Sharp L^p estimates for discrete second order Riesz transforms

Komla Domelevo, Stefanie Petermichl

Sharp L^p estimates for discrete second order Riesz transforms

Komla Domelevo [a] Stefanie Petermichl [a,1,2]

[a] *Institut de Mathématiques de Toulouse, Université Paul Sabatier, Toulouse, France*

Abstract

We show that multipliers of second order Riesz transforms on products of discrete abelian groups enjoy the L^p estimate $p^* - 1$, where $p^* = \max\{p, q\}$ and p and q are conjugate exponents. This estimate is sharp if one considers all multipliers of the form $\sum_i \sigma_i R_i R_i^*$ with $|\sigma_i| \leqslant 1$ and infinite groups. In the real valued case, we obtain better sharp estimates for some specific multipliers, such as $\sum_i \sigma_i R_i R_i^*$ with $0 \leqslant \sigma_i \leqslant 1$. These are the first known precise L^p estimates for discrete Calderón-Zygmund operators.

Key words: Discrete groups, Second order Riesz transforms, Sharp L^p estimates

1 Introduction

Sharp L^p estimates for norms of classical Calderón-Zygmund operators have been intriguing for a very long time. Only in rare cases, such optimal constants are known and they all rely in one way or another on the existence of a specific function of several variables that governs the proof. It is a result by Pichorides [18], that determined the exact L^p norm of the Hilbert transform in the disk. There is also a very streamlined proof by Essén [9] on the same subject. That this estimate carries over to $\mathbb{R}^N$ and the Riesz transforms is an observation by Iwaniec and Martin [11]. A probabilistic counterpart in the language of

Email address: `stefanie.petermichl@gmail.com` (Stefanie Petermichl).
URL: `http://math.univ-toulouse.fr/~petermic` (Stefanie Petermichl).

[1] Research supported in part by ANR-12-BS01-0013-02. The author is a member of IUF

[2] Correponding author, Tel:+33 5 61 55 76 59, Fax: +33 5 61 55 83 85

orthogonal martingales was deduced by Bañuelos and Wang in [3], using the idea of Essén in a probabilistic setting.

The only other sharp L^p estimates for classical Calderón-Zygmund operators we are aware of are those closely related to the real part of the Beurling-Ahlfors operator, $R_1^2 - R_2^2$, where R_i are the two Riesz transforms in $\mathbb{R}^2$. The first estimate of this type is due to Nazarov and Volberg [22] and was used for a partial result towards the famous $p^* - 1$ conjecture, the L^p norm of the Beurling-Ahlfors transform. It was not clear until the work of Geiss, Montgomery-Smith and Saksman [10], that the Nazarov-Volberg estimate for the second order Riesz transform was infact sharp. In contrast to the estimate for the Hilbert transform, in this case an underlying probabilistic estimate - Burkholder's famous theorem on differentially subordinate martingales - was known first and could be used to obtain the optimal estimate for the real part of the Beurling-Ahlfors operator and its appropriate generalisations to $\mathbb{R}^N$.

It has been observed, that these L^p estimates can be obtained without any loss for the case of compact Lie groups, see Arcozzi [1] and Bañuelos and Baudoin [2].

The situation changes, when the continuity assumption is lost. Optimal norm estimates for the discrete Hilbert transform on the integers are a long standing open question. Infact, until now, there were no known sharp L^p estimates for any Calderón-Zygmund operator when entering discrete abelian groups such as $\mathbb{Z}$ or cyclic groups $\mathbb{Z}/m\mathbb{Z}$. In the present paper, we provide such optimal estimates for discrete second order Riesz transforms.

The definition of derivatives and thus the Laplacian respects the discrete setting. For each direction i, there are two choices of Riesz transforms, that depend on whether right or left handed derivatives are used:

$$R_+^i \circ \sqrt{-\Delta} = \partial_+^i \quad \text{and} \quad R_-^i \circ \sqrt{-\Delta} = \partial_-^i.$$

One often sees $R_i$written for R_+^i and R_i^* for $-R_-^i$. The discrete derivatives that come into play induce difficulties that are due to their non-local nature. To underline the artifacts one may encounter, we briefly discuss a related subject, dimensional behavior of Riesz vectors in a variety of different situations.

F. Lust-Piquard showed in [13] and [14] that square functions of Riesz transforms on products of discrete abelian groups of a single generator $\mathbb{Z}, \mathbb{Z}/k\mathbb{Z}$ for k=2,3,4 enjoy dimension-free L^pestimates when $p \geqslant 2$ but suffer from a dimension dependence when $1 < p < 2$. She used non-commutative methods to obtain this nice positive result for $p \geqslant 2$. She provided an example originating on the hyper cube $(\mathbb{Z}/2\mathbb{Z})^N$, demonstrating dimensional growth when $p < 2$.

This difficulty of dimensional dependence for some p in the discrete case is in a sharp contrast to numerous, both interesting and difficult, dimension-free estimates in some very general continuous settings. We refer the reader to the result by Stein [20] in the classical case. Probabilistic methods, applied by P.A. Meyer [15] lead to the first proof of this theorem in the Gaussian setting, whereas Pisier found an analytic proof in [19]. We refer the reader to more general settings on compact Lie groups [1], Heisenberg groups [8] and the best to date estimate for Riesz vectors on Riemannian manifolds with a condition on curvature [6].

2 Definitions and main results

Let G be an additive discrete abelian group generated by e and $f : G \to \mathbb{C}$. Its right and left hand derivatives are

$$\partial_+ f(n) = f(n+e) - f(n) \quad \text{and} \quad \partial_- f(n) = f(n) - f(n-e).$$

The reader may keep in mind that $G \in \{\mathbb{Z}, \mathbb{Z}/m\mathbb{Z}; m \geqslant 2\}$. We then define partial derivatives $\partial^i_\pm$ accordingly on products of such groups. Here, we denote by e_i the ith unit element in G^N and

$$\partial^i_+ f(n) = f(n+e_i) - f(n) \quad \text{and} \quad \partial^i_- f(n) = f(n) - f(n-e_i).$$

A priori, we consider G^N, but it will be clear that our methods apply to mixed cases $\otimes_{i=1}^N G_i$ where G_i are not necessarily the same discrete groups. The discrete Laplacian becomes

$$\Delta = \sum_{i=1}^N \partial^i_+ \partial^i_- = \sum_{i=1}^N \partial^i_- \partial^i_+.$$

For each direction i there are two choices of Riesz transforms, defined in a standard manner:

$$R^i_+ \circ \sqrt{-\Delta} = \partial^i_+ \quad \text{and} \quad R^i_- \circ \sqrt{-\Delta} = \partial^i_-.$$

In this text, we are concerned with second order Riesz transforms defined as

$$R_i^2 := R^i_+ R^i_- = -R_i R_i^*$$

and combinations thereof, such as

$$R_\alpha^2 := \sum_{i=1}^{N} \alpha_i R_i^2,$$

where the $\alpha = \{\alpha_i\} \in \mathbb{C}^N : |\alpha_i| \leqslant 1$. When p and q are conjugate exponents, let

$$p^* = \max\{p, q\} = \begin{cases} p & 2 \leqslant p \\ q = \frac{p}{p-1} & 1 < p \leqslant 2 \end{cases}.$$

So $p^* - 1 = \max\left\{p-1, \frac{1}{p-1}\right\}$. The following are our main results.

Theorem 1 $R_\alpha^2 : L^p(G^N, \mathbb{C}) \to L^p(G^N, \mathbb{C})$ *enjoys the operator norm estimate* $\|R_\alpha^2\| \leqslant p^* - 1$. *The estimate above is sharp when* $G = \mathbb{Z}$ *and* $N \geqslant 2$.

In some cases, better estimates are available in the real valued case.

Theorem 2 $R_\alpha^2 : L^p(G^N, \mathbb{R}) \to L^p(G^N, \mathbb{R})$ *with* $\alpha = \{\alpha_i\}_{i=1,\ldots,N} \in \mathbb{R}^N$ *enjoys the sharp norm estimate* $\|R_\alpha^2\| \leqslant \mathfrak{C}_{\min\alpha, \max\alpha, p}$, *where these are the Choi constants.*

The Choi constants (see [7]) are not explicit, except $\mathfrak{C}_{-1,1,p} = p^* - 1$. However, about $\mathfrak{C}_{0,1,p}$ it is known that

$$\mathfrak{C}_{0,1,p} = \tfrac{p}{2} + \tfrac{1}{2}\log\left(\tfrac{1+e^{-2}}{2}\right) + \tfrac{\beta_2}{p} + \ldots.$$

with $\beta_2 = \log^2\left(\frac{1+e^{-2}}{2}\right) + \frac{1}{2}\log\left(\frac{1+e^{-2}}{2}\right) - 2\left(\frac{e^{-2}}{1+e^{-2}}\right)^2$.

