VARIOUS SHARP ESTIMATES FOR SEMI-DISCRETE RIESZ TRANSFORMS OF THE SECOND ORDER

K. DOMELEVO, A. OSĘKOWSKI, AND S. PETERMICHL

Abstract. We give several sharp estimates for a class of combinations of second order Riesz transforms on Lie groups $\mathbb{G} = \mathbb{G}_x \times \mathbb{G}_y$ that are multiply connected, composed of a discrete abelian component $\mathbb{G}_x$ and a connected component $\mathbb{G}_y$ endowed with a biinvariant measure. These estimates include new sharp L^p estimates via Choi type constants, depending upon the multipliers of the operator. They also include weak-type, logarithmic and exponential estimates. We give an optimal $L^q \to L^p$ estimate as well.

It was shown recently by Arcozzi-Domelevo-Petermichl that such second order Riesz transforms applied to a function may be written as conditional expectation of a simple transformation of a stochastic integral associated with the function.

The proofs of our theorems combine this stochastic integral representation with a number of deep estimates for pairs of martingales under strong differential subordination by Choi, Banuelos and Osękowski.

When two continuous directions are available, sharpness is shown via the laminates technique. We show that sharpness is preserved in the discrete case using Lax-Richtmyer theorem.

1. Introduction

Sharp, classical L^p norm inequalities for pairs of differentially subordinate martingales date back to the celebrated work of Burkholder [15] in 1984 where the optimal constant is exhibited. See also from the same author [17][18]. The relation between differentially subordinate martingales and CZ (i.e. Caldéron–Zygmund) operators is known at least since Gundy–Varopoulos [32]. Banuelos–Wang [12] were the first to exploit this connection to prove new sharp inequalities for singular intergrals. This intersection of probability theory with classical questions in harmonic analysis has lead to much interest and a vast literature has been accumulating on this line of research.

In this article we state a number of sharp estimates that hold in the very recent, new direction concerning the semi–discrete setting, applying it to a family of second order Riesz transforms on multiply–connected Lie groups.

Date: February 12, 2017.

Adam Osekowski is supported by Narodowe Centrum Nauki Poland (NCN), grant DEC-2014/14/E/ST1/00532.

Stefanie Petermichl is supported by ERC project CHRiSHarMa DLV-862402.

We recall their representation through stochastic integrals using jump processes on multiply–connected Lie groups from [3]. In this representation formula jump processes play a role, but the strong differential subordination holds between the martingales representing the test function and the operator applied to the test function.

The usual procedure for obtaining (sharp) inequalities for operators of Calderón–Zygmund type from inequalities for martingales is the following. Starting with a test function f, martingales are built using Brownian motion or background noise and harmonic functions in the upper half space $\mathbb{R}^+ \times \mathbb{R}^n$. Through the use of Itøformula, it is shown that the martingale arising in this way from Rf, where R is a Riesz transform in $\mathbb{R}^n$, is a martingale transform of the martingale arising from f. The two form a pair of martingales that have differential subordination and (in case of Hilbert or Riesz transforms) orthogonality. One then derives sharp martingale inequalities under hypotheses of strong differential subordination (and orthogonality) relations.

In the case of Riesz transforms of the second order, the use of heat extensions in the upper half space instead of Poisson extensions originated in the context of a weighted estimate in Petermichl–Volberg [44] and was used to prove L^p estimates for the second order Riesz transforms based on the results of Burkholder in Nazarov–Volberg [50] as part of their best-at-time estimate for the Beurling–Ahlfors operator, whose real and imaginary parts themselves are second order Riesz transforms. We mention the recent version on discrete abelian groups Domelevo–Petermichl [24] also using a type of heat flow. These proofs are deterministic. The technique of Bellman functions was used. This deterministic strategy does well when no orthogonality is present and when strong subordination is the only important property. Stochastic proofs (aside from giving better estimates in some situations) also have the advantage that once the integral representation is known, the proofs are a very concise consequence of the respective statements on martingales.

In [3] the authors proved sharp L^p estimates for semi–discrete second order Riesz transforms R^2_α using stochastic integrals. There is an array of Riesz transforms of the second order that are treated, indexed my a matrix index α (see below for precisions on acceptable α). The following representation formula of semi-discrete second order Riesz transforms R^2_α *à la* Gundy–Varopoulos (see [32]) is instrumental:

Theorem. ***(Arcozzi–Domelevo–Petermichl, 2016)*** *The second order Riesz transform $R^2_\alpha f$ of a function $f \in L^2(\mathbb{G})$ as defined in (1.1) can be written as the conditional expectation*

$$\mathbb{E}(M_0^{\alpha,f} | \mathcal{Z}_0 = z).$$

Here $M_t^{\alpha,f}$ is a suitable martingale transform of a martingale M_t^f associated to f, and $\mathcal{Z}_t$ is a suitable random walk on $\mathbb{G}$

We remark that the L^p estimates of the discrete Hilbert transform on the integers are still open. It is a famous conjecture that this operator has the same norm as its continuous counterpart.

These known L^p norm inequalities use special functions found in the results of PICHORIDES [45], VERBITSKY[18], ESSÉN [28], BANUELOS–WANG [12] when orthogonality is present in addition to differential subordination or BURKHOLDER [15][16][17], WANG [51] when differential subordination is the only hypothesis.

The aim of the present paper is to establish new estimates for semi–discrete Riesz transforms by using the martingale representation above together with recent martingale inequalities found in the literature.

Here is a brief description of the new results in this paper.

- In the case where the function f is real valued, we can obtain better estimates for R^2_α than in the general case. These estimates depend upon the make of the matrix index α. The precise statement is found in Theorem 1.2.
- We prove a refined sharp weak type estimate using a weak type norm defined just before the statement of Theorem 1.3 .
- We prove logarithmic and exponential estimates, in a sense limiting (in p) cases of the classical sharp L^p estimate. See Theorem 1.4.
- We consider the norm estimates of the $R^2_\alpha : L^q \to L^p$, spaces of different exponent. The statement is found in Theorem 1.5.

1.1. **Differential operators and Riesz transforms.**

First order derivatives and tangent planes. We will consider Lie groups $\mathbb{G} := \mathbb{G}_x \times \mathbb{G}_y$, where $\mathbb{G}_x$ is a discrete abelian group with a fixed set G of m generators, and their reciprocals, and $\mathbb{G}_y$ is a connected, Lie group of dimension n endowed with a biinvariant metric. The choice of the set G of generators in $\mathbb{G}_x$ corresponds to the choice of a bi-invariant metric structure on $\mathbb{G}_x$. We will use on $\mathbb{G}_x$ the multiplicative notation for the group operation. We will define a product metric structure on $\mathbb{G}$, which agrees with the Riemannian structure on the first factor, and with the discrete "word distance" on the second. We will at the same time define a "tangent space" $T_z\mathbb{G}$ for $\mathbb{G}$ at a point $z = (x, y) \in (\mathbb{G}_x \times \mathbb{G}_y) = \mathbb{G}$. We will do this in three steps.

First, since $\mathbb{G}_y$ is an n-dimensional connected Lie group with Lie algebra $\mathfrak{G}_y$. We can identify each left-invariant vector field Y in $\mathfrak{G}_y$ with its value at the identity e, $\mathfrak{G}_y \equiv T_e\mathbb{G}_y$. Since $\mathbb{G}$ is compact, it admits a bi-invariant Riemannian metric, which is unique up to a multiplicative factor. We normalize it so that the measure μ_y associated with the metric satisfies $\mu_y(\mathbb{G}_y) = 1$. The measure μ_y is also the normalized Haar measure of the group. We denote by $< \cdot, \cdot >_y$ be the corresponding inner product on $T_y\mathbb{G}_y$ and by $\boldsymbol{\nabla}_{\boldsymbol{y}} f(y)$ the gradient at $y \in \mathbb{G}_y$ of a smooth function $f : \mathbb{G}_y \to \mathbb{R}$.

Let $Y_1, \ldots, Y_n$ be a orthonormal basis for $\mathfrak{G}_y$. The gradient of f can be written $\boldsymbol{\nabla}_{\boldsymbol{y}} f = Y_1(f)Y_1 + \ldots + Y_n(f)Y_n$.

Second, in the discrete component $\mathbb{G}_x$, let $\mathfrak{G}_x = (g_i)_{i=1,\ldots,m}$ be a set of generators for $\mathbb{G}_x$, such that for $i \neq j$ and $\sigma = \pm 1$ we have $g_i \neq g_j^\sigma$. The choice of a particular set of generators induces a word metric, hence, a geometry, on $\mathbb{G}_x$. Any two sets of generators induce bi-Lipschitz equivalent metrics.

At any point $x \in \mathbb{G}_x$, and given a direction $i \in \{1, \ldots, m\}$, we can define the right and the left derivative at x in the direction i:

$$(\partial^+ f/\partial x_i)(x,y) := f(x+g_i,y) - f(x,y) := (\partial_i^+ f)(x,y)$$

$$(\partial^- f/\partial x_i)(x,y) := f(x,y) - f(x-g_i,y) := (\partial_i^- f)(x,y).$$

Comparing with the continuous component, this suggests that the tangent plane $\hat{T}_x\mathbb{G}_x$ at a point x of the discrete group $\mathbb{G}_x$ might actually be split into a "right" tangent plane $T_x^+\mathbb{G}_x$ and a "left" tangent plane $T_x^-\mathbb{G}_x$, according to the direction with respect to which discrete differences are computed. We consequently define the **augmented** discrete gradient $\widehat{\boldsymbol{\nabla}}_x f(x)$, with a *hat*, as the $2m$–vector of $\hat{T}_x\mathbb{G}_x := T_x^+\mathbb{G}_x \oplus T_x^-\mathbb{G}_x$ accounting for all the local variations of the function f in the direct vicinity of x; that is, the $2m$–column–vector

$$\widehat{\boldsymbol{\nabla}}_x f(x) := (X_1^+ f, X_2^+ f, \ldots, X_1^- f, X_2^- f, \ldots)(x) = \sum_{i=1}^m \sum_{\tau=\pm} X_i^\tau f(x)$$

with $X_i^\tau \in \hat{T}_x\mathbb{G}_x$, where we noted the discrete derivatives $X_i^\pm f := \partial_i^\pm f$ and introduced the discrete $2m$–vectors $X_i^\pm$ as the column vectors of $\mathbb{Z}^{2m}$

$$X_i^+ = (0, \ldots, 1, \ldots, 0) \times \mathbf{0}_m, \quad X_i^- = \mathbf{0}_m \times (0, \ldots, 1, \ldots, 0).$$

Here the 1's in $X_i^\pm$ are located at the i–th position of respectively the first or the second m–tuple. Notice that those vectors are independent of the point x. The scalar product on $\hat{T}_x\mathbb{G}_x := T_x^+\mathbb{G}_x \oplus T_x^-\mathbb{G}_x$ is defined as

$$(U,V)_{\hat{T}_x\mathbb{G}_x} := \frac{1}{2}\sum_{i=1}^m \sum_{\tau=\pm} U_i^\tau V_i^\tau.$$

We chose to put a factor $\frac{1}{2}$ in front of the scalar product to compensate for the fact that we consider both left and right differences.

