for any $f \in C^0_0(\Gamma)$, the solution $u = Uf \in W_r(\Omega)$ given by Lemmas 2.37 and 2.39 satisfies the property that $|\nabla u|^2\,\delta(X)\,dm(X)$ is a Carleson measure with*

$$\sup_{\Delta\subset\Gamma} \frac{1}{\sigma(\Delta)} \iint_{T(\Delta)} |\nabla u|^2\,\delta(X)\,dm(X) \le C\|f\|^2_{\mathrm{BMO}(\sigma)}. \tag{4.23}$$

Then $\omega \in A_\infty(\sigma)$, with the implicit constant depending on d, n, C_0, C_1 and the above constant C.

Let us start with proving the following Lemma.

Lemma 4.24. *Suppose the Dirichlet problem* (D) *is BMO-solvable. Then any nonnegative function $f \in C^0_0(\Gamma)$ whose support is contained in a surface ball Δ satisfies*

$$\int_\Delta f\,d\omega^A \le C\|f\|_{\mathrm{BMO}(\sigma)}. \tag{4.25}$$

Here A is a corkscrew point for Δ.

Proof. Since $f \in C^0_0(\Gamma)$ is a nonnegative function, by Lemma 2.37 $u = Uf \in W_r(\Omega)$ is a nonnegative solution of L. Suppose Δ has radius r. Consider another surface ball $\Delta' = B(q', r)\cap\Gamma$ of the same radius r and which is of distance $2r$ away from Δ. Thus in particular, $Tu = 0$ on $3\Delta'$ and $u \in W_r(B(q', 3r))$, by Lemma 2.37(i) and (iv). Applying the BMO-solvability assumption to $u = Uf$ and the surface ball Δ', we have

$$\iint_{T(\Delta')} |\nabla u|^2\,\delta(X)\,dm(X) \le C\sigma(\Delta')\|f\|^2_{\mathrm{BMO}(\sigma)}. \tag{4.26}$$

We have shown in (2.97) that

$$\iint_{T(\Delta')} |\nabla u|^2\,\delta(X)\,dm(X) \gtrsim \int_{\frac{1}{2}\Delta'} |S_{\frac{1}{2}r}u|^2\,d\sigma, \tag{4.27}$$

where $S_{\frac{1}{2}r}u$ is the truncated square function of aperture $\bar\alpha > \alpha$, whose value is determined in Lemma 2.59 and only depends on n, d, C_0, C_1 and α. In order to get a lower bound of the square function $S_{\frac{1}{2}r}u$, we decompose the nontangential cone $\Gamma_{\frac{1}{2}r}(q)$ into stripes as in (2.57) and use the Poincaré-type inequality proved in Lemma 2.59 for surface ball Δ'. Let m_1, m_2 be integers determined in Lemma 2.59. We obtain

$$\begin{aligned}
|S_{\frac{1}{2}r}u|^2(q) &= \iint_{\Gamma^{\bar\alpha}_{(1/2)r}(q)} |\nabla u|^2\, \delta(X)^{1-d}\, dm(X) \\
&\geq \frac{1}{m_1+m_2} \sum_{j=m_1+1}^{\infty} \iint_{\Gamma^{\bar\alpha}_{j-m_1\to j+m_2}(q)} |\nabla u|^2\, \delta(X)^{1-d}\, dm(X) \\
&\gtrsim \sum_{j=m_1+1}^{\infty} (2^{-j}r)^{1-d} \iint_{\Gamma^{\bar\alpha}_{j-m_1\to j+m_2}(q)} |\nabla u|^2\, dm(X) \\
&\gtrsim \sum_{j=m_1+1}^{\infty} (2^{-j}r)^{1-d}\cdot(2^{-j}r)^{-2} \iint_{\Gamma^{\alpha}_{j}(q)} u^2\, dm(X) \\
&\gtrsim \sum_{j=m_1+1}^{\infty} u^2(A_j),
\end{aligned}$$

where $A_j \in \Gamma_j(q)$ is a corkscrew point at the scale $2^{-j}r$. In the last inequality, we use the interior corkscrew condition, as each stripe of the cone $\Gamma_j(q)$ contains a ball of radius comparable to $2^{-j-1}r$ (as long as α is chosen to be big, say $\alpha > 2M$, where M is the corkscrew constant). Moreover,

$$\sum_{j=m_1+1}^{\infty} u^2(A_j) \geq u^2(A_{m_1}) \gtrsim u^2(A_1). \tag{4.28}$$