The proof of our results uses Bellman functions and passes via embedding theorems that are a somewhat stronger estimate than Theorem 1 and Theorem 2.

Abelian groups of a single generator are isomorphic to $\mathbb{Z}$ or $\mathbb{Z}/m\mathbb{Z}$ for some integer $m \geqslant 2$. We are going to make use of the Fourier transform on these groups. In the case of the infinite group $\mathbb{Z}$, the charater set is $\mathbb{T} = \left[-\frac{1}{2}, \frac{1}{2}\right]$. We briefly review the Fourier Transform in $\mathbb{Z}^N$. We write $\mathbb{T}^N$for $\left[-\frac{1}{2}, \frac{1}{2}\right]^N$. $\mathbb{T}^N$ is the character group of $\mathbb{Z}^N$.

$$\hat{}: (f : \mathbb{Z}^N \to \mathbb{C}) \to (\hat{f} : \mathbb{T}^N \to \mathbb{C}), f(n) \mapsto \hat{f}(\xi) = \sum_{\mathbb{Z}^N} f(n) e^{-2\pi i n \cdot \xi}$$

with inversion formula

$$f(n) = \int_{\mathbb{T}^N} \hat{f}(\xi) e^{2\pi i n\cdot\xi} d\xi.$$

We observe that discrete derivatives are multiplier operators

$$\widehat{\partial^j_\pm f}(\xi) = \pm \hat{f}(\xi) \cdot (e^{\pm 2\pi i \xi_j} - 1) = \hat{f}(\xi) \cdot 2ie^{\pm\pi i\xi_j} \sin(\pi\xi_j).$$

For the discrete Laplacian there holds

$$\begin{aligned}(\Delta f)(n) &= \sum_i (\partial^i_- \partial^i_+ f)(n) = \sum_i \partial^i_-(f(n+e_i) - f(n)) \\ &= \sum_i f(n+e_i) - 2f(n) + f(n-e_i).\end{aligned}$$

We obtain the multiplier of Δ:

$$\hat{\Delta} = \sum_i (-2 + e^{2\pi i\xi_i} + e^{-2\pi i \xi_i}) = -4 \sum_i \sin^2(\pi\xi_i),$$

and the multiplier of the left and right handed jth Riesz transforms:

$$\widehat{R^j_\pm} = \frac{2\,\mathrm{ie}^{\pm\pi i\xi_j} \sin(\pi\xi_j)}{\sqrt{4\sum_i \sin^2(\pi\xi_i)}}.$$

Second order Riesz transforms $R^2_j := R^j_+ R^j_-$ have the following multipliers:

$$\widehat{R^2_j} = \frac{-4\sin^2(\pi\xi_j)}{4\sum_i \sin^2(\pi\xi_i)}.$$

In $(\mathbb{Z}/m\mathbb{Z})^N$, not many changes are required. Recall that a character χ of a locally compact abelian group G is a homomorphism $\chi : G \to S^1$. Let e be a generator of our group $G = \mathbb{Z}/m\mathbb{Z}$, so $me = 0$ and $ne \neq 0$ for $0 < n < m$. Therefore $\chi(e)^m = 1$ so that the character group of $\mathbb{Z}/m\mathbb{Z}$ consists of the m roots of unity. The character set of $(\mathbb{Z}/m\mathbb{Z})^N$ is therefore isomorphic to itself, like for all finite abelian groups. The Fourier transform looks as follows:

$$\hat{f}(\xi) = \sum_{(\mathbb{Z}/m\mathbb{Z})^N} f(n) e^{-\frac{2\pi i}{m} n\cdot\xi}$$

as well as

$$f(n) = \frac{1}{m^N} \sum_{(\mathbb{Z}/m\mathbb{Z})^N} \hat{f}(\xi) e^{\frac{2\pi i}{m} n\cdot\xi}.$$

These are special cases of the more general formulae $\hat{f}(\chi) = \sum_G f(g)\bar{\chi}(g)$ and $f(x) = \frac{1}{\sharp G}\sum_G \hat{f}(\chi)\chi(g)$. We observe the following about discrete derivatives:

$$\widehat{\partial^i_\pm} = 2\,\mathrm{i}\mathrm{e}^{\pm\frac{\pi i}{m}\xi_i}\sin\left(\frac{\pi}{m}\xi_i\right) \quad \text{and} \quad \hat{\Delta} = \widehat{\sum_i \partial^i_- \partial^i_+} = -4\sum_i \sin^2\left(\frac{\pi}{m}\xi_i\right).$$

The Riesz transforms are defined as above and we have for all directions $1 \leqslant j \leqslant N$,

$$\widehat{R^2_j} = \frac{-4\sin^2\left(\frac{\pi}{m}\xi_j\right)}{4\sum_i \sin^2\left(\frac{\pi}{m}\xi_i\right)}.$$

We will be using continuous in time heat extensions $\tilde{f} : G^N \times [0,\infty) \to \mathbb{C}$ of functions $f : G^N \to \mathbb{C}$. Namely for all $t \geqslant 0$, $\tilde{f}(t) = e^{t\Delta}f$. Using this notation, we are ready to state two further result, the following bilinear estimates:

Theorem 3 *Let f and g be test functions on G^N and p and q conjugate exponents. Then we have the estimate*

$$2\sum_{i=1}^{N}\int_0^\infty \sum_{G^N} |\partial^i_+\tilde{f}(n,t)||\partial^i_+\tilde{g}(n,t)|\mathrm{d}t \leqslant (p^*-1)\|f\|_p\|g\|_q. \tag{1}$$

Theorem 4 *Let f and g be real valued test functions on G^N and p and q conjugate exponents. Then we have the estimate*

$$2\sum_{i=1}^{N}\int_0^\infty \sum_{G^N} \left[\partial^i_+\tilde{f}\,\partial^i_+\tilde{g}\right]_\pm \mathrm{d}t \leqslant \mathfrak{C}_p\|f\|_p\|g\|_q,$$

where $[\,\cdot\,]_\pm$ denotes positive and negative parts.

3 The $p^* - 1$ estimates.

We take advantage of an intimate connection between differentially subordinate martingales and representation formulae for our type of singular integrals using heat extensions. This idea was first used in [17] in a weighted context. This here is the first application to a discrete group.

We will be using continuous in time heat extensions $P_t f : G^N \times [0,\infty) \to \mathbb{C}$ of a function $f : G^N \to \mathbb{C}$. Precisely, we set $P_t := e^{t\Delta}$, so that $(P_t f)(t) := e^{t\Delta} f$. Denoting for convenience $\tilde{f}(t) := P_t f$, we have that the function $\tilde{f}$ solves the semi-continuous heat equation $\partial_t \tilde{f} - \Delta \tilde{f} = 0$ with initial condition $\tilde{f}(0) = f$. The group structure allows us to express the semi-group P_t in terms of a convolution kernel involving the elementary solution $K : G^N \times [0,\infty) \to \mathbb{C}$ solution to the semi-continuous heat equation with initial data $K(n,0) := \delta_0(n)$, where $\delta_0(0) = 1$ and $\delta_0(n) = 0$ for $n \neq 0$. We have $K \geqslant 0$ and for all (n,t),

$$\tilde{f}(n,t) = (K(t) * f)(n) = \sum_{m \in G^N} K(m,t) f(0, n-m), \sum_{m \in G^N} K(m,t) = 1 \quad (2)$$

We will note in general $K(\cdot,\cdot\,; 0, m)$ the elementary solution with initial condition $K(0) := \delta_0(\cdot - m)$ translated by $m \in G^N$, namely $K(n,t\,; m, 0) = K(n-m,t)$. It is well known that this kernel, defined on a regular graph such as G^N decays exponentially with respect to its variable n, for G an infinite group.

Let R_i^2 be the square of the ith Riesz transform, as defined above. Let f and g be test functions. Let us first note, that R_i maps constants to 0, so we may assume that in the formulae below, $\hat{g}(0) = 0$.

Lemma 1 *If $\hat{g}(0) = 0$ then*

$$(f, R_j^2 g) = -2 \int_0^\infty \sum_{G^N} \partial_+^j \tilde{f}(n,t) \overline{\partial_+^j \tilde{g}(n,t)} \mathrm{d}t \quad (3)$$

and the sums and integrals that arise converge absolutely.