Finally, for a function f defined on the cartesian product $\mathbb{G} := \mathbb{G}_x \times \mathbb{G}_y$, the (augmented) gradient $\widehat{\boldsymbol{\nabla}}_z f(z)$ at the point $z = (x,y)$ is an element of the tangent plane $\hat{T}_z\mathbb{G} := \hat{T}_x\mathbb{G}_x \oplus T_y\mathbb{G}_y$, that is a $(2m+n)$–column–vector

$$\begin{aligned}\widehat{\boldsymbol{\nabla}}_z f(z) &:= \sum_{i=1}^m \sum_{\tau=\pm} X_i^\tau f(z)\hat{X}_i^\tau + \sum_{j=1}^n Y_j f(z)\hat{Y}_j(z) \\ &= (X_1^+ f, X_2^+ f, \ldots, X_1^- f, X_2^- f, \ldots, Y_1 f, Y_2 f, \ldots)(z)\end{aligned}$$

where $\hat{X}_i^\tau$ and $\hat{Y}_j(z)$ can be identified with column vectors of size $(2m+n)$ with obvious definitions and scalar product $(\cdot,\cdot)_{\hat{T}_z\mathbb{G}_z}$.

Let $\mathrm{d}\mu_z := \mathrm{d}\mu_x \mathrm{d}\mu_y$, $\mathrm{d}\mu_x$ being the counting measure on $\mathbb{G}_x$ and $\mathrm{d}\mu_y$ being the Haar measure on $\mathbb{G}_y$. The inner product of φ, ψ in $L^2(\mathbb{G})$ is

$$(\varphi,\psi)_{L^2(\mathbb{G})} \quad := \quad \int_{\mathbb{G}} \varphi(z)\psi(z) d\mu_z(z).$$

Finally, we make the following hypotheses

Hypothesis. We assume everywhere in the sequel:

1. The discrete component $\mathbb{G}_x$ of the Lie group $\mathbb{G}$ is an abelian group
2. The connected component $\mathbb{G}_y$ of the Lie group $\mathbb{G}$ is a Lie group that can be endowed with a biinvariant Riemannian metric, so that the family $(Y_j)_{j=1,\dots,n}$ commutes with Δ_y.

Notice that this includes compact Lie groups $\mathbb{G}_y$ since those can be endowed with a biinvariant metric. It also includes the usual Euclidian spaces since those are commutative.

Riesz transforms. Following [1][2], recall first that for a compact Riemannian manifold $\mathbb{M}$ without boundary, one denotes by $\boldsymbol{\nabla}_{\mathbb{M}}$, $\mathrm{div}_{\mathbb{M}}$ and $\Delta_{\mathbb{M}} := \mathrm{div}_{\mathbb{M}} \boldsymbol{\nabla}_{\mathbb{M}}$ respectively the gradient, the divergence and the Laplacian associated with $\mathbb{M}$. Then $-\Delta_{\mathbb{M}}$ is a positive operator and the vector Riesz transform is defined as the linear operator

$$\boldsymbol{R}_{\mathbb{M}} := \boldsymbol{\nabla}_{\mathbb{M}} \circ (-\Delta_{\mathbb{M}})^{-1/2}$$

acting on $L^2_0(\mathbb{M})$ (L^2 functions with vanishing mean). It follows that if f is a function defined on $\mathbb{M}$ and $y \in \mathbb{M}$ then $\boldsymbol{R}_{\mathbb{M}} f(y)$ is a vector of the tangent plane $T_y\mathbb{M}$.

Similarly on $\mathbb{M} = \mathbb{G}$, we define $\boldsymbol{\nabla}_{\mathbb{G}} := \widehat{\boldsymbol{\nabla}}_z$ as before, and then we define the divergence operator as its formal adjoint, that is $-\mathrm{div}_{\mathbb{G}} = -\widehat{\mathrm{div}}_z := \widehat{\boldsymbol{\nabla}}^*_z$, with respect to the natural L^2 inner product of vector fields:

$$(U,V)_{L^2(\hat{T}\mathbb{G})} \quad := \quad \int_{\mathbb{G}} (U(z),V(z))_{\hat{T}_z\mathbb{G}} \, \mathrm{d}\mu_z(z)$$

We have the L^2-adjoints $(X_i^{\pm})^* = -X_i^{\mp}$ and $Y_j^* = -Y_j$. If $U \in \hat{T}\mathbb{G}$ is defined by

$$U(z) = \sum_{i=1}^m \sum_{\tau=\pm} U_i^\tau(z)\hat{X}_i^\tau + \sum_{j=1}^n U_j(z)\hat{Y}_j,$$

we define its divergence $\widehat{\boldsymbol{\nabla}}^*_z U$ as

$$\widehat{\boldsymbol{\nabla}}^*_z U(z) \quad := \quad -\frac{1}{2}\sum_{i=1}^m \sum_{\tau=\pm} X_i^{-\tau} U_i^\tau(z) - \sum_{j=1}^n Y_j U_j(z).$$

The Laplacian $\Delta_{\mathbb{G}}$ is as one might expect:

$$
\begin{aligned}
\Delta_z f(z) &:= -\widehat{\boldsymbol{\nabla}}_z^* \widehat{\boldsymbol{\nabla}}_z f(z) = -\widehat{\boldsymbol{\nabla}}_x^* \widehat{\boldsymbol{\nabla}}_x f(z) - \widehat{\boldsymbol{\nabla}}_y^* \widehat{\boldsymbol{\nabla}}_y f(z) \\
&= \sum_{i=1}^m X_i^- X_i^+ f(z) + \sum_{j=1}^n Y_j^2 f(z) \\
&= \sum_{i=1}^m X_i^2 f(z) + \sum_{j=1}^n Y_j^2 f(z) \\
&=: \Delta_x f(z) + \Delta_y f(z)
\end{aligned}
$$

where we denoted $X_i^2 := X_i^+ X_i^- = X_i^- X_i^+$. We have chosen signs so that $-\Delta_{\mathbb{G}} \geqslant 0$ as an operator. The Riesz vector $(\hat{\boldsymbol{R}}_z f)(z)$ is the $(2m+n)$–column–vector of the tangent plane $\hat{T}_z\mathbb{G}$ defined as the linear operator

$$\hat{\boldsymbol{R}}_z f := \left(\widehat{\boldsymbol{\nabla}}_z f\right) \circ (-\Delta_z f)^{-1/2}$$

We also define transforms along the coordinate directions:

$$R_i^{\pm} = X_i^{\pm} \circ (-\Delta_z)^{-1/2} \quad \text{and} \quad R_j = Y_j \circ (-\Delta_z)^{-1/2}.$$

Plan of the paper. In the next two sections, we present successively the main results of the paper and recall the weak formulations involving second order Riesz transforms and semi-discrete heat extensions. Section 2 introduces the stochastic setting for our problems. This includes in Subsection 2.1 semi-discrete random walks, martingale transforms and quadratic covariations. Subsection 2.2 presents a set of martingale inequalities already known in the literature. Finally, in Section 3 we give the proof of the main results.

1.2. **Main results.** In this text, we are concerned with second order Riesz transforms and combinations thereof. We first define the square Riesz transform in the (discrete) direction i to be

$$R_i^2 := R_i^+ R_i^- = R_i^- R_i^+.$$

Then, given $\alpha := ((\alpha_i^x)_{i=1\ldots m}, (\alpha_{jk}^y)_{j,k=1\ldots n}) \in \mathbb{C}^m \times \mathbb{C}^{n\times n}$, we define R_α^2 to be the following combination of second order Riesz transforms:

$$R_\alpha^2 := \sum_{i=1}^m \alpha_i^x \; R_i^2 + \sum_{j,k=1}^n \alpha_{jk}^y \; R_j R_k, \tag{1.1}$$

where the first sum involves squares of discrete Riesz transforms as defined above, and the second sum involves products of continuous Riesz transforms. This combination is written in a condensed manner as the quadratic form

$$R_\alpha^2 = \left(\hat{\boldsymbol{R}}_z, \boldsymbol{A}_\alpha \hat{\boldsymbol{R}}_z\right)$$

where $\boldsymbol{A}_\alpha$ is the $(2m+n)\times(2m+n)$ block matrix

$$\boldsymbol{A}_\alpha := \begin{pmatrix} \boldsymbol{A}_\alpha^x & \mathbf{0} \\ \mathbf{0} & \boldsymbol{A}_\alpha^y \end{pmatrix} \tag{1.2}$$

with

$\boldsymbol{A}_\alpha^x = \mathbf{diag}(\alpha_1^x,\dots,\alpha_m^x,\alpha_1^x,\dots,\alpha_m^x) \in \mathbb{C}^{2m\times 2m}$, $\boldsymbol{A}_\alpha^y = (\alpha_{jk}^y)_{j,k=1\dots n} \in \mathbb{C}^{n\times n}$.

In the theorems below, we assume that $\mathbb{G}$ is a Lie group and R_α^2 is a combination of second order Riesz transforms as defined above. The first application of the stochastic integral formula, Theorem 1.1 was done in [3], while the other applications, Theorems 1.2 1.3 1.4 and 1.5 are new.

Theorem 1.1. ***(Arcozzi–Domelevo–Petermichl, 2016)*** *For any $1 < p < \infty$ we have*

$$\|R_\alpha^2\|_p \leqslant \|\boldsymbol{A}_\alpha\|_2 \ (p^*-1),$$

where, as previously, $p^* = \max\{p, p/(p-1)\}$.

Above, we have set:

$$\|\boldsymbol{A}_\alpha\|_2 = \max\left(\|\boldsymbol{A}_\alpha^x\|_2, \|\boldsymbol{A}_\alpha^y\|_2\right) = \max\left(|\alpha_1^x|,\dots,|\alpha_m^x|, \|\boldsymbol{A}_\alpha^y\|_2\right).$$

In the case where $\mathbb{G} = \mathbb{G}_x$ only consists of the discrete component, this was proved in [25][24] using the deterministic Bellman function technique. In the case where $\mathbb{G} = \mathbb{G}_y$ is a connected compact Lie group, this was proved in [8] using Brownian motions defined on manifolds and projections of martingale transforms.

In the case where the function f is real valued, we can obtain better estimates. For any real numbers $a < b$ and any $1 < p < \infty$, let $\mathfrak{C}_{a,b,p}$ be the constants introduced in Bañuelos and Osękowski [10].

Theorem 1.2. *Assume that $a\boldsymbol{I} \leqslant \boldsymbol{A}_\alpha \leqslant b\boldsymbol{I}$ in the sense of quadratic forms. Then $R_\alpha^2 : L^p(\mathbb{G},\mathbb{R}) \to L^p(\mathbb{G},\mathbb{R})$ enjoys the norm estimate $\|R_\alpha^2\|_p \leqslant \mathfrak{C}_{a,b,p}$.*

We should point out here that the constants $\mathfrak{C}_{a,b,p}$ appear in earlier works of Burkholder [15] (for $a=-b$: then $\mathfrak{C}_{a,b,p} = b(p^*-1)$), and in the paper [19] by Choi (in the case when one of a, b is zero). The Choi constants are not explicit; an approximation of $\mathfrak{C}_{0,1,p}$ is known and writes as

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

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

Coming back to complex-valued functions, we will also establish the following weak-type bounds. We consider the norms

$$||f||_{L^{p,\infty}(\mathbb{G},\mathbb{C})} = \sup\left\{\mu_z(E)^{1/p-1}\int_E f\,d\mu_z\right\},$$

where the supremum is taken over the class of all measurable subsets E of $\mathbb{G}$ of positive measure.

