# DIFFERENTIAL SUBORDINATION UNDER CHANGE OF LAW

BY KOMLA DOMELEVO AND STEFANIE PETERMICHL[1,2]

*Université Paul Sabatier*

We prove optimal $L^2$ bounds for a pair of Hilbert space valued differentially subordinate martingales under a change of law. The change of law is given by a process called a weight and sharpness, and in this context refers to the optimal growth with respect to the characteristic of the weight. The pair of martingales is adapted, uniformly integrable and càdlàg. Differential subordination is in the sense of Burkholder, defined through the use of the square bracket. In the scalar dyadic setting with underlying Lebesgue measure, this was proved by Wittwer [*Math. Res. Lett.* **7** (2000) 1–12], where homogeneity was heavily used. Recent progress by Thiele–Treil–Volberg [*Adv. Math.* **285** (2015) 1155–1188] and Lacey [*Israel J. Math.* **217** (2017) 181–195] independently resolved the so-called nonhomogenous case using discrete in time filtrations, where one martingale is a predictable multiplier of the other. The general case for continuous-in-time filtrations and pairs of martingales that are not necessarily predictable multipliers, remained open and is addressed here. As a very useful second main result, we give the explicit expression of a Bellman function of four variables for the weighted estimate of subordinate martingales with jumps. This construction includes an analysis of the regularity of this function as well as a very precise convexity, needed to deal with the jump part.

**1. Introduction.** The paper by Nazarov–Treil–Volberg [24] has set the groundwork for the early advances in modern weighted theory in harmonic analysis and probability that started around twenty years ago. In their paper, the authors show necessary and sufficient conditions for a dyadic martingale transform to be bounded in the $L^2$ two-weight setting. The methodology of their proof could be used to get the first sharp result in the real valued one-weight setting, for the dyadic martingale transform [35]. Sharpness in this setting means best control on growth with respect to the $A_2$ characteristics

$$Q_2^{\mathcal{F}}[w] = \sup_{\tau} \text{ess}.\sup_{\omega} \mathbb{E}(w|\mathcal{F}_\tau)\mathbb{E}(w^{-1}|\mathcal{F}_\tau),$$

where the underlying filtration $\mathcal{F}$ is dyadic and the supremum is taken over all adapted stopping times $\tau$. Equivalently, without the use of stopping times, the $A_2$

characteristic can be defined by

$$Q_2^{\mathcal{F}}[w] = \sup_I \Big( \frac{1}{|I|} \int w \Big) \Big( \frac{1}{|I|} \int w^{-1} \Big),$$

where the supremum runs over all dyadic intervals.

The area of sharp weighted estimates has seen substantial progress with new, beautiful proofs of Wittwer's result and its extensions to the time shifted martingales referred to as "dyadic shift" [22, 32]. Related, important questions in harmonic analysis, such as boundedness of the Beurling–Ahlfors transform [28], Hilbert transform [26], general Calderon–Zygmund operators [19, 21, 23] and beyond [6, 19] have been solved, beautifully advancing profound understanding of the objects at hand.

During the early days of weighted theory in harmonic analysis, before optimal weighted estimates were within reach, say, for the maximal operator or the Hilbert transform [18] similar questions were asked in probability theory, concerning stochastic processes with continuous-in-time filtrations [7, 20]. The difficulty that arises in the nonhomogenous setting, typically seen when these processes have jumps, were already observed back then and this restriction was made in one form or another in these papers. Certain basic facts about weights do not hold true for all jump processes, such as the classical self improvement of the $A_2$ characteristic of the weight [7]. Another obstacle typical for working with weights is the nonconvexity of the set inspired by the $A_2$ characteristic: $\{r, s \in \mathbb{R}_+ : 1 \leq rs \leq Q\}$ with $Q > 1$. Such continuity-in-space assumptions still appear regularly for these or other reasons when addressing weights; see [4, 25].

Wittwer's proof also uses the homogeneity that arises from the dyadic filtration where the underlying measure is Lebesgue in a subtle but crucial way. This homogeneity assumption has only recently been removed in the papers [31] and [21]. These authors work with discrete-in-time general filtrations with arbitrary underlying measure, where one martingale is a predictable multiplier of the other. The definition of differential subordination for discrete-time martingales was introduced by Burkholder [8]. It was later generalised by Bañuelos–Wang [5] to arbitrary continuous-time martingales where it reads as follows:

(1.1)        $Y$ differentially subordinate to $X$

$$:\Leftrightarrow [X, X]_t - [Y, Y]_t \text{ nonnegative and nondecreasing}$$

is only possible in very special cases, such as predictable multipliers of stochastic integrals—this passage is standard and explained in one of the early works of Burkholder on $L^p$ estimates for pairs of differentially subordinate martingales [8]. (In full generality, this unweighted $L^p$ problem was only much later resolved in [5].)