Proof. In the infinite case, $\mathbb{Z}$, we first notice that for $\xi \neq 0$, we have

$$2 \int_0^\infty e^{-8t \sum_i \sin^2(\pi \xi_i)} dt = - \frac{e^{-8t \sum_i \sin^2(\pi \xi_i)}}{4 \sum_i \sin^2(\pi \xi_i)} \Bigg|_0^\infty = \frac{1}{4 \sum_i \sin^2(\pi \xi_i)}.$$

Now, we calculate

$$
\begin{aligned}
(f, R_j^2 g) &= \sum_{\mathbb{Z}^N} f(n)\overline{R_j^2 g(n)} = \int_{\mathbb{T}^N} \hat{f}(\xi)\frac{-4\sin^2(\pi\xi_j)}{4\sum_i \sin^2(\pi\xi_i)}\overline{\hat{g}(\xi)}\mathrm{d}A(\xi) \\
&= -2\int_{\mathbb{T}^N}\int_0^\infty 4\sin^2(\pi\xi_j)e^{-8t\sum_i \sin^2(\pi\xi_i)}\widehat{f}(\xi)\overline{\widehat{g}(\xi)}\mathrm{d}t\mathrm{d}A(\xi) \\
&= -2\int_{\mathbb{T}^N}\int_0^\infty 2ie^{\pi i\xi_j}\sin(\pi\xi_j)e^{-4t\sum_i \sin^2(\pi\xi_i)}\widehat{f}(\xi)\times \\
&\qquad \times \overline{2ie^{\pi i\xi_j}\sin(\pi\xi_j)e^{-4t\sum_i \sin^2(\pi\xi_i)}\hat{g}(\xi)}\mathrm{d}t\mathrm{d}A(\xi) \\
&= -2\int_0^\infty \sum_{\mathbb{Z}^N} \partial_+^j\tilde{f}(n,t)\overline{\partial_+^j\tilde{g}(n,t)}\mathrm{d}t
\end{aligned}
$$

Here, $\tilde{f}(n,t)$ denotes the semi-continuous heat extension of the function f.

In the cyclic case, there are not many changes. Observe that even in $\mathbb{Z}/2\mathbb{Z}$, where there is no 'room' for three points in a row, the Laplacian is a well defined negative operator and the equation diffuses correctly. Also here, we may assume $\hat{g}(0) = 0$. As before, we have when $\xi \neq 0$

$$
2\int_0^\infty e^{-8t\sum_i \sin^2\left(\frac{\pi}{m}\xi_i\right)}\mathrm{d}t = -\frac{e^{-8t\sum_i \sin^2\left(\frac{\pi}{m}\xi_i\right)}}{4\sum_i \sin^2\left(\frac{\pi}{m}\xi_i\right)}\Bigg|_0^\infty = \frac{1}{4\sum_i \sin^2\left(\frac{\pi}{m}\xi_i\right)}.
$$

We calculate in the same manner

$$
\begin{aligned}
(f, R_j^2 g) &= \sum_{(\mathbb{Z}/m\mathbb{Z})^N} f(n)\overline{R_j^2 g(n)} = \sum_{(\mathbb{Z}/m\mathbb{Z})^N} \hat{f}(\xi)\frac{-4\sin^2\left(\frac{\pi}{m}\xi_j\right)}{4\sum_i \sin^2\left(\frac{\pi}{m}\xi_i\right)}\overline{\hat{g}(\xi)} \\
&= -2\sum_{(\mathbb{Z}/m\mathbb{Z})^N}\int_0^\infty 4\sin^2\left(\frac{\pi}{m}\xi_j\right)e^{-8t\sum_i \sin^2\left(\frac{\pi}{m}\xi_i\right)}\widehat{f}(\xi)\overline{\widehat{g}(\xi)}\mathrm{d}t \\
&= -2\sum_{(\mathbb{Z}/m\mathbb{Z})^N}\int_0^\infty 2ie^{\frac{\pi i}{m}\xi_j}\sin\left(\frac{\pi}{m}\xi_j\right)e^{-4t\sum_i \sin^2\left(\frac{\pi}{m}\xi_i\right)}\widehat{f}(\xi)\times \\
&\qquad \times\overline{2ie^{\frac{\pi i}{m}\xi_j}\sin\left(\frac{\pi}{m}\xi_j\right)e^{-4t\sum_i \sin^2\left(\frac{\pi}{m}\xi_i\right)}\hat{g}(\xi)}\mathrm{d}t \\
&= -2\int_0^\infty \sum_{(\mathbb{Z}/m\mathbb{Z})^N} \partial_+^j\tilde{f}(n,t)\overline{\partial_+^j\tilde{g}(n,t)}\mathrm{d}t
\end{aligned}
$$

This concludes the proof of Lemma 1. *QED*

We are going to use the Bellman function method to derive our estimates. In order to do so, we are going to need a tool to control the terms we see on the right hand side of our representation formulae. This is the content of the next subsection.

In this section we derive the existence of a function with a certain convexity condition. There is an explicit construction of (upto smoothness) the type of function we need in [21] of an elaborate complexity. There is also a much simpler explicit expression in [16] at the cost of a factor, which does bother us for our purposes. In our proof, we only need the existence, not the explicit expression. This part is standard for readers with some background in Bellman functions, however this argument is not usually written carefully. We will also apply similar considerations to derive a Bellman function that will give us our estimates in terms of the Choi constant.

Theorem 5 *For any $1 < p < \infty$ we define*

$$D_p := \{\boldsymbol{v} = (\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{f}, \boldsymbol{g}) \subset \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{C} \times \mathbb{C} : |\boldsymbol{f}|^p < \boldsymbol{F}, |\boldsymbol{g}|^q < \boldsymbol{G}\}.$$

Let K be any compact subset of D_p and let ε be an arbitrary small positive number. Then there exists a function $B_{\varepsilon,p,K}(\boldsymbol{v})$ that is infinitely differentiable in an ε-neighborhood of K and such that

$$0 \leqslant B_{\varepsilon,p,K}(\boldsymbol{v}) \leqslant (1 + \varepsilon C_K)(p^* - 1)\boldsymbol{F}^{1/p}\boldsymbol{G}^{1/q},$$

and

$$-\mathrm{d}^2_{\boldsymbol{v}} B_{\varepsilon,p,K}(\boldsymbol{v}) \geqslant 2|\mathrm{d}\boldsymbol{f}||\mathrm{d}\boldsymbol{g}|. \tag{4}$$

Before we proceed with the proof of this theorem, we will require a lemma that follows directly from Burkholder's famous $p^* - 1$ estimate for differentially subordinate martingales. Readers who do not wish to admit any probability, can take the Vasyunin-Volberg function [21].

Lemma 2 *For every $1 < p < \infty$ there exists a function B_p of four variables $\boldsymbol{v} = (\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{f}, \boldsymbol{g})$ in the domain $\overline{D_p} = \{(\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{f}, \boldsymbol{g}) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{C} \times \mathbb{C} : |\boldsymbol{f}|^p \leqslant \boldsymbol{F}, |\boldsymbol{g}|^q \leqslant \boldsymbol{G}\}$ so that*

$$0 \leqslant B(\boldsymbol{v}) \leqslant (p^* - 1)\boldsymbol{F}^{1/p}\boldsymbol{G}^{1/q},$$

and

$$B(\boldsymbol{v}) \geqslant \frac{1}{2}B(\boldsymbol{v}_+) + \frac{1}{2}B(\boldsymbol{v}_-) + \frac{1}{4}|\boldsymbol{f}_+ - \boldsymbol{f}_-||\boldsymbol{g}_+ - \boldsymbol{g}_-| \tag{5}$$

if $\boldsymbol{v}_+ + \boldsymbol{v}_- = 2\boldsymbol{v}$ *and* $\boldsymbol{v}$, $\boldsymbol{v}_+$, $\boldsymbol{v}_-$ *are in the domain. Here* $\|\cdot\|$ *denotes the usual* ℓ_2 *norm in* $\mathbb{C}$.