Theorem 1.3. *For any $1 < p < \infty$ we have*

$$\|R_\alpha^2\|_{L^p(\mathbb{G},\mathbb{C})\to L^{p,\infty}(\mathbb{G},\mathbb{C})} \leqslant \|\boldsymbol{A}_\alpha\|_2 \cdot \begin{cases} \left(\frac{1}{2}\Gamma\left(\frac{2p-1}{p-1}\right)\right)^{1-1/p} & \text{if } 1 < p \le 2, \\ \left(\frac{p^{p-1}}{2}\right)^{1/p} & \text{if } p \ge 2. \end{cases}$$

We will also prove the following logarithmic and exponential estimates, which can be regarded as versions of Theorem 1.1 for $p = 1$ and $p = \infty$. Consider the Young functions Φ, $\Psi : [0,\infty) \to [0,\infty)$, given by $\Phi(t) = e^t - 1 - t$ and $\Psi(t) = (t+1)\log(t+1) - t$.

Theorem 1.4. *Let $K > 1$ be fixed.*
(i) For any measurable subset E of $\mathbb{G}$ and any f on $\mathbb{G}$ we have

$$\int_E |R_\alpha^2 f| d\mu_z \leqslant \|\boldsymbol{A}_\alpha\|_2 \cdot \left(K \int_{\mathbb{G}} \Psi(|f|) d\mu_z + \frac{\mu_z(E)}{2(K-1)}\right).$$

(ii) For any $f : \mathbb{G} \to \mathbb{C}$ bounded by 1,

$$\int_{\mathbb{G}} \Phi\left(\frac{|R_\alpha^2 f|}{K\,\|\boldsymbol{A}_\alpha\|_2}\right) d\mu_z \leqslant \frac{||f||_{L^1(\mathbb{G},\mathbb{C})}}{2K(K-1)}.$$

Our final result concerns another extension of Theorem 1.1, which studies the action of R_α^2 between two different L^p spaces. For $1 \le p < q < \infty$, let $C_{p,q}$ be the constant defined by Osękowski in [40].

Theorem 1.5. *For any $1 \le p < q < \infty$, any measurable subset E of $\mathbb{G}$ and any $f \in L^q(\mathbb{G})$ we have*

$$\|R_\alpha^2 f\|_{L^p(E,\mathbb{C})} \leqslant C_{p,q}\,\|\boldsymbol{A}_\alpha\|_2\,\|f\|_{L^q(\mathbb{G},\mathbb{C})}\mu_z(E)^{1/p-1/q}.$$

An interesting feature is that all the estimates in the five theorems above are sharp when the group $\mathbb{G} = \mathbb{G}_x \times \mathbb{G}_y$ and $\dim(\mathbb{G}_y) + \dim^\infty(\mathbb{G}_x) \geqslant 2$, where $\dim^\infty(\mathbb{G}_x)$ denotes the number of infinite components of $\mathbb{G}_x$.

1.3. **Weak formulations.** Let $f : \mathbb{G} \to \mathbb{C}$ be given. The heat extension $\tilde{f}(t)$ of f is defined as $\tilde{f}(t) := e^{t\Delta_z} f =: P_t f$. We have therefore $\tilde{f}(0) = f$. The aim of this section is to derive weak formulations for second order Riesz transforms. We start with the weak formulation of the identity operator $\mathcal{I}$, that is obtained by using semi-discrete heat extensions (see [3] for details).

Assume f in $L^2(\mathbb{G})$ and g in $L^2(\mathbb{G})$. Let $\bar{f}$ be the average of f on $\mathbb{G}$ if $\mathbb{G}$ has finite measure and zero otherwise. Then

$$\begin{aligned}(\mathcal{I}f, g) &= (f, g)_{L^2(\mathbb{G})} \\ &= \bar{f}\bar{g} + 2\int_0^\infty \left(\widehat{\boldsymbol{\nabla}}_z P_t f, \widehat{\boldsymbol{\nabla}}_z P_t g\right)_{L^2(\hat{T}\mathbb{G})} \mathrm{d}t \\ &= \bar{f}\bar{g} + 2\int_0^\infty \int_{z\in\mathbb{G}} \left\{ \frac{1}{2}\sum_{i=1}^m \sum_{\tau=\pm} (X_i^\tau P_t f)(z)(X_i^\tau P_t g)(z) + \right. \\ &\left. \sum_{j=1}^n (Y_j P_t f)(z)(Y_j P_t g)(z) \right\} \mathrm{d}\mu_z(z)\mathrm{d}t\end{aligned}$$

and the sums and integrals that arise converge absolutely.

In order to pass to the weak formulation for the squares of Riesz transforms, we first observe that the following commutation relations hold

$$\begin{aligned}Y_j \circ \Delta_z &= \Delta_z \circ Y_j \\ X_i^\tau \circ \Delta_z &= \Delta_z \circ X_i^\tau, \quad \tau \in \{+, -\}\end{aligned}$$

This is an easy consequence of the hypothesis made on the Lie group. Following [3], the following weak formulation for second order Riesz transforms holds

Assume f in $L^2(\mathbb{G})$ and g in $L^2(\mathbb{G})$, then

$$\begin{aligned}(R_\alpha^2 f, g)_{L^2(\mathbb{G})} &= -2\int_0^\infty \left(\boldsymbol{A}_\alpha \widehat{\boldsymbol{\nabla}}_z P_t f, \widehat{\boldsymbol{\nabla}}_z P_t g\right)_{L^2(\hat{T}\mathbb{G})} \mathrm{d}t \\ &= -2\int_0^\infty \int_{z\in\mathbb{G}} \left\{ \frac{1}{2}\sum_{i=1}^m \sum_{\tau=\pm} \alpha_i^x \, (X_i^\tau P_t f)(z)(X_i^\tau P_t g)(z) \right. \\ &\quad \left. + \sum_{j,k=1}^n \alpha_{jk}^y \, (Y_j P_t f)(z)(Y_k P_t g)(z) \right\} \mathrm{d}\mu_z(z)\, \mathrm{d}t\end{aligned}$$

and the sums and integrals that arise converge absolutely.

2. Stochastic integrals and martingale transforms

In what follows, we assume that we have a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a càdlàg (i.e. right continuous left limit) filtration $(\mathcal{F}_t)_{t\geqslant 0}$ of sub-σ–algebras of $\mathcal{F}$. We assume as usual that $\mathcal{F}_0$ contains all events of probability zero. All random walks and martingales are adapted to this filtration.

We define below a continuous-time random process $\mathcal{Z}$ with values in $\mathbb{G}$, $\mathcal{Z}_t := (\mathcal{X}_t, \mathcal{Y}_t) \in \mathbb{G}_x \times \mathbb{G}_y$, having infinitesimal generator $L = \Delta_z$. The pure-jump component $\mathcal{X}_t$ is a compound Poisson jump process on the discrete set $\mathbb{G}_x$, wheras the continuous component $\mathcal{Y}_t$ is a standard brownian motion on the manifold $\mathbb{G}_y$. Then, Itô's formula ensures that semi-discrete "harmonic" functions $f : \mathbb{R}^+ \times \mathbb{G} \to \mathbb{C}$ solving the backward heat equation $(\partial_t + \Delta_z) f =$

0 give rise to martingales $M_t^f := f(t, \mathcal{Z}_t)$ for which we define a class of martingale transforms.

2.1. **Stochastic integrals, Martingale transforms and quadratic covariations.**

Stochastic integrals on Riemannian manifolds and Itô integral. Following Emery [26][27], see also Arcozzi [1][2], we define the Brownian motion $\mathcal{Y}_t$ on $\mathbb{G}_y$, a compact Riemannian manifold, as the process $\mathcal{Y}_t : \Omega \to (0,T) \times \mathbb{G}_y$ such that for all smooth functions $f : \mathbb{G}_y \to \mathbb{R}$, the quantity

$$f(\mathcal{Y}_t) - f(\mathcal{Y}_0) - \frac{1}{2}\int_0^t (\Delta_y f)(\mathcal{Y}_s)\ \mathrm{d}s =: (I_{\mathrm{d}_y f})_t \tag{2.1}$$

is an $\mathbb{R}$–valued continuous martingale. For any adapted continuous process Ψ with values in the cotangent space $T^*\mathbb{G}_y$ of $\mathbb{G}_y$, if $\Psi_t(\omega) \in T^*_{Y_t(\omega)}\mathbb{G}_y$ for all $t \geqslant 0$ and $\omega \in \Omega$, then one can define the *continuous* Itô integral I_Ψ of Ψ as

$$(I_\Psi)_t := \int_0^t \langle \Psi_s, \mathrm{d}\mathcal{Y}_s \rangle$$

so that in particular

$$(I_{\mathrm{d}_y f})_t := \int_0^t \langle \mathrm{d}_y f(\mathcal{Y}_s), \mathrm{d}\mathcal{Y}_s \rangle$$

The integrand above involves the 1–form of $T^*_y\mathbb{G}_y$

$$\mathrm{d}_y f(y) := \sum_j (Y_j f)(y)\ Y_j^*.$$

A pure jump process on $\mathbb{G}_x$. We will now define the *discrete* m–dimensional process $\mathcal{X}_t$ on the *discrete* abelian group $\mathbb{G}_x$ as a generalized compound Poisson process. In order to do this we need a number of independent variables and processes:

First, for any given $1 \leqslant i \leqslant m$, let $\mathcal{N}_t^i$ be a càdlàg Poisson process of parameter λ, that is

$$\forall t, \quad \mathbb{P}(\mathcal{N}_t^i = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

The sequence of instants where the jumps of the $\mathcal{N}_t^i$ occur is noted $(T_k^i)_{k\in\mathbb{N}}$, with the convention $T_0^i = 0$.

Second, we set

$$\mathcal{N}_t = \sum_{i=1}^m \mathcal{N}_t^i$$

Almost surely, for any two distinct i and j, we have $\{T_k^i\}_{k\in\mathbb{N}} \cap \{T_k^j\}_{k\in\mathbb{N}} = \emptyset$. Let therefore $\{T_k\}_{k\in\mathbb{N}} = \cup_{i=1}^m \{T_k^i\}_{k\in\mathbb{N}}$ be the ordered sequence of instants of jumps of $\mathcal{N}_t$ and let $i_t \equiv i_t(\omega)$ be the index of the coordinate where the jump occurs at time t. We set $i_t = 0$ if no jump occurs. The random variables i_t are measurable: $i_t = (\mathcal{N}_t^1 - \mathcal{N}_{t^-}^1, \mathcal{N}_t^2 - \mathcal{N}_{t^-}^2, \dots, \mathcal{N}_t^m - \mathcal{N}_{t^-}^m) \cdot (1, 2, \dots, m)$. In differential form,

$$\mathrm{d}\mathcal{N}_t = \sum_{i=1}^{m} \mathrm{d}\mathcal{N}_t^i = \mathrm{d}\mathcal{N}_t^{i_t}.$$

Third, we denote by $(\tau_k)_{k\in\mathbb{N}}$ a sequence of independent Bernoulli variables

$$\forall k, \quad \mathbb{P}(\tau_k = 1) = \mathbb{P}(\tau_k = -1) = 1/2.$$

Finally, the random walk $\mathcal{X}_t$ started at $\mathcal{X}_0 \in \mathbb{G}_x$ is the càdlàg compound Poisson process (see e.g. Protter [48], Privault [46, 47]) defined as

$$\mathcal{X}_t := \mathcal{X}_0 + \sum_{k=1}^{\mathcal{N}_t} G_{i_k}^{\tau_k},$$

where $G_i^\tau = (0, \dots, 0, \tau g_i, 0, \dots, 0)$ when $i \neq 0$ and $(0, \dots, 0)$ when $i = 0$.