In this article, we tackle the sharp weighted estimate in full generality, using the notion of differential subordination of Bañuelos and Wang (1.1) and the martingale

$A_2$ characteristic

$$Q_2^{\mathcal{F}}[w] = \sup_\tau \mathrm{ess}\,.\,\sup_\omega (w)_\tau (w^{-1})_\tau.$$

We prove that for $L^2$ integrable Hilbert space valued martingales $X, Y$ with $Y$ differentially subordinate to $X$ there holds

$$\|Y\|_{L^2(w)} \lesssim Q_2^{\mathcal{F}}[w] \|X\|_{L^2(w)},$$

where the implied constant is numeric and does not depend upon the dimension, the pair of martingales or the weight. The linear growth in the quantity $Q_2^{\mathcal{F}}(w)$ is sharp.

The proof in this paper is different from the proofs in [21] and [31]. In [21], so-called sparse operators are used while in [31] the authors reduce the estimate through the use of the so-called outer measure space theory.

Our approach is the following. We derive an explicit Bellman function of four variables adapted to the problem. It has certain conditions on its range, a continuous sub-convexity as well as discrete one-leg convexity, such as seen in [31] for two smaller Bellman functions (their functions make up a part of ours). We heavily use the explicit form of the Bellman function and its regularity properties in several parts in our proof to handle the delicacy of the continuous-in-time processes with values in Hilbert space. The resulting function is in the "dualized" or "weak form", which is in a contrast to the "strong form" of a Burkholder-type functional often seen when using the strong subordination condition (1.1). (The explicit form of a Burkholder-type functional for this weighted question is still open.) Indeed, the form of the strong differential subordination condition is adapted to work well for Burkholder-type functionals and arises naturally in this setting. The passage to its use in the weak form is accomplished through the use of the so-called el- lipse lemma and requires a Bellman function solving the entire problem at once as opposed to splitting the problem into pieces. This is the first use of this strategy for problems in probability and should allow generalisations of numerous existing results as well as an alternative (allbeit more complicated) proof of Wang's exten- sion to Burkholder's famous estimates using [33] or [3]. Note that for these $L^p$ problems, fewer difficulties arise, even in the presence of jumps. This is thanks to the convexity of the domain in the $L^p$ problem. The one-leg-convexity required to control the jumps is almost free in the $L^p$ case, when using a trick from [11]. This trick is not available here because of the nonconvex domain.

Our result gives through the formula in [2] a probabilistic proof of the weighted estimate for the Beurling–Ahlfors transform with its implication, a famous bor- derline regularity problem for the Beltrami equation, solved in [28]. Other appli- cations are discussed in the last section. They include a dimensionless weighted bound for discrete and semi-discrete second-order Riesz transforms. The Bell- man function constructed here has already been used to obtain a dimensionless weighted bound of the Bakry Riesz vector on Riemannian manifolds under a con- dition on curvature.

1.1. *Differentially subordinate martingales.* Consider first discrete-in-time martingales. For that, let $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ be a probability space with a nondecreasing sequence $\mathcal{F} = (\mathcal{F}_n)_{n \geq 0}$ of sub $\sigma$-fields of $\mathcal{F}_\infty$ such that $\mathcal{F}_0$ contains all $\mathcal{F}_\infty$-null sets. We are interested in $\mathbb{H}$-valued martingales, where $\mathbb{H}$ is a separable Hilbert space with norm $|\cdot|_\mathbb{H}$ and scalar product $\langle \cdot, \cdot \rangle_\mathbb{H}$: if $f = \{f_n\}_{n \in \mathbb{N}}$ is a $\mathbb{H}$-valued martingale adapted to $\mathcal{F}$, we note $f_n = \sum_{k=0}^{n} \mathrm{d} f_k$, with the convention $\mathrm{d} f_0 := f_0$, and $\mathrm{d} f_k := f_k - f_{k-1}$, for $k \geq 1$. Similarly, if $g$ is another adapted $\mathbb{H}$-valued martingale, we note $g_n = \sum_{k=0}^{n} \mathrm{d} g_k$ with the same conventions. One says that $g$ is differentially subordinate to $f$ if one has almost surely for $k \geq 0$, $|\mathrm{d} g_k|_\mathbb{H} \leq |\mathrm{d} f_k|_\mathbb{H}$.