Proof. The existence of this function follows from Burkholder's theorem directly in one of its simplest forms (see [4] and [5]):

Theorem 6 *(Burkholder) Let* $(\Omega, \mathfrak{F}, P)$ *be a probability space with filtration* $\mathfrak{F} = (\mathfrak{F}_n)_{n\in\mathbb{N}}$. *Let* X *and* Y *be complex valued martingales with differential subordination* $|Y_0(\omega)| \leqslant |X_0(\omega)|$ *and* $|Y_n(\omega) - Y_{n-1}(\omega)| \leqslant |X_n(\omega) - X_{n-1}(\omega)|$ *for almost all* $\omega \in \Omega$ *. Then* $\|Y\|_{L^p} \leqslant (p^* - 1)\|X\|_{L^p}$ *where the* L^p *norms are in the sense of martingales.*

Now, if $\mathcal{D}$ is the dyadic grid on the real line, let $\mathfrak{F}_n$contain all dyadic intervals of size at least $2^{-n}|J|$. For a test function f supported on a dyadic interval J, its dyadic approximation gives rise to a martingale $\forall \omega \in J, X_0(\omega) = \mathbb{E}f$, with

$$\forall n > 0, \quad X_n(\omega) = \sum_{I\in\mathcal{D}(J), |I|>2^{-n}|J|} (f, h_I)h_I(\omega).$$

With the help of random multiplications $\sigma = \{\sigma_I \in S^1; \quad I \in \mathcal{D}(J)\}$ one defines the martingale transform Y of X as $\forall\omega \in J, Y_0(w) = X_0(w)$ and

$$\forall n > 0, \quad Y_n(\omega) = \sum_{I\in\mathcal{D}(J), |I|>2^{-n}|J|} \sigma_I(f, h_I)h_I(\omega).$$

Here X and Y enjoy the differential subordination property. Consequently, Burkholder's theorem asserts that $\|Y\|_p \leqslant (p^* - 1)\|X\|_p$. This implies that the operator

$$T_\sigma : f \mapsto \sum_{I\in\mathcal{D}(J)} \sigma_I(f, h_I)h_I$$

has a uniform L^p bound of at most $p^* - 1$. So for test functions f, g supported in J we have the estimate

$$\sup_\sigma \frac{1}{|J|}|(T_\sigma f, g)| \leqslant (p^* - 1)\langle|f|^p\rangle_J^{1/p}\langle|g|^q\rangle_J^{1/q}.$$

Expanding the orthonormal series on the left and choosing the worst σ yields the estimate

$$\begin{aligned}\frac{1}{|J|}\sum_{I\in\mathcal{D}(J)} |(f, h_I)||(g, h_I)| &= \frac{1}{4|J|}\sum_{I\in\mathcal{D}(J)} |I||\langle f\rangle_{I_+} - \langle f\rangle_{I_-}||\langle g\rangle_{I_+} - \langle g\rangle_{I_-}| \\ &\leqslant (p^* - 1)\langle|f|^p\rangle_J^{1/p}\langle|g|^q\rangle_J^{1/q}.\end{aligned}$$

This means that the function we are looking for takes the form:

$$B_p(\boldsymbol{F},\boldsymbol{G},\boldsymbol{f},\boldsymbol{g}) = \sup_{f,g}\left\{\frac{1}{4|J|}\sum_{I\in\mathcal{D}(J)}|I||\langle f\rangle_{I_+}-\langle f\rangle_{I_-}||\langle g\rangle_{I_+}-\langle g\rangle_{I_-}|:\right.$$
$$\left.\langle f\rangle_J=\boldsymbol{f},\langle g\rangle_J=\boldsymbol{g},\langle|f|^p\rangle_J=\boldsymbol{F},\langle|g|^q\rangle_J=\boldsymbol{G}\right\}.$$

The supremum runs over functions supported in J, but scaling shows that the function B_p itself does not depend upon the interval J. The function B_p has the following properties stated in the lemma:

1. We have $B_p \geqslant 0$. This means that for any $(\boldsymbol{F},\boldsymbol{G},\boldsymbol{f},\boldsymbol{g})\in D_p$, one can find complex valued functions f and g such that $(\langle f\rangle_J,\langle g\rangle_J,\langle|f|^p\rangle_J,\langle|g|^q\rangle_J) = (\boldsymbol{F},\boldsymbol{G},\boldsymbol{f},\boldsymbol{g})$. Otherwise the supremum would run over the empty set yielding $B_p(\boldsymbol{F},\boldsymbol{G},\boldsymbol{f},\boldsymbol{g}) = -\infty$. On the contrary, as soon as there exist (f,g) such that $(\langle f\rangle_J,\langle g\rangle_J,\langle|f|^p\rangle_J,\langle|g|^q\rangle_J) = (\boldsymbol{F},\boldsymbol{G},\boldsymbol{f},\boldsymbol{g})$, the expression of B_p ensures that $B_p(\boldsymbol{F},\boldsymbol{G},\boldsymbol{f},\boldsymbol{g})\geqslant 0$.

   Let therefore $(\boldsymbol{F},\boldsymbol{G},\boldsymbol{f},\boldsymbol{g})\in D_p$. Choose f so that $|f| = \boldsymbol{F}^{1/p}$. Then trivially $\langle|f|^p\rangle_J=\boldsymbol{F}$. Choose φ so that $e^{i\varphi}\boldsymbol{f}\in\mathbb{R}$. Since $|\boldsymbol{f}|\leqslant\boldsymbol{F}^{1/p}$ there is $c\in\mathbb{R}:|c|\leqslant 1$ with $e^{i\varphi}\boldsymbol{f}=c\boldsymbol{F}^{1/p}$. Now choose Ka subinterval of J so that $c=\frac{-|K|+|J\backslash K|}{|J|}$ and choose $f\equiv -e^{-i\varphi}\boldsymbol{F}^{1/p}$on K and $f\equiv e^{-i\varphi}\boldsymbol{F}^{1/p}$ on $J\backslash K$. Similar considerations for variables $\boldsymbol{g},\boldsymbol{G}$ show that $B_p>-\infty$ and therefore $B_p\geqslant 0$.
2. The upper estimate on B_p follows from Burkholder's theorem and the fact that we only allow values $(\boldsymbol{F},\boldsymbol{G},\boldsymbol{f},\boldsymbol{g})$ that can occur when these numbers are averages of functions f,g respectively. This is why the domain is restricted so as not to violate Hölder's inequality: $|\boldsymbol{f}|^p\leqslant\boldsymbol{F}$, $|\boldsymbol{g}|^q\leqslant\boldsymbol{G}$.
3. The dissipation estimate can be seen as follows. Fix the interval J and let $2\boldsymbol{v}=\boldsymbol{v}_++\boldsymbol{v}_-$ with $\boldsymbol{v},\boldsymbol{v}_+,\boldsymbol{v}_-$ in the domain of B. Now construct functions f,g with the prescribed averages on J_+ and J_- respectively as shown above. The resulting functions f,g defined on J correspond to $\boldsymbol{v}$. So

$$\begin{aligned}B(\boldsymbol{v})&\geqslant\frac{1}{4|J|}\sum_{I\in\mathcal{D}(J)}|I||\langle f\rangle_{I_+}-\langle f\rangle_{I_-}||\langle g\rangle_{I_+}-\langle g\rangle_{I_-}|\\&\geqslant\frac{1}{4}|\boldsymbol{f}_+-\boldsymbol{f}_-||\boldsymbol{g}_+-\boldsymbol{g}_-|\\&\quad+\frac{1}{4|J|}\sum_{I\in\mathcal{D}(J),I\neq J}|I||\langle f\rangle_{I_+}-\langle f\rangle_{I_-}||\langle g\rangle_{I_+}-\langle g\rangle_{I_-}|.\end{aligned}$$

   Taking supremum over all f,g made to match $\boldsymbol{v}_\pm$ gives the required convexity property (5) of B_p.

This concludes the proof of Lemma 2. *QED*

Notice that the function B_p attained in this manner need not be smooth. This is the point of Theorem 5. We refer the reader also to [22], where this has previously been done.

Proof. (of Theorem 5) Using Lemma 2 above, the rest of the proof is a standard mollifying argument, combined with the observation that midpoint concavity is the same as concavity defined through the Hessian. This is a well known fact that only uses an application of Taylor's theorem and integration against a tent function (Green's function on $\mathbb{Z}$). Let us fix a compact, convex set $K \subset D_p$ and a corresponding $\varepsilon > 0$, that is very small in comparison to the distance of K to the boundary of the domain D_p. Take a standard mollifier φ and mollify B_p by convolution with dilates of φ. Then one can show that the resulting function $B_{\varepsilon,p,K}$ has the properties

$$0 \leqslant B_{\varepsilon,p,K}(\boldsymbol{v}) \leqslant (1+\varepsilon C_K)(p^*-1)\boldsymbol{F}^{1/p}\boldsymbol{G}^{1/q},$$

and

$$-\mathrm{d}^2_{\boldsymbol{v}} B_{\varepsilon,p,K}(\boldsymbol{v}) \geqslant 2|\mathrm{d}\boldsymbol{f}||\mathrm{d}\boldsymbol{g}|. \tag{6}$$

QED

Now, we are ready to derive our dynamics condition in the lemma below. In continuous settings, this becomes just an application of the chain rule. The absence of a chain rule in the discrete setting is a true obstacle that is very characteristic of discrete groups. It can be overcome in this very particular case, because our Hessian estimate (6) for the function $B(\boldsymbol{v})$ is universal and does not depend on $\boldsymbol{v}$.