Stochastic integrals on discrete groups. We recall for the convenience of the reader the derivation of stochastic integrals for jump processes. We will emphasize the fact that the corresponding Itô's formula involves the action of a discrete 1–form written in a well-chosen local coordinate system of the discrete *augmented* cotangent plane (see details below). Let $1 \leqslant k \leqslant \mathcal{N}_t$ and let (T_k, i_k, τ_k) be respectively the instant, the axis and the direction of the k–th jump. We set $T_0 = 0$. Let $f := f(t,x)$, $t \in \mathbb{R}^+$, $x \in \mathbb{G}_x$ a function defined on $\mathbb{R}^+ \times \mathbb{G}_x$. Then

$$\begin{aligned} & f(t, \mathcal{X}_t) - f(0, \mathcal{X}_0) \\ = & \int_0^t (\partial_t f)(s, \mathcal{X}_s)\mathrm{d}s + \sum_{i=1}^m \int_0^t (f(s, \mathcal{X}_s) - f(s, \mathcal{X}_{s_-}))\, \mathrm{d}\mathcal{N}_s^i. \end{aligned}$$

At an instant s, the integrand in the last term writes as

$$\begin{aligned} & (f(s, \mathcal{X}_s) - f(s, \mathcal{X}_{s_-}))\mathrm{d}\mathcal{N}_s^i \\ = & \left(f\left(s, \mathcal{X}_{s_-} + G_i^{\tau_{\mathcal{N}_s}}\right) - f(s, \mathcal{X}_{s_-})\right) \mathrm{d}\mathcal{N}_s^i \\ = & \left(X_i^{\tau_{\mathcal{N}_s}} f\right)(s, \mathcal{X}_{s_-})\tau_{\mathcal{N}_s}\mathrm{d}\mathcal{N}_s^i \\ = & \frac{1}{2}\left\{(X_i^2 f)(s, \mathcal{X}_{s_-}) + \tau_{\mathcal{N}_s}(X_i^0 f)(s, \mathcal{X}_{s_-})\right\} \mathrm{d}\mathcal{N}_s^i \end{aligned}$$

where we introduced, for all $1 \leqslant i \leqslant m$,

$$\begin{aligned} X_i^0 &:= X_i^+ + X_i^- \\ X_i^2 &:= X_i^+ - X_i^-. \end{aligned}$$

Notice that, for any given $1 \leqslant i \leqslant m$, up to a normalisation factor, the system of coordinate (X_i^2, X_i^0) is obtained thanks to a *rotation* of $\pi/4$ of the canonical system of coordinate (X_i^+, X_i^-). Finally,

$$\begin{aligned} & f(t,\mathcal{X}_t) - f(0,\mathcal{X}_0) \\ = & \int_0^t \left\{ (\partial_t f)(s,\mathcal{X}_s) + \frac{\lambda}{2}(\Delta_x f)(s,\mathcal{X}_s) \right\} \mathrm{d}s + \int_0^t \left\langle \widehat{\mathbf{d}} f(s,\mathcal{X}_{s_-}), \mathrm{d}\widehat{\mathcal{W}}_s \right\rangle \\ =: & \int_0^t \left\{ (\partial_t f)(s,\mathcal{X}_s) + \frac{\lambda}{2}(\Delta_x f)(s,\mathcal{X}_s) \right\} \mathrm{d}s + \left(I_{\widehat{\mathbf{d}}_x f} \right)_t . \end{aligned}$$

where we set $\mathrm{d}\mathcal{X}_s^i = \tau_{\mathcal{N}_s} \mathrm{d}\mathcal{N}_s^i$. It is easy to see that $\mathrm{d}\mathcal{X}_s^i$ is the stochastic differential of a martingale. Here and in the sequel, we take $\lambda = 2$.

Discrete Itô integral. The stochastic integral above shows that Itô formula (2.1) for continuous processes has a discrete counterpart involving stochastic integrals for jump processes, namely we have the *discrete* Itô integral

$$\left(I_{\widehat{\mathbf{d}}_x f} \right)_t \quad := \quad \frac{1}{2} \sum_{i=1}^m \int_0^t (X_i^2 f)(s,\mathcal{X}_{s_-}) \ \mathrm{d}(\mathcal{N}_s^i - \lambda s) + (X_i^0 f)(s,\mathcal{X}_{s_-}) \ \mathrm{d}\mathcal{X}_s$$

This has a more intrinsic expression similar to the continuous Itô integral (2.1). If we regard the discrete component $\mathbb{G}_x$ as a "discrete Riemannian" manifold, then this discrete Itô integral involves discrete vectors (resp. 1–forms) defined on the *augmented* discrete tangent (resp. cotangent) space $\hat{T}_x \mathbb{G}_x$ (resp. $\hat{T}_x^* \mathbb{G}_x$) of dimension $2m$ defined as

$$\begin{aligned} \hat{T}_x \mathbb{G}_x &= \operatorname{span}\{X_1^+, X_2^+, \ldots, X_1^-, X_2^-, \ldots\} \\ &= \operatorname{span}\{X_1^2, X_2^2, \ldots, X_1^0, X_2^0, \ldots\} \\ \hat{T}_x^* \mathbb{G}_x &= \operatorname{span}\{(X_1^+)^*, (X_2^+)^*, \ldots, (X_1^-)^*, (X_2^-)^*, \ldots\} \\ &= \operatorname{span}\{(X_1^2)^*, (X_2^2)^*, \ldots, (X_1^0)^*, (X_2^0)^*, \ldots\}. \end{aligned}$$

Let $\mathrm{d}\widehat{\mathcal{W}}_s \in \hat{T}_{\mathcal{X}_s} \mathbb{G}_x$ be the vector and $\widehat{\mathbf{d}} f \in \hat{T}_{\mathcal{X}_s}^* \mathbb{G}_x$ be the 1–form respectively defined as:

$$\mathrm{d}\widehat{\mathcal{W}}_s = \mathrm{d}(\mathcal{N}_s^1 - \lambda s) X_1^2 + \ldots + \mathrm{d}(\mathcal{N}_s^m - \lambda s) X_m^2 + \mathrm{d}\mathcal{X}_s^1 X_1^0 + \ldots + \mathrm{d}\mathcal{X}_s^m X_m^0$$

$$\widehat{\mathbf{d}}_x f = X_1^2 f (X_1^2)^* + \ldots + X_m^2 f (X_m^2)^* + X_1^0 f (X_1^0)^* + \ldots + X_m^0 f (X_m^0)^*$$

We have with these notations

$$\left(I_{\widehat{\mathbf{d}}_x f} \right)_t \quad := \quad \left\langle \widehat{\mathbf{d}}_x f, \mathrm{d}\widehat{\mathcal{W}}_s \right\rangle_{\hat{T}_x^* \mathbb{G}_x \times \hat{T}_x \mathbb{G}_x}$$

where the factor 1/2 is included in the pairing $\langle \cdot, \cdot \rangle_{\hat{T}_x^* \mathbb{G}_x \times \hat{T}_x \mathbb{G}_x}$.

Semi–discrete stochastic integrals. Let finally $\mathcal{Z}_t = (\mathcal{X}_t, \mathcal{Y}_t)$ be a semi-discrete random walk on the cartesian product $\mathbb{G} = \mathbb{G}_x \times \mathbb{G}_y$, where $\mathcal{X}_t$ is the random walk above defined on $\mathbb{G}_x$ with generator Δ_x and where $\mathcal{Y}_t$ is the Brownian motion defined on $\mathbb{G}_y$ with generator Δ_y. For $f := f(t,z) =$

$f(t,x,y)$ defined from $\mathbb{R}^+ \times \mathbb{G}$ onto $\mathbb{C}$, we have easily the stochastic integral involving both discrete and continuous parts:

$$f(t,\mathcal{Z}_t) \quad = \quad \int_0^t \{(\partial_t f)(s,\mathcal{Z}_s) + (\Delta_z f)(s,\mathcal{Z}_s)\}\,\mathrm{d}s + \Big(I_{\widehat{\mathbf{d}}_z f}\Big)_t$$

where the *semi-discrete* Itô integral writes as

$$\begin{aligned}\Big(I_{\widehat{\mathbf{d}}_z f}\Big)_t \quad &:= \quad \Big(I_{\widehat{\mathbf{d}}_x f}\Big)_t + \big(I_{\mathbf{d}_y f}\big)_t \\ &:= \quad \int_0^t \Big\langle \widehat{\mathbf{d}}_x f(s,\mathcal{Z}_{s_-}), \mathrm{d}\widehat{\mathcal{W}}_s \Big\rangle_{\hat{T}^*_{\mathcal{X}_s}\mathbb{G}_x \times \hat{T}_{\mathcal{X}_s}\mathbb{G}_x} \\ &\qquad + \int_0^t \big\langle \mathbf{d}_y f(s,\mathcal{Z}_{s_-}), \mathrm{d}\mathcal{Y}_s \big\rangle_{\hat{T}^*_{\mathcal{Y}_s}\mathbb{G}_y \times \hat{T}_{\mathcal{Y}_s}\mathbb{G}_y}.\end{aligned}$$

Martingale transforms. We are interested in martingale transforms allowing us to represent second order Riesz transforms. Let $f(t,z)$ be a solution to the heat equation $\partial_t - \Delta_z = 0$. Fix $T > 0$ and $\mathcal{Z}_0 \in \mathbb{G}$. Then define

$$\forall 0 \leqslant t \leqslant T, \quad M_t^{f,T,\mathcal{Z}_0} = f(T-t,\mathcal{Z}_t).$$

This is a martingale since $f(T-t)$ solves the backward heat equation $\partial_t + \Delta_z = 0$, and we have in terms of stochastic integrals

$$M_t^{f,T,\mathcal{Z}_0} = f(T-t,\mathcal{Z}_t) = f(T,\mathcal{Z}_0) + \int_0^t \Big\langle \widehat{\mathbf{d}}_z f(T-s,\mathcal{Z}_{s_-}), \mathrm{d}\mathcal{Z}_s \Big\rangle$$

Given $\boldsymbol{A}_\alpha$ the $\mathbb{C}^{(2m+n)\times(2m+n)}$ matrix defined earlier, we note $M_t^{\alpha,f,T,\mathcal{Z}_0}$ the martingale transform $\boldsymbol{A}_\alpha * M_t^{f,T,\mathcal{Z}_0}$ defined as

$$\begin{aligned} M_t^{\alpha,f,T,\mathcal{Z}_0} \quad &:= \quad f(T,\mathcal{Z}_0) + \int_0^t \Big(\boldsymbol{A}_\alpha \widehat{\boldsymbol{\nabla}}_z f(s,\mathcal{Z}_{s_-}), \mathrm{d}\mathcal{Z}_s\Big) \\ &= \quad f(T,\mathcal{Z}_0) + \int_0^t \Big\langle \widehat{\mathbf{d}}_z f(T-s,\mathcal{Z}_{s_-}) \boldsymbol{A}_\alpha^*, \mathrm{d}\mathcal{Z}_s \Big\rangle \end{aligned}$$

where the first integral involves the L^2 scalar product on $\hat{T}_z\mathbb{G} \times \hat{T}_z\mathbb{G}$ and the second integral involves the duality $\hat{T}_z^*\mathbb{G} \times \hat{T}_z\mathbb{G}$. In differential form:

$$\begin{aligned} &\mathrm{d}M_t^{\alpha,f,T,\mathcal{Z}_0} \\ &\quad = \quad \Big(\boldsymbol{A}_\alpha \widehat{\boldsymbol{\nabla}}_z f(s,\mathcal{Z}_{s_-}), \mathrm{d}\mathcal{Z}_s\Big) \\ &\quad = \quad \sum_{i=1}^m \alpha_i^x \left\{(X_i^2 f)(T-t,\mathcal{Z}_{t_-})\,\mathrm{d}(\mathcal{N}_t^i - \lambda t) + (X_i^0 f)(t,\mathcal{Z}_{t_-})\,\mathrm{d}\mathcal{X}_t^i\right\} \\ &\qquad\quad + \sum_{j=1}^n \alpha_{j,k}^y\,(X_j f)(T-t,\mathcal{Z}_{t_-})\,\mathrm{d}\mathcal{Y}_t^k \end{aligned}$$