Recall for any $q \in \Delta'$, the point $A_1 = A_1(q)$ is a corkscrew point of $B(q, 2^{-1}r)$. Let A' be the corkscrew point for $T\left(\frac{1}{2}\Delta'\right)$; by Lemma 2.1 and the Harnack inequality, $u(A') \approx u(A_1)$. Therefore

$$|S_{\frac{1}{2}r}u|^2(q) \gtrsim u^2(A_1) \gtrsim u^2(A') \quad \text{for any } q \in \Delta'.$$

Combining this with (4.26) and (4.27), we get

$$\sigma(\Delta')\|f\|^2_{\mathrm{BMO}(\sigma)} \gtrsim \int_{\frac{1}{2}\Delta'} |S_{\frac{1}{2}r}u|^2\, d\sigma \gtrsim \sigma\left(\tfrac{1}{2}\Delta'\right)u^2(A') \gtrsim \sigma(\Delta')u^2(A'),$$

and thus

$$u(A') \lesssim \|f\|_{\mathrm{BMO}(\sigma)}. \tag{4.29}$$

Let A be a corkscrew point for Δ. Since Δ and Δ' have the same radius r and they are of distance $2r$ apart, we have $u(A) \sim u(A')$. By assumption f is supported on Δ; hence

$$u(A) = \int_{\Delta} f\, d\omega^A. \tag{4.30}$$

The lemma follows by combining (4.29) and (4.30). □

With that at hand, we pass to the proof of Theorem 4.22.

Proof of Theorem 4.22. By the change-of-pole formula in Lemma 2.86 and the Harnack inequality, to prove $\omega \in A_\infty(\sigma)$ and in particular (2.90), it suffices to show: for any $\epsilon > 0$ fixed, we can find $\eta = \eta(\epsilon)$ such that for any Borel set $E \subset \Delta$

$$\frac{\sigma(E)}{\sigma(\Delta)} < \eta \quad \text{implies} \quad \frac{\omega^A(E)}{\omega^A(\Delta)} < \epsilon. \tag{4.31}$$

Here Δ is a surface ball and A is a corkscrew point for Δ. In fact, since σ and ω are regular Borel measures, we may assume E is an open subset of Δ.

Recall from Lemma 2.79 that

$$\omega^A(\Delta) \geq C^{-1}$$

for some $C > 1$. Thus to show $\omega^A(E)/\omega^A(\Delta) < \epsilon$ it suffices to show $\omega^A(E) < C^{-1}\epsilon$. Let $\delta > 0$ be a small constant to be determined later; we define a function

$$f(x) = \max\{0, 1 + \delta \log M_\sigma \chi_E(x)\}, \tag{4.32}$$

where M_σ is the Hardy–Littlewood maximal function with respect to σ. Similar to Section 5.3 of [Zhao 2018], f satisfies:

- $0 \leq f \leq 1$, and $f \equiv 1$ on the open set E.
- $\|f\|_{\mathrm{BMO}(\sigma)} \leq A\delta$, where A is a constant independent of E.
- If
$$\frac{\sigma(E)}{\sigma(\Delta)} < \eta(\delta) \sim e^{-1/\delta}, \tag{4.33}$$
then f is supported in 2Δ.

Next we use a mollification argument to approximate f by continuous functions. Let φ be a radially symmetric smooth function on $\mathbb{R}^n$ such that $\varphi = 1$ on $B_{\frac{1}{2}}$, $\operatorname{supp} \varphi \subset B_1$ and $0 \leq \varphi \leq 1$. Let

$$\varphi_\epsilon(z) = \frac{1}{\epsilon^d}\varphi\left(\frac{z}{\epsilon}\right), \qquad f_\epsilon(x) = \frac{\int_{y\in\Gamma} f(y)\varphi_\epsilon(x-y)\,d\sigma(y)}{\int_{y\in\Gamma} \varphi_\epsilon(x-y)\,d\sigma(y)} \quad \text{for } x \in \Gamma. \tag{4.34}$$

Then these f_ϵ's satisfy the following properties:

- Each f_ϵ is continuous, and is supported in 3Δ.
- There is a constant C (independent of ϵ) such that $\|f_\epsilon\|_{\mathrm{BMO}(\sigma)} \leq C\|f\|_{\mathrm{BMO}(\sigma)}$.
- $f(x) \leq \liminf_{\epsilon\to 0} f_\epsilon(x)$ for all x in their support 3Δ.