Lemma 3 *Let us assume a compact and convex subset $K \subset D_p$ has been chosen. Let*

$$\tilde{\boldsymbol{v}}(n,t) := (\widetilde{|f|^p}(n,t), \widetilde{|g|^q}(n,t), \tilde{f}(n,t), \tilde{g}(n,t))$$

and assume that $\tilde{\boldsymbol{v}}(n,t)$ and its neighbors $\tilde{\boldsymbol{v}}(n \pm e_i, t)$ lie in the domain K of $B_{\varepsilon,p,K}$. Then

$$(\partial_t - \Delta)(B_{\varepsilon,p,K} \circ \tilde{\boldsymbol{v}})(n,t) \geqslant \textstyle\sum_i (|\partial^i_+ \tilde{f}||\partial^i_+ \tilde{g}| + |\partial^i_- \tilde{f}||\partial^i_- \tilde{g}|). \tag{7}$$

Proof. We will write $B = B_{\varepsilon,p,K}$. Taylor's theorem and the smoothness of B allow us to write,

$$B(\boldsymbol{v}+\delta\boldsymbol{v}) = B(\boldsymbol{v}) + \nabla_{\boldsymbol{v}} B(\boldsymbol{v})\delta\boldsymbol{v} + \int_0^1 (1-s)(\mathrm{d}^2_{\boldsymbol{v}} B(\boldsymbol{v}+s\delta\boldsymbol{v})\delta\boldsymbol{v},\delta\boldsymbol{v})\mathrm{d}s.$$

In particular, for any $1 \leqslant i \leqslant N$ we have for fixed (n,t) with $\delta_i^{\pm}\tilde{\boldsymbol{v}}$ such that $\tilde{\boldsymbol{v}} + \delta_i^{\pm}\tilde{\boldsymbol{v}} = \tilde{\boldsymbol{v}}(\cdot \pm e_i, \cdot)$

$$\begin{aligned} B(\tilde{\boldsymbol{v}}+\delta_i^{\pm}\tilde{\boldsymbol{v}}) = {} & B(\tilde{\boldsymbol{v}}) + \nabla_{\boldsymbol{v}} B(\tilde{\boldsymbol{v}})\delta_i^{\pm}\tilde{\boldsymbol{v}} \\ & + \int_0^1 (1-s)(\mathrm{d}^2_{\boldsymbol{v}} B(\tilde{\boldsymbol{v}}+s\delta_i^{\pm}\tilde{\boldsymbol{v}})\delta_i^{\pm}\tilde{\boldsymbol{v}},\delta_i^{\pm}\tilde{\boldsymbol{v}})\mathrm{d}s. \end{aligned}$$

and summing over all i and all $\pm$'s, we have

$$\begin{aligned} \Delta(B\circ\tilde{\boldsymbol{v}}) = {} & \sum_{i,\pm} B(\tilde{\boldsymbol{v}}+\delta_i^{\pm}\tilde{\boldsymbol{v}}) - B(\tilde{\boldsymbol{v}}) \\ = {} & \sum_{i,\pm} \nabla_{\boldsymbol{v}} B(\tilde{\boldsymbol{v}})\delta_i^{\pm}\tilde{\boldsymbol{v}} + \sum_{i,\pm}\int_0^1 (1-s)(\mathrm{d}^2_{\boldsymbol{v}} B(\tilde{\boldsymbol{v}}+s\delta_i^{\pm}\tilde{\boldsymbol{v}})\delta_i^{\pm}\tilde{\boldsymbol{v}},\delta_i^{\pm}\tilde{\boldsymbol{v}})\mathrm{d}s \\ = {} & \nabla_{\boldsymbol{v}} B(\tilde{\boldsymbol{v}})\Delta\tilde{\boldsymbol{v}} + \sum_{i,\pm}\int_0^1 (1-s)(\mathrm{d}^2_{\boldsymbol{v}} B(\tilde{\boldsymbol{v}}+s\delta_i^{\pm}\tilde{\boldsymbol{v}})\delta_i^{\pm}\tilde{\boldsymbol{v}},\delta_i^{\pm}\tilde{\boldsymbol{v}})\mathrm{d}s \end{aligned}$$

On the other hand, the boundedness of $\tilde{\boldsymbol{v}}$ together with the fact that the discrete laplacian is a bounded operator from $L^\infty(G^N)$ into itself easily imply that $\tilde{\boldsymbol{v}}$ is continuously differentiable w.r.t. the variable t, and we have,

$$\frac{\partial}{\partial t}(B\circ\tilde{\boldsymbol{v}}) = \nabla_{\boldsymbol{v}} B(\tilde{\boldsymbol{v}})\frac{\partial\tilde{\boldsymbol{v}}}{\partial t}.$$

Now $\tilde{\boldsymbol{v}}(n,t)$ is a vector consisting of solutions to the semi-continuous heat equation. This means we have $(\partial_t - \Delta)\tilde{\boldsymbol{v}} = 0$, that is

$$\partial_t\tilde{\boldsymbol{v}}(n,t) - \sum_i (\tilde{\boldsymbol{v}}(n+e_i,t) - 2\tilde{\boldsymbol{v}}(n,t) + \tilde{\boldsymbol{v}}(n-e_i,t)) = 0.$$

The difference of the last two equations reads

$$\begin{aligned} \left(\frac{\partial}{\partial t} - \Delta\right)(B\circ\tilde{\boldsymbol{v}}) = {} & \nabla_{\boldsymbol{v}} B(\tilde{\boldsymbol{v}})(\partial_t - \Delta)\tilde{\boldsymbol{v}} \\ & + \sum_{i,\pm}\int_0^1 (1-s)(-\mathrm{d}^2_{\boldsymbol{v}} B(\tilde{\boldsymbol{v}}+s\delta_i^{\pm}\tilde{\boldsymbol{v}})\delta_i^{\pm}\tilde{\boldsymbol{v}},\delta_i^{\pm}\tilde{\boldsymbol{v}})\mathrm{d}s. \end{aligned}$$

Notice that the domain of B is convex. Recall that (n,t) is such that both $\tilde{\boldsymbol{v}}$ and $\tilde{\boldsymbol{v}} + \delta^i_{\pm}\tilde{\boldsymbol{v}} = \tilde{\boldsymbol{v}}(\cdot \pm e_i, \cdot)$ at (n,t) are contained in the domain of B. Thanks to convexity, so is $\tilde{\boldsymbol{v}} + s\delta^i_{+}\tilde{\boldsymbol{v}}$ for $0 \leqslant s \leqslant 1$. Together with the universal inequality $-\mathrm{d}^2_{\boldsymbol{v}} B(\boldsymbol{v}) \geqslant 2|\mathrm{d}\boldsymbol{f}||\mathrm{d}\boldsymbol{g}|$ in the domain of B, we finally obtain the estimate

$$(\partial_t - \Delta)(B \circ \tilde{\boldsymbol{v}}) \geqslant \sum_i (|\partial_+^i \tilde{f}||\partial_+^i \tilde{g}| + |\partial_-^i \tilde{f}||\partial_-^i \tilde{g}|).$$

after observing that $\delta_\pm = \pm\partial_\pm$. *QED*

3.3 Bilinear embedding

We are going to prove the following stated as Theorem 3 in the Introduction:

Theorem 7 *Let f and g be test functions on G^N and p and q conjugate exponents. Then we have the estimate*

$$2\sum_{i=1}^{N} \int_0^\infty \sum_{G^N} |\partial_+^i \tilde{f}(n,t)||\partial_+^i \tilde{g}(n,t)| \mathrm{d}t \leqslant (p^* - 1)\|f\|_p \|g\|_q.$$