Quadratic covariation and subordination. We have the quadratic covariations (see PROTTER [48], DELLACHERIE–MEYER [22], or PRIVAULT [46, 47]). Since

$$\begin{aligned} \mathrm{d}[\mathcal{N}^i - \lambda t, \mathcal{N}^i - \lambda t]_t &= \mathrm{d}\mathcal{N}_t^i \\ \mathrm{d}[\mathcal{N}^i - \lambda t, \mathcal{X}^i]_t &= \tau_{\mathcal{N}_t}\, \mathrm{d}\mathcal{N}_t^i \\ \mathrm{d}[\mathcal{X}^i, \mathcal{X}^i]_t &= \mathrm{d}\mathcal{N}_t^i \\ \mathrm{d}[\mathcal{Y}^j, \mathcal{Y}^j]_t &= \mathrm{d}t, \end{aligned}$$

it follows that

$$\begin{aligned} \mathrm{d}[M^f, M^g]_t &= \sum_{i=1}^m \sum_{\tau=\pm} (X_i^\tau f)\, (X_i^\tau g)(T-t, \mathcal{Z}_{t_-}) \mathbb{1}(\tau_{\mathcal{N}_t} = \tau 1) \mathrm{d}\mathcal{N}_t^i \\ &\quad + (\boldsymbol{\nabla}_y f, \boldsymbol{\nabla}_y g)\,(T-t, \mathcal{Z}_{t_-})\mathrm{d}t \end{aligned} \tag{2.2}$$

Differential subordination. Following WANG [51], given two adapted càdlàg Hilbert space valued martingales X_t and Y_t, we say that Y_t is differentially subordinate by quadratic variation to X_t if $|Y_0|_{\mathbb{H}} \leqslant |X_0|_{\mathbb{H}}$ and $[Y,Y]_t - [X,X]_t$ is nondecreasing nonnegative for all t. In our case, we have

$$\begin{aligned} \mathrm{d}[M^{\alpha,f}, M^{\alpha,f}]_t &= \sum_{i=1}^m |\alpha_i^x|^2 \, \big\{ (X_i^+ f)^2 (T-t, \mathcal{Z}_{t_-}) \mathbb{1}(\tau_{\mathcal{N}_t} = 1) \\ &\quad + (X_i^- f)^2 (T-t, \mathcal{Z}_{t_-}) \mathbb{1}(\tau_{\mathcal{N}_t} = -1) \big\}\, \mathrm{d}\mathcal{N}_t^i \\ &\quad + (\boldsymbol{A}_\alpha^y \boldsymbol{\nabla}_y f, \boldsymbol{A}_\alpha^y \boldsymbol{\nabla}_y f)\,(T-t, \mathcal{Z}_{t_-})\mathrm{d}t. \end{aligned}$$

Hence

$$\mathrm{d}[M^{\alpha,f}, M^{\alpha,f}]_t \leqslant \|\boldsymbol{A}_\alpha\|_2^2 \, \mathrm{d}[M^f, M^f]_t. \tag{2.3}$$

This means that $M_t^{\alpha,f}$ is differentially subordinate to $\|\boldsymbol{A}_\alpha\|_2 M_t^f$.

2.2. **Martingale inequalities under differential subordination.** In the final part of the section we discuss a number of sharp martingale inequalities which hold under the assumption of the differential subordination imposed on the processes. Our starting point is the following celebrated L^p bound.

Theorem 2.1. ***(Wang, 1995)*** *Suppose that X and Y are martingales taking values in a Hilbert space $\mathbb{H}$ such that Y is differentially subordinate to X. Then for any $1 < p < \infty$ we have*

$$||Y||_p \le (p^* - 1)||X||_p$$

and the constant $p^* - 1$ is the best possible, even if $\mathbb{H} = \mathbb{R}$.

This result was first proved by BURKHOLDER in [15] in the following discrete-time setting. Suppose that $(X_n)_{n\ge 0}$ is an $\mathbb{H}$-valued martingale and

$(\alpha_n)_{n\geq 0}$ is a predictable sequence with values in $[-1,1]$. Let $Y := \alpha * X$ be the martingale transform of X defined for almost all $\omega \in \Omega$ by

$$Y_0(\omega) = \alpha_0 X_0(\omega) \quad \text{and} \quad (Y_{n+1} - Y_n)(\omega) = \alpha_n (X_{n+1} - X_n)(\omega).$$

Then the above L^p bound holds true and the constant $p^* - 1$ is optimal. The general continuous-time version formulated above is due to WANG [51]. To see that the preceding discrete-time version is indeed a special case, treat a discrete-time martingale $(X_n)_{n\geq 0}$ and its transform $(Y_n)_{n\geq 0}$ as continuous-time processes via $X_t = X_{\lfloor t \rfloor}$, $Y_t = Y_{\lfloor t \rfloor}$ for $t \geq 0$; then Y is differentially subordinate to X.

In 1992, CHOI [19] established the following non-symmetric, discrete-time version of the L^p estimate.

Theorem 2.2. ***(Choi, 1992)*** *Suppose that $(X_n)_{n\geq 0}$ is a real-valued discrete time martingale and let $(Y_n)_{n\geq 0}$ be its transform by a predictable sequence $(\alpha_n)_{n\geq 0}$ taking values in $[0,1]$. Then there exists a constant $\mathfrak{C}_p$ depending only on p such that $\|Y\|_p \leqslant \mathfrak{C}_p \|X\|_p$ and the estimate is best possible.*

This result can be regarded as a non-symmetric version of the previous theorem, since the transforming sequence $(\alpha_n)_{n\geq 0}$ takes values in a non-symmetric interval $[0,1]$. There is a natural question whether the estimate can be extended to the continuous-time setting; in particular, this gives rise to the problem of defining an appropriate notion of non-symmetric differential subordination. The following statement obtained by BAÑUELOS and OSĘKOWSKI addresses both these questions. For any real numbers $a < b$ and any $1 < p < \infty$, let $\mathfrak{C}_{a,b,p}$ be the constant introduced in [10].

Theorem 2.3. ***(Banuelos–Osękowski, 2012)*** *Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be two real-valued martingales satisfying*

$$\mathrm{d}\left[Y - \frac{a+b}{2}X, Y - \frac{a+b}{2}X\right]_t \leqslant \mathrm{d}\left[\frac{b-a}{2}X, \frac{b-a}{2}X\right]_t \tag{2.4}$$

for all $t \geqslant 0$. Then for all $1 < p < \infty$, we have $\|Y\|_p \leqslant \mathfrak{C}_{a,b,p}\|X\|_p$.

The condition (2.4) is the continuous counterpart of the condition that the transforming sequence $(\alpha_n)_{n\geq 0}$ takes values in the interval $[a,b]$. Thus, in particular, Choi's constant $\mathfrak{C}_p$ is, in the terminology of the above theorem, equal to $\mathfrak{C}_{p,0,1}$.

We return to the context of the "classical" differential subordination introduced in the preceding subsection and study other types of martingale inequalities. The following statements, obtained by BAÑUELOS–OSĘKOWSKI, [11] will allow us to deduce sharp weak-type and logarithmic estimates for Riesz transforms, respectively.

Theorem 2.4. ***(Banuelos–Osękowski, 2015)*** *Suppose that X and Y are martingales taking values in a Hilbert space $\mathbb{H}$ such that Y is differentially subordinate to X.*

(i) Let $1 < p < 2$. Then for any $t \geq 0$,

$$\mathbb{E}\max\left\{|Y_t| - \frac{p^{-1/(p-1)}}{2}\Gamma\left(\frac{p}{p-1}\right), 0\right\} \leq \mathbb{E}|X_t|^p.$$

(ii) Suppose that $2 < p < \infty$. Then for any $t \geq 0$,

$$\mathbb{E}\max\left\{|Y_t| - 1 + p^{-1}, 0\right\} \leq \frac{p^{p-2}}{2}\mathbb{E}|X_t|^p.$$

Both estimates are sharp: for each p, the numbers $\frac{p^{-1/(p-1)}}{2}\Gamma\left(\frac{p}{p-1}\right)$ and $1 - p^{-1}$ cannot be decreased.

Recall that Φ, $\Psi : [0,\infty) \to [0,\infty)$ are conjugate Young functions given by $\Phi(t) = e^t - 1 - t$ and $\Psi(t) = (t+1)\log(t+1) - t$.

Theorem 2.5. ***(Banuelos–Osękowski, 2015)*** *Suppose that X and Y are martingales taking values in a Hilbert space $\mathbb{H}$ such that Y is differentially subordinate to X. Then for any $K > 1$ and any $t \geq 0$ we have*

$$\mathbb{E}\left\{|Y_t| - (2(K-1))^{-1}, 0\right\} \leq K\mathbb{E}\Psi(|X_t|).$$

For each K, the constant $(2(K-1))^{-1}$ appearing on the left, is the best possible (it cannot be replaced by any smaller number).

The following exponential estimate, established by Osękowski in [41], can be regarded as a dual statement to the above logarithmic bound.

Theorem 2.6. ***(Osękowski, 2013)*** *Assume that X, Y are $\mathbb{H}$-valued martingales such that $||X||_\infty \leq 1$ and Y is differentially subordinate to X. Then for any $K > 1$ and any $t \geq 0$ we have*

$$\mathbb{E}\Phi(|Y_t|/K) \leq \frac{1}{2K(K-1)}\mathbb{E}|X_t|. \tag{2.5}$$

Finally, we will need the following sharp $L^q \to L^p$ estimate, established by Osękowski in [43], which will allow us to deduce the corresponding estimate for Riesz transforms.

Theorem 2.7. ***(Osękowski, 2014)*** *Assume that X, Y are $\mathbb{H}$-valued martingales such that Y is differentially subordinate to X. Then for any $1 \leq p < q < \infty$ there is a constant $L_{p,q}$ such that*

$$\mathbb{E}\max\{|Y_t|^p - L_{p,q}, 0\} \leq \mathbb{E}|X_t|^q. \tag{2.6}$$

Actually, the paper [43] identifies, for any p and q as above, the optimal (i.e., the least) value of the constant $L_{p,q}$ in the estimate above. As the description of this constant is a little complicated (and will not be needed in our considerations below), we refer the reader to that paper for the formal definition of $L_{p,q}$.