The proof of the above properties is a slight modification of Appendix A of [Zhao 2018]: here the mollifier $\{\varphi_\epsilon\}$ is an approximation of identity of dimension d, instead of dimension $n-1$. The proof uses standard mollification arguments and the Ahlfors regularity of Γ. Moreover, the proof of the last property also uses the precise definition of f in (4.32).

Let A' be a corkscrew point with respect to 3Δ. The last property and Fatou's lemma imply

$$\int_{3\Delta} f(x)\, d\omega^{A'}(x) \le \int_{3\Delta} \liminf_{\epsilon\to 0} f_\epsilon(x)\, d\omega^{A'}(x) \le \liminf_{\epsilon\to 0} \int_{3\Delta} f_\epsilon(x)\, d\omega^{A'}(x). \tag{4.35}$$

Since each f_ϵ is nonnegative, continuous and supported on 3Δ, we apply Lemma 4.24 and get

$$\int_{3\Delta} f_\epsilon(x)\, d\omega^{A'}(x) \le C\|f_\epsilon\|_{\mathrm{BMO}(\sigma)} \le C'\|f\|_{\mathrm{BMO}(\sigma)}. \tag{4.36}$$

Combining (4.35) and (4.36), we get

$$\int_{3\Delta} f(x)\, d\omega^{A'}(x) \le C'\|f\|_{\mathrm{BMO}(\sigma)} \le C''\delta.$$

On the other hand, since $f \ge \chi_E$

$$\int_{3\Delta} f(x)\, d\omega^{A'}(x) \ge \omega^{A'}(E) \gtrsim \omega^A(E).$$

The last inequality follows from the Harnack inequality and the fact that A, A' are corkscrew points to surface balls $\Delta, 3\Delta$ respectively. Therefore $\omega^A(E) \le C\delta$, as long as the condition (4.33), i.e., $\sigma(E)/\sigma(\Delta) < \eta$, is satisfied. In other words, $\omega \in A_\infty(\sigma)$. □

Acknowledgement

We would like to thank Guy David for very helpful discussions in connection with this paper.

References

[Akman et al. 2016] M. Akman, S. Bortz, S. Hofmann, and J. M. Martell, "Rectifiability, interior approximation and harmonic measure", preprint, 2016. arXiv

[Akman et al. 2017] M. Akman, M. Badger, S. Hofmann, and J. M. Martell, "Rectifiability and elliptic measures on 1-sided NTA domains with Ahlfors–David regular boundaries", *Trans. Amer. Math. Soc.* **369**:8 (2017), 5711–5745. MR Zbl

[Azzam 2017] J. Azzam, "Semi-uniform domains and the A_∞ property for harmonic measure", preprint, 2017. arXiv

[Azzam and Mourgoglou 2017] J. Azzam and M. Mourgoglou, "Tangent measures of elliptic harmonic measure and applications", preprint, 2017. arXiv

[Azzam et al. 2016a] J. Azzam, J. Garnett, M. Mourgoglou, and X. Tolsa, "Uniform rectifiability, elliptic measure, square functions, and ε-approximability via an ACF monotonicity formula", preprint, 2016. arXiv

[Azzam et al. 2016b] J. Azzam, S. Hofmann, J. M. Martell, S. Mayboroda, M. Mourgoglou, X. Tolsa, and A. Volberg, "Rectifiability of harmonic measure", *Geom. Funct. Anal.* **26**:3 (2016), 703–728. MR Zbl

[Azzam et al. 2017] J. Azzam, S. Hofmann, J. M. Martell, K. Nyström, and T. Toro, "A new characterization of chord-arc domains", *J. Eur. Math. Soc.* **19**:4 (2017), 967–981. MR Zbl