In this paper, we consider continuous-in-time filtrations. Let us introduce again $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ a probability space with a nondecreasing right continuous family $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ of sub $\sigma$-fields of $\mathcal{F}_\infty$ such that $\mathcal{F}_0$ contains all $\mathcal{F}_\infty$-null sets. We are interested in $\mathbb{H}$-valued càdlàg martingales, where $\mathbb{H}$ is a separable Hilbert space. In order to clearly define differential subordination in this setting, we make use of the square bracket or quadratic variation process.

Recall that the quadratic variation process of a real-valued semimartingale $X$ is the process denoted by $[X, X] := ([X, X]_t)_{t \geq 0}$ and defined as (see, e.g., Protter [29])

$$[X, X]_t = X_t^2 - 2 \int_0^t X_{s-} \, \mathrm{d} X_s,$$

where we have set $X_{0-} = 0$. Similarly, the quadratic covariation of two real-valued semimartingales $X$ and $Y$ is the following process also known as the bracket process:

$$[X, Y]_t := X_t Y_t - \int_0^t X_{s-} \, \mathrm{d} Y_s - \int Y_{s-} \, \mathrm{d} X_s.$$

In the Hilbert-space-valued setting, one identifies $\mathbb{H}$ with $\ell_2$ and defines $[X, X] = \sum_{n=1}^{\infty} [X^n, X^n]$, where $X^n$ is the $n$th coordinate of $X$.

DEFINITION 1 (Differential subordination). Let $X$ and $Y$ two adapted càdlàg semimartingales taking values in a separable Hilbert space. We say that $Y$ is differentially subordinate by quadratic variation to $X$ iff

$$[X, X]_t - [Y, Y]_t$$

is a nondecreasing and nonnegative function of $t \geq 0$.

Let us denote by $X^c$ the unique continuous part of $X$ with

$$[X, X]_t = |X_0|^2 + [X^c, X^c]_t + \sum_{0 < s \leq t} |\Delta X_s|^2,$$

where $\Delta X_t := X_t - X_{t-}$. There holds $[X, X]_t^c = [X^c, X^c]_t$ and $\Delta [X, X]_t = |\Delta X_t|^2$. We have the following characterisation distinguishing the continuous and jump parts proved by Wang in [34], Lemma 1.

LEMMA 1. *If $X$ and $Y$ are semimartingales, then $Y$ is differentially subordinate to $X$ if and only if* (i) *$[X, X]_t^c - [Y, Y]_t^c$ is a nonnegative and nondecreasing function of $t$,* (ii) *the inequality $|\Delta Y_t| \leq |\Delta X_t|$ holds for all $t > 0$ and* (iii) *$|Y_0| \leq |X_0|$.*

1.2. *Martingales in nonhomogeneous weighted spaces.* Let again $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ be a probability space with a nondecreasing right continuous family $\mathcal{F} := (\mathcal{F}_t)_{t \geq 0}$ of sub $\sigma$-fields of $\mathcal{F}_\infty$ such that $\mathcal{F}_0$ contains all $\mathcal{F}_\infty$-null sets. If $X$ and $Y$ are adapted càdlàg square integrable $\mathbb{H}$-valued martingales and $Y$ is $\mathbb{P}$-differentially-subordinate to $X$, then it is obvious that

$$\|Y\|_{L^2} \leq \|X\|_{L^2}. \tag{1.2}$$

Recall here that $\|X\|_{L^2} := \sup_t \|X_t\|_{L^2}$, where

$$\|X_t\|_{L^2}^2 := \mathbb{E}|X_t|^2 = \int_\Omega |X_t(\omega)|^2 \, d\mathbb{P}(\omega). \tag{1.3}$$

Assume again that $Y$ is differentially subordinate to $X$. We might insist on the underlying probability space at hand by saying in short that $X$ and $Y$ are $\mathbb{P}$-martingales and that $Y$ is $\mathbb{P}$-differentially-subordinate to $X$. The main concern of this paper is to obtain sharp inequalities similar to (1.2) under a change of law in the definition of the $L^2$-norm according to [10]. To be more precise, we ask the following.