Let us conclude with the observation which will be crucial in the proofs of our main results. Namely, all the martingale inequalities presented above are of the form $\mathbb{E}\zeta(|Y_t|) \leq \mathbb{E}\xi(|X_t|)$, $t \geq 0$, where ζ, ξ are certain convex functions. This will allow us to successfully apply a conditional version of Jensen's inequality.

3. Proofs of the main results

We turn our attention to the proofs of the estimates for R^2_α formulated in the introductory section. We will focus on Theorems 1.1, 1.2 and 1.3 only; the remaining statements are established by similar arguments. Also, we postpone the proof of the sharpness of these estimates to the next section.

3.1. **Proof of Theorem 1.1.** Recall that the subordination estimate (2.3) shows that the martingale transform $Y_t := M_t^\alpha$ is differentially subordinate to the martingale $X_t := \|\boldsymbol{A}_\alpha\|_2 \, M_t^f$. Therefore, by Theorem 2.1, we immediately obtain that

$$\|M_t^{\alpha,f}\|_p \leqslant \|\boldsymbol{A}_\alpha\|_2 \, (p^* - 1) \, \|M_t^f\|_p$$

for all $t \geq 0$. Since the operator $\mathcal{T}^\alpha$ is a conditional expectation of $M_t^{\alpha,f}$, an application of Jensen's inequality proves the estimate $\|\mathcal{T}^\alpha\|_p \leqslant \|\boldsymbol{A}_\alpha\|_2 \, (p^* - 1)$, which is the desired bound.

3.2. **Proof of Theorem 1.2.** The argument is the same as above and exploits the fine-tuned L^p estimate of Theorem 2.3 applied to $X_t = M_t^f$ and $Y_t = M_t^{\alpha,f}$. It is not difficult to prove that the difference of quadratic variations above writes in terms of a jump part and a continuous part as

$$\begin{aligned}
&\left[Y - \frac{a+b}{2}X, Y - \frac{a+b}{2}X\right]_t - \mathrm{d}\left[\frac{b-a}{2}X, \frac{b-a}{2}X\right]_t \\
&= \sum_{i=1}^m \sum_{\pm} (\alpha_i^x - a)(\alpha_i^x - b)\,(X_i^{\pm} f)^2(\mathcal{B}_t)\,\mathbb{1}(\tau_{N_t} = \pm 1)\,\mathrm{d}\mathcal{N}_t^i \\
&\quad + \langle (\boldsymbol{A}_\alpha^y - a\boldsymbol{I})\,(\boldsymbol{A}_\alpha^y - b\boldsymbol{I})\,\boldsymbol{\nabla}_y f(\mathcal{B}_t), \boldsymbol{\nabla}_y f(\mathcal{B}_t)\rangle\,\mathrm{d}t,
\end{aligned}$$

which is nonpositive since we assumed precisely $a\boldsymbol{I} \leqslant \boldsymbol{A}_\alpha \leqslant b\boldsymbol{I}$. Thus, the estimate of Theorem 1.2 follows. The sharpness is established in a similar manner. □

3.3. **Proof of Theorem 1.3.** We will focus on the case $1 < p < 2$; for remaining values of p the argument is similar. An application of Theorem 2.4 to the processes $X_t = \|\boldsymbol{A}_\alpha\|_2 \, M_t^f$ and $Y_t = M_t^{\alpha,f}$ yields

$$\mathbb{E}\max\left\{|M_t^{\alpha,f}| - \frac{p^{-1/(p-1)}}{2}\Gamma\left(\frac{p}{p-1}\right), 0\right\} \leq \|\boldsymbol{A}_\alpha\|_2^p \, \mathbb{E}|M_t^f|^p$$

and hence, by Jensen's inequality, we obtain

$$\int_{\mathbb{G}} \max\left\{|R_\alpha^2 f| - \frac{p^{-1/(p-1)}}{2}\Gamma\left(\frac{p}{p-1}\right), 0\right\} d\mu_z \leq \|\boldsymbol{A}_\alpha\|_2^p \, ||f||^p_{L^p(\mathbb{G})}.$$

Therefore, if E is an arbitrary measurable subset of $\mathbb{G}$, we get

$$\begin{aligned}\int_E |R_\alpha^2 f| d\mu_z &\leq \int_E \left(|R_\alpha^2 f| - \frac{p^{-1/(p-1)}}{2}\Gamma\left(\frac{p}{p-1}\right)\right) \mathrm{d}\mu_z \\ &\quad + \frac{p^{-1/(p-1)}}{2}\Gamma\left(\frac{p}{p-1}\right)\mu_z(E) \\ &\leq ||f||_{L^p(\mathbb{G})}^p + \frac{p^{-1/(p-1)}}{2}\Gamma\left(\frac{p}{p-1}\right)\mu_z(E).\end{aligned}$$

Apply this bound to λf, where λ is a nonnegative parameter, then divide both sides by λ and optimize the right-hand side over λ to get the desired assertion.

4. Sharpness

The proof of the sharpness of the different results is made in several steps. In some cases the sharpness for certain second order Riesz transform estimates in the continuous setting (such as in Theorem 1.1) is already known. In these cases we prove below the sharpness for the discrete (or semidiscrete) case by using sequences of finite difference approximates of continuous functions and their finite difference second order Riesz transforms. In other cases, we need to prove first sharpness for certain continuous second order Riesz transforms. The key point here is to transfer the sharp result for zigzag martingales into a sharp result for certain continuous second order Riesz transforms by the laminate technique. We will illustrate this for the weak-type estimate of Theorem 1.3 and establish the following statement.

Theorem 4.1. *Let $\Theta : [0,\infty) \to [0,\infty)$ be a given function and let $\lambda > 0$ be a fixed number. Assume further that there is a pair (F,G) of finite martingales starting from $(0,0)$ such that G is a ± 1-transform of F and*

$$\mathbb{E}(|G_\infty| - \lambda)_+ > \mathbb{E}\Theta(|F_\infty|).$$

Then there is a function $f : \mathbb{R}^2 \to \mathbb{R}$ supported on the unit disc $\mathbb{D}$ of $\mathbb{R}^2$ such that

$$\int_{\mathbb{R}^2} \big(|(R_1^2 - R_2^2)f| - \lambda\big)_+ dx > \int_{\mathbb{D}} \Theta(|f|)dx.$$

We will prove this statement with the use of laminates, important family of probability measures on matrices. It is convenient to split this section into several separate parts. For the sake of convenience, and to make this section as self contained as possible, we recall the preliminaries on laminates and their connections to martingales from [14] and [39], Section 4.2.

4.1. **Laminates.** Assume that $\mathbb{R}^{m\times n}$ stands for the space of all real matrices of dimension $m \times n$ and $\mathbb{R}^{n\times n}_{sym}$ denote the subclass of $\mathbb{R}^{n\times n}$ which consists of all symmetric matrices of dimension $n \times n$.

Definition 4.2. A function $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is said to be *rank-one convex*, if for all $A, B \in \mathbb{R}^{m\times n}$ with rank $B = 1$, the function $t \mapsto f(A+tB)$ is convex

For other equivalent definitions of rank-one convexity, see [21, p. 100]. Suppose that $\mathcal{P} = \mathcal{P}(\mathbb{R}^{m\times n})$ is the class of all compactly supported probability measures on $\mathbb{R}^{m\times n}$. For a measure $\nu \in \mathcal{P}$, we define

$$\overline{\nu} = \int_{\mathbb{R}^{m\times n}} X d\nu(X),$$

the associated *center of mass* or *barycenter* of ν.

Definition 4.3. We say that a measure $\nu \in \mathcal{P}$ is a *laminate* (and write $\nu \in \mathcal{L}$), if

$$f(\overline{\nu}) \leq \int_{\mathbb{R}^{m\times n}} f \mathrm{d}\nu$$

for all rank-one convex functions f. The set of laminates with barycenter 0 is denoted by $\mathcal{L}_0(\mathbb{R}^{m\times n})$.

Laminates can be used to obtain lower bounds for solutions of certain PDEs, as observed by Faraco in [30]. In addition, laminates appear naturally in the context of convex integration, where they lead to interesting counterexamples, see e.g. [5], [20], [34], [37] and [49]. For our results here we will be interested in the case of 2×2 symmetric matrices. The key observation is that laminates can be regarded as probability measures that record the distribution of the gradients of smooth maps: see Corollary 4.7 below. We briefly explain this and refer the reader to the works [33], [37] and [49] for full details.

Definition 4.4. Let U be a subset of $\mathbb{R}^{2\times 2}$ and let $\mathcal{PL}(U)$ denote the smallest class of probability measures on U which

(i) contains all measures of the form $\lambda\delta_A + (1-\lambda)\delta_B$ with $\lambda \in [0,1]$ and satisfying $\operatorname{rank}(A - B) = 1$;
(ii) is closed under splitting in the following sense: if $\lambda\delta_A + (1 - \lambda)\nu$ belongs to $\mathcal{PL}(U)$ for some $\nu \in \mathcal{P}(\mathbb{R}^{2\times 2})$ and μ also belongs to $\mathcal{PL}(U)$ with $\overline{\mu} = A$, then also $\lambda\mu + (1 - \lambda)\nu$ belongs to $\mathcal{PL}(U)$.

The class $\mathcal{PL}(U)$ is called the *prelaminates* in U.

It follows immediately from the definition that the class $\mathcal{PL}(U)$ only contains atomic measures. Also, by a successive application of Jensen's inequality, we have the inclusion $\mathcal{PL} \subset \mathcal{L}$. The following are two well known lemmas in the theory of laminates; see [5], [33], [37], [49].

Lemma 4.5. *Let $\nu = \sum_{i=1}^{N} \lambda_i \delta_{A_i} \in \mathcal{PL}(\mathbb{R}^{2\times 2}_{sym})$ with $\overline{\nu} = 0$. Moreover, let $0 < r < \frac{1}{2} \min |A_i - A_j|$ and $\delta > 0$. For any bounded domain $\mathcal{B} \subset \mathbb{R}^2$ there exists $u \in W_0^{2,\infty}(\mathcal{B})$ such that $\|u\|_{C^1} < \delta$ and for all $i = 1 \dots N$*

$$\big|\{x \in \mathcal{B} : |D^2 u(x) - A_i| < r\}\big| = \lambda_i |\mathcal{B}|.$$

Lemma 4.6. *Let $K \subset \mathbb{R}^{2\times 2}_{sym}$ be a compact convex set and suppose that $\nu \in \mathcal{L}(\mathbb{R}^{2\times 2}_{sym})$ satisfies $\operatorname{supp}\nu \subset K$. For any relatively open set $U \subset \mathbb{R}^{2\times 2}_{sym}$ with $K \subset U$, there exists a sequence $\nu_j \in \mathcal{PL}(U)$ of prelaminates with $\overline{\nu}_j = \overline{\nu}$ and $\nu_j \overset{*}{\rightharpoonup} \nu$, where $\overset{*}{\rightharpoonup}$ denotes weak convergence of measures.*

Combining these two lemmas and using a simple mollification, we obtain the following statement, proved by Boros, Shékelyhidi Jr. and Volberg [14]. It exhibits the connection between laminates supported on symmetric matrices and second derivatives of functions. It will be our main tool in the proof of the sharpness. Recall that $\mathbb{D}$ denotes the unit disc of $\mathbb{C}$.