[Badger 2012] M. Badger, "Null sets of harmonic measure on NTA domains: Lipschitz approximation revisited", *Math. Z.* **270**:1-2 (2012), 241–262. MR Zbl

[Bennewitz and Lewis 2004] B. Bennewitz and J. L. Lewis, "On weak reverse Hölder inequalities for nondoubling harmonic measures", *Complex Var. Theory Appl.* **49**:7-9 (2004), 571–582. MR Zbl

[Bishop and Jones 1990] C. J. Bishop and P. W. Jones, "Harmonic measure and arclength", *Ann. of Math.* (2) **132**:3 (1990), 511–547. MR Zbl

[Christ 1990] M. Christ, "A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral", *Colloq. Math.* **60/61**:2 (1990), 601–628. MR Zbl

[Coifman and Weiss 1977] R. R. Coifman and G. Weiss, "Extensions of Hardy spaces and their use in analysis", *Bull. Amer. Math. Soc.* **83**:4 (1977), 569–645. MR Zbl

[Coifman et al. 1985] R. R. Coifman, Y. Meyer, and E. M. Stein, "Some new function spaces and their applications to harmonic analysis", *J. Funct. Anal.* **62**:2 (1985), 304–335. MR Zbl

[Dahlberg 1977] B. E. J. Dahlberg, "Estimates of harmonic measure", *Arch. Rational Mech. Anal.* **65**:3 (1977), 275–288. MR Zbl

[Dahlberg et al. 1984] B. E. J. Dahlberg, D. S. Jerison, and C. E. Kenig, "Area integral estimates for elliptic differential operators with nonsmooth coefficients", *Ark. Mat.* **22**:1 (1984), 97–108. MR Zbl

[David and Jerison 1990] G. David and D. Jerison, "Lipschitz approximation to hypersurfaces, harmonic measure, and singular integrals", *Indiana Univ. Math. J.* **39**:3 (1990), 831–845. MR Zbl

[David and Semmes 1991] G. David and S. Semmes, *Singular integrals and rectifiable sets in $\mathbb{R}^n$: beyond Lipschitz graphs*, Astérisque **193**, Soc. Math. France, Paris, 1991. MR Zbl

[David and Semmes 1993] G. David and S. Semmes, *Analysis of and on uniformly rectifiable sets*, Math. Surveys and Monographs **38**, Amer. Math. Soc., Providence, RI, 1993. MR Zbl

[David et al. 2017] G. R. David, J. Feneuil, and S. Mayboroda, "Elliptic theory for sets with higher co-dimensional boundaries", preprint, 2017. arXiv

[David et al. 2019] G. David, J. Feneuil, and S. Mayboroda, "Dahlberg's theorem in higher co-dimension", *J. Funct. Anal.* **276**:9 (2019), 2731–2820. Zbl

[Dindos et al. 2011] M. Dindos, C. Kenig, and J. Pipher, "BMO solvability and the A_∞ condition for elliptic operators", *J. Geom. Anal.* **21**:1 (2011), 78–95. MR Zbl

[García-Cuerva and Rubio de Francia 1985] J. García-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland Math. Studies **116**, North-Holland, Amsterdam, 1985. MR Zbl

[Garnett et al. 2018] J. Garnett, M. Mourgoglou, and X. Tolsa, "Uniform rectifiability from Carleson measure estimates and ε-approximability of bounded harmonic functions", *Duke Math. J.* **167**:8 (2018), 1473–1524. MR Zbl

[Hofmann and Martell 2014] S. Hofmann and J. M. Martell, "Uniform rectifiability and harmonic measure, I: Uniform rectifiability implies Poisson kernels in L^p", *Ann. Sci. Éc. Norm. Supér.* (4) **47**:3 (2014), 577–654. MR Zbl

[Hofmann and Martell 2017] S. Hofmann and J. M. Martell, "Harmonic measure and quantitative connectivity: geometric characterization of the L^p-solvability of the Dirichlet problem, I", preprint, 2017. arXiv