Proof. We write for simplicity $Q = G^N$. Notice first that for any $1 \leqslant p \leqslant \infty$, the discrete character of Q ensures that $L^p(Q) \subset L^\infty(Q)$, namely $\|f\|_\infty \leqslant \|f\|_p$. From the elementary properties (2) of the heat kernel, the heat semi-group is a contraction in $L^\infty(Q)$. As a conclusion, for all $1 < p < \infty$, for all $f \in L^p(Q)$, we have for all $t \geqslant 0$, $\|\tilde{f}(t)\|_\infty \leqslant \|f\|_\infty \leqslant \|f\|_p$. Similar observations hold for the functions $\tilde{g}$, $\widetilde{|f|^p}$ and $\widetilde{|g|^q}$. Setting as in Lemma 3

$$\tilde{\boldsymbol{v}}(n,t) := (\widetilde{|f|^p}(n,t), \widetilde{|g|^q}(n,t), \tilde{f}(n,t), \tilde{g}(n,t)),$$

we have that $\tilde{\boldsymbol{v}}(Q \times [0,\infty))$ lies in a compact $K \subset D_p$, the domain of the Bellman function. Pick a point $(n,t) \in Q \times (0,\infty)$ and set $b(n,t) = B_{\varepsilon,p,K} \circ \tilde{\boldsymbol{v}}(n,t)$. Invoking Theorem 5, we have therefore on the one hand the upper estimate $\forall (n,t)$

$$(B \circ \tilde{\boldsymbol{v}})(n,t) := b(n,t) \leqslant (1 + \varepsilon C_K)(p^* - 1)(\widetilde{|f|^p}(n,t))^{1/p}(\widetilde{|g|^q}(n,t))^{1/q},$$

or in a condensed manner as an inequality involving functions,

$$\forall t, \quad (B \circ \tilde{\boldsymbol{v}})(t) := b(t) \leqslant (1 + \varepsilon C_K)(p^* - 1)(\widetilde{|f|^p}(t))^{1/p}(\widetilde{|g|^q}(t))^{1/q}.$$

On the other hand, for a lower estimate, we wish to make explicit the dependence of $b(t)$ with respect to the dissipation estimate (7). Introduce now the function $\mathcal{D}$ defined on $Q \times [0,t]$ as

$$\forall 0 \leqslant s \leqslant t, \quad \mathcal{D}(s) := b(s) * K(t - s)$$

so that $\mathcal{D}(t) = b(t) = (B \circ \boldsymbol{v})(t)$. We have, using standard discrete integration by parts that are allowed thanks to the exponential decay of the kernel,

$$\begin{aligned}\mathcal{D}'(s) &= (\partial_t b)(s) * K(t-s) - b(s) * (\partial_t K)(t-s)\\ &= (\partial_t b)(s) * K(t-s) - b(s) * (\Delta K)(t-s)\\ &= ((\partial_t - \Delta) b)(s) * K(t-s),\end{aligned}$$

from which follows

$$\begin{aligned}b(t) &= b(0) * K(t) + \int_0^t ((\partial_t - \Delta) b)(s) * K(t-s) \mathrm{d}s\\ &\geqslant \int_0^t \Gamma(\tilde{f}, \tilde{g})(s) * K(t-s) \mathrm{d}s.\end{aligned}$$

We used that $K \geqslant 0$, $b \geqslant 0$, and the dissipation estimate (7) with $\Gamma(\tilde{f}, \tilde{g})$ as a shorthand for:

$$\Gamma(\tilde{f}, \tilde{g}) := \sum_{i=1}^{N} |\partial_+^i \tilde{f}||\partial_+^i \tilde{g}| + |\partial_-^i \tilde{f}||\partial_-^i \tilde{g}|$$

Summarizing, we have compared the functions $\forall t$,

$$\begin{aligned}\int_0^t \Gamma(\tilde{f}, \tilde{g})(s) * K(t-s) \mathrm{d}s &\leqslant b(t)\\ &\leqslant (1 + \varepsilon C_K)(p^* - 1)(\widetilde{|f|^p}(t))^{1/p}(\widetilde{|g|^q}(t))^{1/q}.\end{aligned}$$

Hölder's inequality ensures that the right hand side is integrable over $Q := G^N$ uniformly in t. $\forall t$,

$$\begin{aligned}\sum_{n \in Q} (\widetilde{|f|^p}(n,t))^{1/p}(\widetilde{|g|^q}(n,t))^{1/q} &\leqslant \left(\sum_{n \in Q} \widetilde{|f|^p}(n,t)\right)^{1/p} \left(\sum_{n \in Q} \widetilde{|g|^q}(n,t)\right)^{1/q}\\ &\leqslant (\|\widetilde{|f|^p}(t)\|_1)^{1/p}(\|\widetilde{|g|^q}(t)\|_1)^{1/q}\\ &\leqslant (\| \, |f|^p \, \|_1)^{1/p} (\| \, |g|^q \, \|_1)^{1/q}\\ &\leqslant \|f\|_p \|g\|_q < \infty,\end{aligned}$$

where we used that the heat semi-group is a contraction in L^1 (actually preserves norms of nonnegative functions). For the left hand side, integrating also on $Q := G^N$ and using Fubini for absolutely converging series, we have

$$\sum_{n\in Q}\int_0^t (\Gamma(\tilde f,\tilde g)(s)*K(t-s))(n)\mathrm{d}s$$
$$=\sum_{n\in Q}\int_0^t \sum_{m\in Q}\Gamma(\tilde f,\tilde g)(m,s)K(n-m,t-s)\mathrm{d}s$$
$$=\int_0^t \sum_{m\in Q}\Gamma(\tilde f,\tilde g)(m,s)\left(\sum_{n\in Q}K(n-m,t-s)\right)\mathrm{d}s$$
$$=\int_0^t \sum_{m\in Q}\Gamma(\tilde f,\tilde g)(m,s)\mathrm{d}s$$

Letting finally ε go to 0 and t go to infinity yields the result. *QED*

3.4 Proof of Theorem 1

The identity formulae for second order Riesz transforms allow us to deduce the estimate from the bilinear embedding.

Proof. (of Theorem 1) Remember that

$$R_\alpha^2 := \sum_{i=1}^N \alpha_i R_i^2 = \sum_{i=1}^N \alpha_i R_+^i R_-^i, \qquad \alpha_i\in\mathbb{C},\quad |\alpha_i|\leqslant 1,\quad \forall i\in[1,\ldots,N].$$

Using successively the representation formula (3) of Lemma 1 and the bilinear estimate (1) of Theorem 3, we have

$$|(f,R_\alpha^2 g)| := \left|\sum_i \alpha_i \int_0^\infty \sum_{G^N} \partial_+^i\tilde f(n,t)\overline{\partial_+^i\tilde g(n,t)} + \partial_-^i\tilde f(n,t)\overline{\partial_-^i\tilde g(n,t)}\mathrm{d}t\right|$$
$$\leqslant \sum_i |\alpha_i| \int_0^\infty \sum_{G^N} |\partial_+^i\tilde f(n,t)||\partial_+^i\tilde g(n,t)| + |\partial_-^i\tilde f(n,t)||\partial_-^i\tilde g(n,t)|\mathrm{d}t$$
$$\leqslant (p^*-1)\|f\|_p\|g\|_q$$

which proves the result. *QED*

4 The Choi constant

For some multipliers, the best constant is better than p^*-1. The simplest example is when $\alpha\equiv 1$, in which case we estimate the identity. When considered as an operator on real valued functions only, the constant p^*-1 is not necessarily best possible also for some non-trivial cases. The estimates of multipliers with real non-negative coefficients will depend upon the relative distance of the largest and the smallest multiplier. Instead of Burkholder's result, we use a theorem by Choi [7].

Theorem 8 *(Choi) Let $1 < p < \infty$ and let d_k be a real valued martingale difference sequence and let θ_k be a predictable sequence taking values in $\{0,1\}$. The best constant in the inequality*

$$\left\| \sum_{k=1}^{n} \theta_k \mathrm{d}_k \right\|_p \leqslant \mathfrak{C}_p \left\| \sum_{k=1}^{n} \mathrm{d}_k \right\|_p$$

that holds for all n is the Choi constant $\mathfrak{C}_p$.

It is a stronger hypothesis than differential subordination. We transform this martingale estimate into the existence of an appropriate Bellman function. We prove now

Lemma 4 *For every $1 < p < \infty$ there exists a function C_p^+ of four variables $\boldsymbol{v} = (\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{f}, \boldsymbol{g})$ in the domain $\overline{D_p} = \{(\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{f}, \boldsymbol{g}) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} : |\boldsymbol{f}|^p \leqslant \boldsymbol{F}, |\boldsymbol{g}|^q \leqslant \boldsymbol{G}\}$ so that*

$$0 \leqslant C_p^+(\boldsymbol{v}) \leqslant \mathfrak{C}_p \boldsymbol{F}^{1/p} \boldsymbol{G}^{1/q},$$

and

$$C_p^+(\boldsymbol{v}) \geqslant \frac{1}{2} C_p^+(\boldsymbol{v}_+) + \frac{1}{2} C_p^+(\boldsymbol{v}_-) + \frac{1}{4}[(\boldsymbol{f}_+ - \boldsymbol{f}_-)(\boldsymbol{g}_+ - \boldsymbol{g}_-)]_+$$

if $\boldsymbol{v}_+ + \boldsymbol{v}_- = 2\boldsymbol{v}$ and $\boldsymbol{v}$, $\boldsymbol{v}_+$, $\boldsymbol{v}_-$ are in the domain. Here $[\,\cdot\,]_+ := \max\{\cdot, 0\}$ denotes non-negative part.