Corollary 4.7. *Let $\nu \in \mathcal{L}_0(\mathbb{R}^{2\times 2}_{sym})$. Then there exists a sequence $u_j \in C_0^\infty(\mathbb{D})$ with uniformly bounded second derivatives, such that*

$$\frac{1}{|\mathbb{D}|}\int_{\mathbb{D}} \phi(D^2 u_j(x))\,dx \;\to\; \int_{\mathbb{R}^{2\times 2}_{sym}} \phi\,d\nu$$

for all continuous $\phi : \mathbb{R}^{2\times 2}_{sym} \to \mathbb{R}$.

4.2. **Biconvex functions and a special laminate.** The next step in our analysis is devoted to the introduction of a certain special laminate. We need some additional notation. A function $\zeta : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ is said to be *biconvex* if for any fixed $z \in \mathbb{R}$, the functions $x \mapsto \zeta(x,z)$ and $y \mapsto \zeta(z,y)$ are convex. Now, take the martingales F and G appearing in the statement of Theorem 4.1. Then the martingale pair

$$(\mathtt{F},\mathtt{G}) := \left(\frac{F+G}{2}, \frac{F-G}{2}\right)$$

is finite, starts from $(0,0)$ and has the following *zigzag* property: for any $n \geq 0$ we have $\mathtt{F}_n = \mathtt{F}_{n+1}$ with probability 1 or $\mathtt{G}_n = \mathtt{G}_{n+1}$ almost surely; that is, in each step $(\mathtt{F},\mathtt{G})$ moves either vertically, or horizontally. Indeed, this follows directly from the assumption that G is a ± 1-transform of F. This property combines nicely with biconvex functions: if ζ is such a function, then a successive application of Jensen's inequality gives

$$\mathbb{E}\zeta(\mathtt{F}_n,\mathtt{G}_n) \geq \mathbb{E}\zeta(\mathtt{F}_{n-1},\mathtt{G}_{n-1}) \geq \ldots \geq \mathbb{E}\zeta(\mathtt{F}_0,\mathtt{G}_0) = \zeta(0,0). \tag{4.1}$$

The distribution of the terminal variable $(\mathtt{F}_\infty,\mathtt{G}_\infty)$ gives rise to a probability measure ν on $\mathbb{R}^{2\times 2}_{sym}$: put

$$\nu\left(\operatorname{diag}(x,y)\right) = \mathbb{P}\big((\mathtt{F}_\infty,\mathtt{G}_\infty) = (x,y)\big), \qquad (x,y) \in \mathbb{R}^2,$$

where $\operatorname{diag}(x,y)$ stands for the diagonal matrix $\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$. Observe that ν is a laminate of barycenter 0. Indeed, if $\psi : \mathbb{R}^{2\times 2} \to \mathbb{R}$ is a rank-one convex, then $(x,y) \mapsto \psi(\operatorname{diag}(x,y))$ is biconvex and thus, by (4.1),

$$\int_{\mathbb{R}^{2\times 2}} \psi \mathrm{d}\nu = \mathbb{E}\psi(\operatorname{diag}(\mathtt{F}_\infty,\mathtt{G}_\infty)) \geq \psi(\operatorname{diag}(0,0)) = \psi(\bar{\nu}).$$

Here we used the fact that $(\mathtt{F},\mathtt{G})$ is finite, so $(\mathtt{F}_\infty,\mathtt{G}_\infty) = (\mathtt{F}_n,\mathtt{G}_n)$ for some n.

4.3. **A proof of Theorem 4.1.** Consider a continuous function $\phi : \mathbb{R}^{2\times 2}_{sym} \to \mathbb{R}$ given by

$$\phi(A) = (|A_{11} - A_{22}| - \lambda)_+ - \Theta(|A_{11} + A_{22}|).$$

By Corollary 4.7, there is a functional sequence $(u_j)_{j\geq 1} \subset C_0^\infty(\mathbb{D})$ such that

$$\begin{aligned}\frac{1}{|\mathbb{D}|}\int_{\mathbb{R}^2} \phi(D^2u_j)\mathrm{d}x &= \frac{1}{|\mathbb{D}|}\int_{\mathbb{D}} \phi(D^2u_j)\mathrm{d}x \\ &\xrightarrow{j\to\infty} \int_{\mathbb{R}^{2\times 2}_{sym}} \phi \mathrm{d}\nu = \mathbb{E}(|G_\infty| - \lambda)_+ - \mathbb{E}\Theta(|F_\infty|) > 0.\end{aligned}$$

Therefore, for sufficiently large j, we have

$$\int_{\mathbb{R}^2} \left(\left\| \frac{\partial^2 u_j}{\partial x^2} - \frac{\partial^2 u_j}{\partial y^2} \right\| - \lambda \right)_+ \mathrm{d}x\mathrm{d}y > \int_{\mathbb{R}^2} \Theta\left(|\Delta u_j|\right) \mathrm{d}x\mathrm{d}y.$$

Setting $f = \Delta u_j$, we obtain the desired assertion.

In the remaining part of this subsection, let us briefly explain how Theorem 4.1 yields the sharpness of weak-type and logarithmic estimates for second-order Riesz transforms (in the classical setting). We will focus on the weak-type bounds for $1 < p < 2$ - the remaining estimates can be treated analogously. Suppose that λ_p is the best constant in the estimate

$$\mathbb{E}(|G_\infty| - \lambda_p)_+ \leq \mathbb{E}|F_\infty|^p,$$

valid for all pairs (F, G) of finite martingales starting from 0 such that G is a ± 1-transform of F. The value of λ_p appears in the statement of Theorem 9 above, the fact that it is already the best for martingale transforms follows from the examples exhibited in [38]. For any $\varepsilon > 0$, Theorem 4.1 yields the existence of $f : \mathbb{R}^2 \to \mathbb{R}$, supported on the unit disc, such that

$$\int_{\mathbb{R}^2} \left(|(R_1^2 - R_2^2)f| - \lambda_p + \varepsilon\right)_+ \mathrm{d}x\mathrm{d}y > \int_{\mathbb{R}^2} |f|^p \mathrm{d}x\mathrm{d}y.$$

That is, if we set $A = \{|(R_1^2 - R_2^2)f| \geq \lambda_p - \varepsilon\}$, we get

$$\int_A |(R_1^2 - R_2^2)f|\mathrm{d}x\mathrm{d}y > \int_{\mathbb{R}^2} |f|^p \mathrm{d}x\mathrm{d}y + (\lambda_p - \varepsilon)|A|. \tag{4.2}$$

However, if the weak-type estimate holds with a constant c_p, Young's inequality implies

$$\int_A |(R_1^2 - R_2^2)f|\mathrm{d}x\mathrm{d}y \leq \int_{\mathbb{R}^2} |f|^p \mathrm{d}x\mathrm{d}y + \frac{(p-1)c_p^{p/(p-1)}}{p^{p/(p-1)}}|A|.$$

Therefore, the inequality (4.2) enforces that

$$\frac{(p-1)c_p^{p/(p-1)}}{p^{p/(p-1)}} \geq \lambda_p$$

(since ε was arbitrary). This estimate is equivalent to

$$c_p \geq \left(\frac{1}{2}\Gamma\left(\frac{2p-1}{p-1}\right)\right)^{1-1/p},$$

which is the desired sharpness.

4.4. **From continuous to discrete sharp estimates.** We claim that the sharp bounds found for the continuous second order Riesz transforms also hold in the case of purely discrete groups. Groups of mixed type would be treated in the same manner. We illustrate those results only for the sharpness in Theorem 1.1 and in Theorem 1.3 since other results follow the same lines. Precisely, we show that the sharpness in the discrete case is inherited from the sharpness of the continuous case through the use of the so–called fundamental theorem of finite difference methods from Lax and Richtmyer [35] (see also [36]). This result states that **stability** and **consistency** of the finite difference scheme implies **convergence** of the approximate finite difference solution towards the continuous solution, in a sense that we detail below.

Finite difference Riesz transforms. Let $u = R_i^2 f$ be the i-th second order Riesz transform in $\Omega := \mathbb{R}^N$ of a function $f \in L^p$. The function u is the unique solution to the Poisson problem in $\mathbb{R}^N$, $\Delta u = \partial_i^2 f$ in $\mathbb{R}^N$ (see [29]). This is a problem of the form $Au = Bf$, where $A = \Delta$ and $B = \partial_{ij}^2$. Introduce now a finite difference grid of step–size $h > 0$, that is the grid $\Omega_h := h\mathbb{Z}^N$. The functions v_h defined on Ω_h are equipped with the L^p_h norm defined as

$$\|v_h\|^p_{L^p_h} := \sum_{x\in\Omega_h} |v_h(x)|^p h^N.$$

It is common to identify a finite difference function v_h defined on the grid Ω_h with the piecewise constant function (also denoted) $v_h : \mathbb{R}^N \to \mathbb{C}$ such that $v_h(x) = v_h(y)$ for all x's in the open cube $\Omega(y)$ of volume h^N centered around the grid point $y \in \Omega_h$. With this notation, we might write finite difference integrals in the form

$$\|v_h\|_{L^p_h} = \int_{x\in\mathbb{R}^N} |v_h(x)| \mathrm{d}\mu_h(x).$$

The finite difference second order Riesz transform $u_h = R_i^2 f_h$ of f_h is the solution to the problem $A_h u_h = B_h^2 f_h$, where $A_h := \Delta_h$ is the finite difference Laplacian and $B_h := \partial^2_{i,h}$ the 3–point finite difference second order derivative. Precisely, for any $x \in \Omega_h$, any $v_h : \Omega_h \to \mathbb{R}$,

$$(\partial^2_{i,h} v_h)(x) := \frac{v_h(x + he_i) - 2v_h(x) + v_h(x - he_i)}{h^2}$$

$$(\Delta_h v_h)(x) := \sum_{i=1}^N (\partial^2_{i,h} v_h)(x).$$

It is classical that we have the **consistency** of the discrete problem with respect to the continuous problem, that is for given smooth functions u and f we have $\Delta_h u = \Delta u + \mathcal{O}(h)$ and $\partial^2_{i,h} f = \partial_i^2 f + \mathcal{O}(h)$, where the coefficients in $\mathcal{O}(h)$ include as a factor up to fourth–order derivatives of u

or f. This implies in particular that $B_h f = Bf + \mathcal{O}(h)$ in L^p_h for any given smooth function f with compact support. It is also classical that $(-\Delta_h)^{-1}$ is bounded in L^p_h uniformly w.r.t. h. This is the L^p **stability** of the finite difference scheme. The fundamental theorem of finite difference methods implies the L^p **convergence** of the sequence of discrete second order Riesz transforms u_h towards the continuous second order Riesz transform u.

Discrete Riesz tranforms on Lie-Group. Observe that the finite difference Riesz transform $u_h = R^2_{i,h} f_h$ defined on the grid Ω_h, also gives rise to a Riesz transform on the Lie group $\Omega_1 = \mathbb{Z}^N$. This is a consequence of the homogeneity of order zero of the Riesz transforms. Indeed, the equation $\Delta_h u_h = \partial^2_{i,h} f_h$ rewrites as $\Delta_1 u_1 = \partial^2_{i,1} f_1$, where $u_1(y) := u_h(y/h)$, $f_1(y) := f_h(y/h)$ for all $y \in \mathbb{Z}^N$, and where Δ_1 and $\partial^2_{i,1}$ are the discrete differential operators defined on $\mathbb{Z}^N$. We have also $\|u_h\|_{L^p_h} = h^{N/p}\|u_1\|_{L^p_1}$ and $\|f_h\|_{L^p_h} = h^{N/p}\|f_1\|_{L^p_1}$. Notice that for all h, this ensures that $\|u_h\|_{L^p_h}/\|f_h\|_{L^p_h} = \|u_1\|_{L^p_1}/\|f_1\|_{L^p_1}$.