[Hofmann et al. 2014] S. Hofmann, J. M. Martell, and I. Uriarte-Tuero, "Uniform rectifiability and harmonic measure, II: Poisson kernels in L^p imply uniform rectifiability", *Duke Math. J.* **163**:8 (2014), 1601–1654. MR Zbl

[Hofmann et al. 2016] S. Hofmann, J. M. Martell, and S. Mayboroda, "Uniform rectifiability, Carleson measure estimates, and approximation of harmonic functions", *Duke Math. J.* **165**:12 (2016), 2331–2389. MR Zbl

[Hofmann et al. 2017a] S. Hofmann, P. Le, J. M. Martell, and K. Nyström, "The weak-A_∞ property of harmonic and p-harmonic measures implies uniform rectifiability", *Anal. PDE* **10**:3 (2017), 513–558. MR Zbl

[Hofmann et al. 2017b] S. Hofmann, J. M. Martell, S. Mayboroda, T. Toro, and Z. Zhao, "Uniform rectifiability and elliptic operators with small Carleson norm", preprint, 2017. arXiv

[Hofmann et al. 2017c] S. Hofmann, J. M. Martell, and T. Toro, "A_∞ implies NTA for a class of variable coefficient elliptic operators", *J. Differential Equations* **263**:10 (2017), 6147–6188. MR Zbl

[Jerison and Kenig 1982] D. S. Jerison and C. E. Kenig, "Boundary behavior of harmonic functions in nontangentially accessible domains", *Adv. in Math.* **46**:1 (1982), 80–147. MR

[John and Nirenberg 1961] F. John and L. Nirenberg, "On functions of bounded mean oscillation", *Comm. Pure Appl. Math.* **14** (1961), 415–426. MR Zbl

[Lavrentyev 1936] M. Lavrentyev, "Boundary problems in the theory of univalent functions", *Mat. Sb.* **1**:43 (1936), 815–846. In Russian; translated in *Amer. Math. Soc. Trans.* **32** (1963) 1–35.

[Milakis et al. 2013] E. Milakis, J. Pipher, and T. Toro, “Harmonic analysis on chord arc domains”, *J. Geom. Anal.* **23**:4 (2013), 2091–2157. MR Zbl

[Mourgoglou 2015] M. Mourgoglou, “Uniform domains with rectifiable boundaries and harmonic measure”, preprint, 2015. arXiv

[Mourgoglou and Tolsa 2017] M. Mourgoglou and X. Tolsa, “Harmonic measure and Riesz transform in uniform and general domains”, *J. Reine Ang. Math.* (online publication October 2017).

[Riesz and Riesz 1920] F. Riesz and M. Riesz, “Über die Randwerte einer analytischen Funktion”, pp. 27–44 in *Compte rendus de Quatrième Congrès des Mathématiciens Scandianves* (Stockholm, 1916), edited by G. Mittag-Leffler, Almqvist and Wiksell, Uppsala, 1920. Zbl

[Semmes 1990] S. Semmes, “Analysis vs. geometry on a class of rectifiable hypersurfaces in $\mathbb{R}^n$”, *Indiana Univ. Math. J.* **39**:4 (1990), 1005–1035. MR Zbl

[Stein 1970] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Math. Series **30**, Princeton Univ. Press, 1970. MR Zbl

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Math. Series **43**, Princeton Univ. Press, 1993. MR Zbl

[Toro and Zhao 2017] T. Toro and Z. Zhao, “Boundary rectifiability and elliptic operators with $W^{1,1}$ coefficients”, *Adv. Calc. Var.* (online publication October 2017).

[Zhao 2018] Z. Zhao, “BMO solvability and A_∞ condition of the elliptic measures in uniform domains”, *J. Geom. Anal.* **28**:2 (2018), 866–908. MR Zbl

Received 26 Feb 2018. Revised 13 Oct 2018. Accepted 20 Nov 2018.

SVITLANA MAYBORODA: svitlana@math.umn.edu
School of Mathematics, University of Minnesota, Minneapolis, MN, United States

ZIHUI ZHAO: zhaozh@uw.edu, zzhao@ias.edu
Department of Mathematics, University of Washington, Seattle, WA, United States

and

Institute for Advanced Study, Princeton, NJ, 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 12 No. 7 2019

2157-5045(2019)12:7;1-J