Proof. In all the sequel, we consider the case where θ_k takes values in $\{0,1\}$. It will be clear from the context how to handle the general case. Similar to above, this implies that the operator

$$T_c : f \mapsto \sum_{I \in \mathcal{D}(J)} c_I (f, h_I) h_I$$

where $c = \{c_I\}_{I \in \mathcal{D}} \in \{0,1\}^{\mathcal{D}}$, has L^p bound, uniformly in the choice of the sequence c, and has norm at most $\mathfrak{C}_p$. By duality this becomes

$$\sup_c |(T_c f, g)| \leqslant \mathfrak{C}_p \|f\|_p \|g\|_q.$$

Looking at the left hand side, the supremum is attained, for a given pair of test functions f, g when one omits either the positive or the negative terms of the sum. That is either the sequence

$$c = \{c_I : \quad c_I = 1 \quad \text{for} \quad (f, h_I)(g, h_I) \geqslant 0, \quad c_I = 0 \quad \text{otherwise}\}$$

or the sequence

$$c = \{c_I : \quad c_I = 1 \quad \text{for} \quad (f, h_I)(g, h_I) \leqslant 0, \quad c_I = 0 \quad \text{otherwise}\}$$

That is, the supremum writes as

$$|(T_c f, g)| = \sum_I [(f, h_I)(g, h_I)]_+ \text{ or } |(T_c f, g)| = \sum_I [(f, h_I)(g, h_I)]_-.$$

Here, $[\,\cdot\,]_\pm$ denotes the positive or negative part of the sequence. We deal now with the estimate of the positive part. Let us define the following function

$$C_p^+(\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{f}, \boldsymbol{g}) = \sup_{f,g} \left\{ \frac{1}{4|J|} \sum_{I \in \mathcal{D}(J)} |I| [(\langle f \rangle_{I_+} - \langle f \rangle_{I_-})(\langle g \rangle_{I_+} - \langle g \rangle_{I_-})]_+ : \right.$$
$$\left. \langle f \rangle_J = \boldsymbol{f}, \langle g \rangle_J = \boldsymbol{g}, \langle |f|^p \rangle_J = \boldsymbol{F}, \langle |g|^q \rangle_J = \boldsymbol{G} \right\}.$$

The supremum runs over functions supported in J,but scaling shows that the function C_p^+ itself does not depend upon the interval J. The domain of C_p^+ is chosen so that every set of numbers $\boldsymbol{v} = (\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{f}, \boldsymbol{g})$ with $\boldsymbol{F}, \boldsymbol{G} \geqslant 0, |\boldsymbol{f}|^p \leqslant \boldsymbol{F}, |\boldsymbol{g}|^q \leqslant \boldsymbol{G}$ can be associated with real valued functions f, g so that

$$\langle f \rangle_J = \boldsymbol{f}, \langle g \rangle_J = \boldsymbol{g}, \langle |f|^p \rangle_J = \boldsymbol{F}, \langle |g|^q \rangle_J = \boldsymbol{G}.$$

As discussed before, we know that

$$0 \leqslant C_p^+(\boldsymbol{v}) \leqslant \mathfrak{C}_p \boldsymbol{F}^{1/p} \boldsymbol{G}^{1/q}.$$

We will now investigate the contributions of $J_\pm$ to the above sum and derive a concavity condition for C_p^+. Choose as above $\boldsymbol{v}_\pm$ in the domain so that $\boldsymbol{v}_+ + \boldsymbol{v}_- = 2\boldsymbol{v}$. Then

$$
\begin{aligned}
&C_p^+(\boldsymbol{F},\boldsymbol{G},\boldsymbol{f},\boldsymbol{g})\\
&\quad\geqslant \sup_{f,g}\left\{\frac{1}{4|J|}\sum_{I\in\mathcal{D}(J)}|I|[(\langle f\rangle_{I_+}-\langle f\rangle_{I_-})(\langle g\rangle_{I_+}-\langle g\rangle_{I_-})]_+ :\right.\\
&\qquad\qquad\left.\langle f\rangle_{J_\pm}=\boldsymbol{f}_\pm,\langle g\rangle_{J_\pm}=\boldsymbol{g}_\pm,\langle |f|^p\rangle_{J_\pm}=\boldsymbol{F}_\pm,\langle |g|^q\rangle_{J\pm}=\boldsymbol{G}_\pm\right\}\\
&\quad\geqslant \sup_{f,g}\left\{\frac{1}{8|J|}\sum_{I\in\mathcal{D}(J_+)}|I|[(\langle f\rangle_{I_+}-\langle f\rangle_{I_-})(\langle g\rangle_{I_+}-\langle g\rangle_{I_-})]_+\right.\\
&\qquad\qquad+\frac{1}{8|J|}\sum_{I\in\mathcal{D}(J_-)}|I|[(\langle f\rangle_{I_+}-\langle f\rangle_{I_-})(\langle g\rangle_{I_+}-\langle g\rangle_{I_-})]_+\\
&\qquad\qquad+\frac{1}{4}[(\langle f\rangle_{I_+}-\langle f\rangle_{I_-})(\langle g\rangle_{I_+}-\langle g\rangle_{I_-})]_+ :\\
&\qquad\qquad\left.\langle f\rangle_{J_\pm}=\boldsymbol{f}_\pm,\langle g\rangle_{J_\pm}=\boldsymbol{g}_\pm,\langle |f|^p\rangle_{J_\pm}=\boldsymbol{F}_\pm,\langle |g|^q\rangle_{J\pm}=\boldsymbol{G}_\pm\right\}\\
&\quad\geqslant \frac{1}{2}C_p^+(\boldsymbol{F}_+,\boldsymbol{G}_+,\boldsymbol{f}_+,\boldsymbol{g}_+)+\frac{1}{2}C_p^+(\boldsymbol{F}_-,\boldsymbol{G}_-,\boldsymbol{f}_-,\boldsymbol{g}_-)\\
&\qquad\qquad+\frac{1}{4}[(\boldsymbol{f}_+-\boldsymbol{f}_-)(\boldsymbol{g}_+-\boldsymbol{g}_-)]_+.
\end{aligned}
$$

This proves Lemma 4. *QED*

An infinitesimal version of Lemma 4, omitting smoothing parameters in the name of the function andfollowing the same considerations as in the proof of the range and Hessian estimate (6) of Theorem 5, we have

$$0\leqslant C_p^+(\boldsymbol{v})\leqslant(1+\varepsilon C_K)\mathfrak{C}_p\boldsymbol{F}^{1/p}\boldsymbol{G}^{1/q},$$

and

$$-\mathrm{d}^2_{\boldsymbol{v}}C_p^+(\boldsymbol{v})\geqslant(1+\operatorname{sign}(\mathrm{d}\boldsymbol{f}\mathrm{d}\boldsymbol{g}))\mathrm{d}\boldsymbol{f}\mathrm{d}\boldsymbol{g}.$$

Again, by Taylor's theorem

$$
\begin{aligned}
C_p^+(\boldsymbol{v}+\delta\boldsymbol{v})=&\,C_p^+(\boldsymbol{v})+\nabla_{\boldsymbol{v}}C_p^+(\boldsymbol{v})\delta\boldsymbol{v}\\
&+\int_0^1(1-s)(\mathrm{d}^2_{\boldsymbol{v}}C_p^+(\boldsymbol{v}+s\delta\boldsymbol{v})\delta\boldsymbol{v},\delta\boldsymbol{v})ds.
\end{aligned}
$$