Sharpness for Theorem 1.1 in the discrete setting. In the continuous setting, the sharpness was proved in [31] based on the combintation $R^2_\alpha = R^2_1 - R^2_2$ of second order Riesz tranforms. Let $u^{(k)} = R^2_\alpha f^{(k)}$ a sequence of second order Riesz transforms yielding the sharp constant C_p in the estimate, that is $\|u^{(k)}\|_p/\|f^{(k)}\|_p \to C_p$ as k goes to infinity. For each $k \in \mathbb{N}$ and $h > 0$, introduce the finite difference approximation $f^{(k)}_h$ of $f^{(k)}$ and the corresponding finite difference Riesz transform. Thanks to the convergence of the finite difference scheme, we can extract a subsequence $f^{(k)}_{h_k}$ such that $\|u^{(k)}_1\|_p/\|f^{(k)}_1\|_p = \|u^{(k)}_{h_k}\|_p/\|f^{(k)}_{h_k}\|_p \to C_p$. Therefore C_p is also the sharp constant for the second order Riesz transforms in $\mathbb{Z}^N$.

Sharpness for Theorem 1.3 in the discrete setting. Recall that we have a bound of the form

$$\|R^2_\alpha f\|_{L^{p,\infty}(\mathbb{G},\mathbb{C})} := \sup_E \left\{ \mu_z(E)^{1/p-1} \int_E |R^2_\alpha f| \mathrm{d}\mu_z \right\} \leqslant C_p \|f\|_{L^p}$$

for a certain constant C_p that is known to be sharp in the case of continuous second order Riesz transforms. In order to prove sharpness when the Lie group $\mathbb{G}$ does not have enough continuous components, it suffices again to approximate a sequence of continuous extremizers by a sequence of finite difference approximations. Take $\mathbb{G} = \mathbb{R}^N$. For any $\varepsilon > 0$, let f, $u := R^2_\alpha f$, and E with finite measure chosen so that

$$\mu_z(E)^{1/p-1}\|u\|_{L^1(E)}/\|f\|_{L^p} \geqslant C_p - \varepsilon.$$

We can assume without loss of generality that f is a smooth function with compact support. Let f_h a finite difference approximation of f defined as its L^2 projection on the grid, and u_h its discrete second order Riesz transform both defined on $\Omega_h := h\mathbb{Z}^N$. Since $\mu_z(E)$ is the finite N–dimensional Lebesgue measure of E, we use outer measure approximations of E followed by approximations from below by a finite number of small enough cubes of size h centered around the grid points of Ω_h, to define a "finite difference" approximation E_h of E such that

$$\mu_h(E_h) := \sum_{x\in E_h} h^N \to \mu_z(E)$$

when h goes to zero. Since the discrete Riesz transforms are stable in L^2, the Lax-Richtmyer theorem ensures that $\|u_h\|_{L^2_h} \to \|u\|_{L^2}$ which implies $\|u_h\|_{L^1_h(E)} \to \|u\|_{L^1(E)}$ and also $\|u_h\|_{L^1_h(E_h)} \to \|u\|_{L^1(E)}$. Therefore for h small enough,

$$\mu_h(E_h)^{1/p-1}\|u_h\|_{L^1_h(E_h)}/\|f_h\|_{L^p} \geqslant C_p - 2\varepsilon.$$

Let as before $u_1(y) := u_h(y/h)$, $f_1(y) := f_h(y/h)$ for all $y \in \Omega_1 := \mathbb{Z}^N$, and $E_1 := E/h$. We have sucessively $\mu_h(E) = h^N\mu_1(E_1)$, $\|u_h\|_{L^1_h(E_h)} = h^N\|u_1\|_{L^1_h(E_1)}$ and $\|f_h\|_{L^p} = h^{N/p}\|f_1\|_{L^p}$. This yields immediately

$$\mu_1(E_1)^{1/p-1}\|u_1\|_{L^1_h(E_1)}/\|f_1\|_{L^p} \geqslant C_p - 2\varepsilon,$$

allowing us to prove sharpness for the class of discrete groups we are interested in.

References

[1] Nicola Arcozzi. *Riesz transforms on spheres and compact Lie groups.* ProQuest LLC, Ann Arbor, MI, 1995. Thesis (Ph.D.)–Washington University in St. Louis.

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

[3] N. Arcozzi, K. Domelevo, and S. Petermichl. *Second Order Riesz Transforms on MultiplyConnected Lie Groups and Processes with Jumps.* Potential Anal., 45(4):777794, 2016.

[4] K. Astala, *Area distortion of quasiconformal mappings*, Acta. Math. **173** (1994), 37–60.

[5] K. Astala, D. Faraco, L. Székelyhidi, Jr., *Convex integration and the L^p theory of elliptic equations.* Ann. Sc. Norm. Super. Pisa Cl. Sci. (5), Vol. 7 (2008), pp. 1–50.

[6] R. Bañuelos and P. J. Méndez-Hernández. Space-time Brownian motion and the Beurling-Ahlfors transform. *Indiana Univ. Math. J.*, 52(4):981–990, 2003.

[7] Rodrigo Bañuelos. Martingale transforms and related singular integrals. *Trans. Amer. Math. Soc.*, 293(2):547–563, 1986.

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

[9] Rodrigo Bañuelos and Prabhu Janakiraman. L^p-bounds for the Beurling-Ahlfors transform. *Trans. Amer. Math. Soc.*, 360(7):3603–3612, 2008.

[10] Rodrigo Banuelos and Adam Osękowski. Martingales and sharp bounds for Fourier multipliers. *Ann. Acad. Sci. Fenn., Math.*, 37(1):251–263, 2012.

[11] Rodrigo Banuelos and Adam Osękowski. Sharp martingale inequalities and applications to Riesz transforms on manifolds, Lie groups and Gauss space. *J. Funct. Anal.*, 269(6):1652–1713, 2015.

[12] 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.

[13] Alexander Borichev, Prabhu Janakiraman, and Alexander Volberg. Subordination by conformal martingales in L^p and zeros of Laguerre polynomials. *Duke Math. J.*, 162(5):889–924, 2013.

[14] N. Boros, L. Székelyhidi Jr. and A. Volberg, *Laminates meet Burkholder functions*, Journal de Mathématiques Pures et Appliquées, to appear.

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

[16] D. L. Burkholder. A sharp and strict L^p-inequality for stochastic integrals. *Ann. Probab.*, 15(1):268–273, 1987.

[17] 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).

[18] Donald L. Burkholder. Explorations in martingale theory and its applications. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 1–66. Springer, Berlin, 1991.

[19] 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.

[20] S. Conti, D. Faraco, F. Maggi, *A new approach to counterexamples to L^1 estimates: Korn's inequality, geometric rigidity, and regularity for gradients of separately convex functions*, Arch. Rat. Mech. Anal. 175 no. 2 (2005), pp. 287–300.

[21] B. Dacoronga, *Direct Methods in the Calculus of Variations,* Springer 1989.

[22] Claude Dellacherie and Paul-André Meyer. *Probabilities and potential. B*, volume 72 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1982. Theory of martingales, Translated from the French by J. P. Wilson.

[23] Komla Domelevo and Stefanie Petermichl. Bilinear embeddings on graphs. *preprint*, 2014.

[24] Komla Domelevo and Stefanie Petermichl. Sharp L^p estimates for discrete second order Riesz transforms. *Adv. Math.*, 262:932–952, 2014.

[25] Komla Domelevo and Stefanie Petermichl. Sharp L^p estimates for discrete second-order Riesz transforms. *C. R. Math. Acad. Sci. Paris*, 352(6):503–506, 2014.

[26] Michel Emery. Martingales continues dans les variétés différentiables. In *Lectures on probability theory and statistics (Saint-Flour, 1998)*, volume 1738 of *Lecture Notes in Math.*, pages 1–84. Springer, Berlin, 2000.

[27] Michel Emery. An Invitation to Second-Order Stochastic Differential Geometry. https://hal.archives-ouvertes.fr/hal-00145073, 2005. 42 pages.

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

[29] Lawrence C. Evans. *Partial Differential Equations*. Amer. Math. Soc., 1998.

[30] D. Faraco, *Milton's conjecture on the regularity of solutions to isotropic equations*, Ann. Inst. H. Poincaré Anal. Non Linéaire, **20** (2003), pp. 889–909.

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

[32] Richard F. Gundy and Nicolas Th. Varopoulos. Les transformations de Riesz et les intégrales stochastiques. *C. R. Acad. Sci. Paris Sér. A-B*, 289(1):A13–A16, 1979.

[33] B. Kirchheim, *Rigidity and Geometry of Microstructures*, Habilitation Thesis, University of Leipzig (2003), http://www.mis.mpg.de/publications/other-series/ln/lecturenote-1603.html

[34] B. Kirchheim, S. Müller, V. Šverák, *Studying nonlinear pde by geometry in matrix space*, Geometric Analysis and nonlinear partial differential equations, Springer (2003), pp. 347–395.

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

[36] Randall LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems (Classics in Applied Mathematics Classics in Applied Mathemat)*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2007.

[37] S. Müller, V. Šverák, *Convex integration for Lipschitz mappings and counterexamples to regularity*, Ann. of Math. (2), 157 no. 3 (2003), pp. 715–742.

[38] A. Osękowski, *Weak-type inequalities for Fourier multipliers with applications to the Beurling-Ahlfors transform*, J. Math. Soc. Japan 66 (2014), no. 3, 745–764.

[39] A. Osękowski, *On restricted weak-type constants of Fourier multipliers*, Publ. Mat. 58 (2014), no. 2, 415–443.

[40] Adam Osękowski. Sharp moment inequalities for differentially subordinated martingales. *Studia Math.* 201:103–131, 2010.

[41] Adam Osękowski. Logarithmic inequalities for second-order Riesz transforms and related Fourier multipliers. *Colloq. Math.*, 130(1):103–126, 2013.

[42] Adam Osękowski. Sharp weak-type inequalities for Fourier multipliers and second-order Riesz transforms. *Cent. Eur. J. Math.*, 12(8):1198–1213, 2014.

[43] Adam Osękowski. Sharp localized inequalities for Fourier multipliers. *Canadian J. Math.*, 66: 1358–1381, 2014.

[44] 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.

[45] 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.

[46] Nicolas Privault. *Stochastic analysis in discrete and continuous settings with normal martingales*, volume 1982 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009.

[47] Nicolas Privault. *Stochastic finance*. Chapman & Hall/CRC Financial Mathematics Series. CRC Press, Boca Raton, FL, 2014. An introduction with market examples.

[48] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[49] L. Székelyhidi, Jr., *Counterexamples to elliptic regularity and convex integration*, Contemp. Math. 424 (2007), pp. 227–245.

[18] Igor E. Verbitsky. Estimate of the norm of a function in a Hardy space in terms of the norms of its real and imaginary parts. *Mat. Issled.*, (54):16–20, 164–165, 1980. English transl.: Amer. Math. Soc. Transl. (2) 124 (1984), 11-15.

[50] 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.

[51] Gang Wang. Differential subordination and strong differential subordination for continuous-time martingales and related sharp inequalities. *Ann. Probab.*, 23(2):522–551, 1995.