Define as before the vector $\tilde{\boldsymbol{v}}(n,t):=(\widetilde{|f|^p}(n,t),\widetilde{|g|^q}(n,t),\tilde{f}(n,t),\tilde{g}(n,t))$. Letting for fixed (n,t), $\delta^i_\pm\tilde{\boldsymbol{v}}=\pm\partial^i_\pm\tilde{\boldsymbol{v}}=\tilde{\boldsymbol{v}}(\cdot\pm e_i,\cdot)-\tilde{\boldsymbol{v}}$, we have

$$C_p^+(\tilde{\boldsymbol{v}} + \delta_\pm^i \tilde{\boldsymbol{v}}) = C_p^+(\tilde{\boldsymbol{v}}) + \nabla_{\boldsymbol{v}} C_p^+(\tilde{\boldsymbol{v}}) \cdot \delta_\pm^i \tilde{\boldsymbol{v}} + \int_0^1 (1-s)(\mathrm{d}_{\boldsymbol{v}}^2 C_p^+(\tilde{\boldsymbol{v}} + s\delta_\pm^i \tilde{\boldsymbol{v}})\delta_\pm^i \tilde{\boldsymbol{v}}, \delta_\pm^i \tilde{\boldsymbol{v}})\mathrm{d}s,$$

as well as

$$\partial_t C_p^+(\tilde{\boldsymbol{v}}(n,t)) = \nabla_{\boldsymbol{v}} C_p^+(\tilde{\boldsymbol{v}}(n,t))\partial_t \tilde{\boldsymbol{v}}(n,t).$$

Summing over i and $\pm$ and using that $\tilde{\boldsymbol{v}}$ solves the heat equation, we obtain

$$(\partial_t - \Delta)C_p^+(\tilde{\boldsymbol{v}}(n,t)) = \sum_{i,\pm} \int_0^1 (1-s)\{(\mathrm{d}_{\boldsymbol{v}}^2 C_p^+(\tilde{\boldsymbol{v}} + s\delta_\pm^i \tilde{\boldsymbol{v}})\delta_\pm^i \tilde{\boldsymbol{v}}, \delta_\pm^i \tilde{\boldsymbol{v}})\}\mathrm{d}s.$$

Notice that the domain of C_p^+ is convex. Both $\tilde{\boldsymbol{v}}(n,t)$ and $\tilde{\boldsymbol{v}} + \delta_\pm^i \tilde{\boldsymbol{v}} = \tilde{\boldsymbol{v}}(\cdot \pm e_i, \cdot)$ are contained in the domain of C_p^+, and so is $\tilde{\boldsymbol{v}} + s\delta_+^i \tilde{\boldsymbol{v}}$ for $0 \leqslant s \leqslant 1$. Together with the universal inequality $-\mathrm{d}_{\boldsymbol{v}}^2 C_p^+(\boldsymbol{v}) \geqslant (1 + \mathrm{sign}(\mathrm{d}\boldsymbol{f}\mathrm{d}\boldsymbol{g}))\mathrm{d}\boldsymbol{f}\mathrm{d}\boldsymbol{g}$ in the domain of C_p^+, we obtain the estimate

$$(\partial_t - \Delta)C_p^+(\tilde{\boldsymbol{v}}(n,t)) \geqslant \sum_i \left[\partial_+^i \tilde{f} \;\; \partial_+^i \tilde{g}\right]_+ + \left[\partial_-^i \tilde{f} \;\; \partial_-^i \tilde{g}\right]_+ \tag{8}$$

With the help of these considerations, it is now easy to prove:

Proof. (of Theorem 2 and Theorem 4)

By the identity formula, we see that for $c_i \in \{0,1\}$,

$$\begin{aligned}\left|\left(f, \sum_i c_i R_i^2 g\right)\right| &= 2\left|\int_0^\infty \sum_{\mathbb{Z}^N} \sum_i c_i \partial_+^i \tilde{f}(n,t)\partial_+^i \tilde{g}(n,t)\mathrm{d}t\right| \\ &\leqslant \max_\pm \left\{2\int_0^\infty \sum_{\mathbb{Z}^N} \sum_i [\delta_+^i \tilde{f} \delta_+^i \tilde{g}]_\pm \mathrm{d}t\right\} \\ &\leqslant \mathfrak{C}_p \|f\|_p \|g\|_q.\end{aligned}$$

We used successively the representation formula (3) of Lemma 1 the bilinear embedding of Theorem 4 resulting from the dissipation estimate (8) above for C_p^+ above as well as the similar bilinear embedding that results from C_p^-. This concludes the proof of Theorem 2. *QED*

5 Sharpness.

The sharpness in $\mathbb{Z}^N$is inherited from the continuous case $\mathbb{R}^N$. Just consider the isomorphic groups $(t\mathbb{Z})^N$ for $0 < t \leqslant 1$ in conjunction with the Lax equivalence theorem [12]. By the same argument, sharpness for a *uniform* estimate in m for the cyclic case $(\mathbb{Z}/m\mathbb{Z})^N$ is inherited from that on the torus $\mathbb{T}^N$.

References

[1] Nicola Arcozzi. Riesz transforms on compact Lie groups, spheres and Gauss space. *Ark. Mat.*, 36(2):201–231, 1998.

[2] Rodrigo Bañuelos and Fabrice Baudoin. Martingale Transforms and Their Projection Operators on Manifolds. *Potential Anal.*, 38(4):1071–1089, 2013.

[3] Rodrigo Bañuelos and Gang Wang. Sharp inequalities for martingales with applications to the Beurling-Ahlfors and Riesz transforms. *Duke Math. J.*, 80(3):575–600, 1995.

[4] D. L. Burkholder. Boundary value problems and sharp inequalities for martingale transforms. *Ann. Probab.*, 12(3):647–702, 1984.

[5] Donald L. Burkholder. Sharp inequalities for martingales and stochastic integrals. *Astérisque*, (157-158):75–94, 1988. Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987).

[6] Andrea Carbonaro and Oliver Dragičević. Bellman function and linear dimension-free estimates in a theorem of Bakry. *J. Funct. Anal.*, 265(7):1085–1104, 2013.

[7] K. P. Choi. A sharp inequality for martingale transforms and the unconditional basis constant of a monotone basis in $L^p(0,1)$. *Trans. Amer. Math. Soc.*, 330(2):509–529, 1992.

[8] Thierry Coulhon, Detlef Müller, and Jacek Zienkiewicz. About Riesz transforms on the Heisenberg groups. *Math. Ann.*, 305(2):369–379, 1996.

[9] Matts Essén. A superharmonic proof of the M. Riesz conjugate function theorem. *Ark. Mat.*, 22(2):241–249, 1984.

[10] Stefan Geiss, Stephen Montgomery-Smith, and Eero Saksman. On singular integral and martingale transforms. *Trans. Amer. Math. Soc.*, 362(2):553–575, 2010.

[11] Tadeusz Iwaniec and Gaven Martin. Riesz transforms and related singular integrals. *J. Reine Angew. Math.*, 473:25–57, 1996.

[12] P. D. Lax and R. D. Richtmyer. Survey of the stability of linear finite difference equations. *Comm. Pure Appl. Math.*, 9:267–293, 1956.

[13] Françoise Lust-Piquard. Riesz transforms associated with the number operator on the Walsh system and the fermions. *J. Funct. Anal.*, 155(1):263–285, 1998.

[14] Françoise Lust-Piquard. Dimension free estimates for discrete Riesz transforms on products of abelian groups. *Adv. Math.*, 185(2):289–327, 2004.

[15] P.-A. Meyer. Transformations de Riesz pour les lois gaussiennes. In *Seminar on probability, XVIII*, volume 1059 of *Lecture Notes in Math.*, pages 179–193. Springer, Berlin, 1984.

[16] F. L. Nazarov and S. R. Treil, Sergei. The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis. *Algebra i Analiz*, 8(5):32–162, 1996.

[17] Stefanie Petermichl and Alexander Volberg. Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular. *Duke Math. J.*, 112(2):281–305, 2002.

[18] S. K. Pichorides. On the best values of the constants in the theorems of M. Riesz, Zygmund and Kolmogorov. *Studia Math.*, 44:165–179. (errata insert), 1972. Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, II.

[19] Gilles Pisier. Riesz transforms: a simpler analytic proof of P.-A. Meyer's inequality. In *Séminaire de Probabilités, XXII*, volume 1321 of *Lecture Notes in Math.*, pages 485–501. Springer, Berlin, 1988.

[20] E. M. Stein. Some results in harmonic analysis in $\mathbf{R}^n$, for $n \to \infty$. *Bull. Amer. Math. Soc. (N.S.)*, 9(1):71–73, 1983.

[21] Vasily Vasyunin and Alexander Volberg. Burkholder's function via Monge-Ampère equation. *Illinois J. Math.*, 54(4):1393–1428 (2012), 2010.

[22] A. Volberg and F. Nazarov. Heat extension of the Beurling operator and estimates for its norm. *St. Petersburg Math J.*, 15(4):563–573, 2004.