QUESTION 1. Let $\mathbb{P}$ and $\mathbb{Q}$ such that $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ and $(\Omega, \mathcal{F}_\infty, \mathbb{Q})$ are two filtered probability spaces as described above. Does there exist a constant $C_{\mathbb{P},\mathbb{Q}} > 0$ depending only on $(\mathbb{P}, \mathbb{Q})$ such that if $X$ and $Y$ are uniformly integrable $\mathbb{P}$-martingales adapted to $\mathcal{F}$ and $Y$ is $\mathbb{P}$-differentially-subordinate to $X$, then

$$\|Y\|_{L^2(d\mathbb{Q})} \leq C_{\mathbb{P},\mathbb{Q}} \|X\|_{L^2(d\mathbb{Q})}?$$

For this purpose, let $w$ be a positive, uniformly integrable martingale (that we often identify with its closure $w_\infty$) that we call a weight. This allows us to consider the weighted norms $L^2(w)$ with $\|X\|_{L^2(w)} := (\mathbb{E}(X_\infty^2 w))^{1/2}$. Given $X$ and $Y$ two $\mathbb{P}$-martingales as in the question above, we will seek a constant $C_w$ such that

$$\|X\|_{L^2(w)} \leq C_w \|Y\|_{L^2(w)}. \tag{1.4}$$

In the case where $d\mathbb{Q} = w \, d\mathbb{P}$ is a new probability measure, that is, the $\mathbb{P}$-mean of $w$ equals 1, the inequality above answers the question with $C_{\mathbb{P},\mathbb{Q}} = C_w$. It is well known that the class of weights allowing such estimates is the so-called $A_2$ class.

DEFINITION 2 ($A_2$ class). Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space. We say that the weight $w > 0$, a locally integrable function, is in the $A_2$ class, iff the $A_2$ characteristic of the weight $w$, noted $Q_2^{\mathcal{F}}[w]$ and defined as

$$Q_2^{\mathcal{F}}[w] := \sup_\tau \text{ess.} \sup_\omega (w)_\tau (w^{-1})_\tau$$

with the first supremum running over all adapted stopping times, is finite.

We often write $Q_2^{\mathcal{F}}[w] := \sup_\tau \text{ess.}\sup_\omega w_\tau u_\tau$ where $u := w^{-1}$ is the inverse weight. Notice finally that both inequality (1.4) and the $A_2$ characteristic $Q_2^{\mathcal{F}}[w]$ of the weight $w$ are insensitive to the scaling $w \mapsto \lambda w$ by an arbitrary factor $\lambda > 0$. This is the reason why in the sequel we will prove (1.4) only assuming that the weight is in the $A_2$ class.

## 2. Statement of the main results.

THEOREM 1 (Differential subordination under change of law).   *Let $X$ and $Y$ be two adapted uniformly integrable càdlàg $\mathbb{H}$-valued martingales such that $Y$ is differentially subordinate to $X$. Let $w$ be a weight in the $A_2$ class. Then*

$$\|Y\|_{L^2(w)} \lesssim Q_2^{\mathcal{F}}[w]\|X\|_{L^2(w)}$$

*and the linear growth in $Q_2^{\mathcal{F}}[w]$ is sharp.*

The upper estimate of this result will be a consequence of the following bilinear estimate.

PROPOSITION 1 (Bilinear estimate).   *Let $X$ and $Y$ be two adapted uniformly integrable càdlàg $\mathbb{H}$-valued martingales such that $Y$ is differentially subordinate to $X$. Let $w$ an admissible weight in the $A_2$ class. Then*

$$\mathbb{E}\int_0^\infty |\mathrm{d}[Y,Z]_t| \lesssim Q_2^{\mathcal{F}}[w]\|X\|_{L^2(w)}\|Z\|_{L^2(u)}.$$

The proof of the above proposition will be based on the Bellman function method. Let us fix some notation. Let us denote by $V$ the quadruplet

$$V := (x,y,r,s) \in \mathbb{H} \times \mathbb{H} \times \mathbb{R}_+^* \times \mathbb{R}_+^* =: \mathbb{S}.$$

The variables $(x,y)$ will be associated to $\mathbb{H}$-valued martingales whereas the variables $(r,s)$ to $\mathbb{R}$-valued martingales for the weights. We introduce the domain

$$\mathcal{D}_Q := \{V \in \mathbb{S} : 1 \leq rs \leq Q\}.$$

We will often restrict our attention to truncated weights, that is given $0 < \varepsilon < 1$, we consider weights $w$ such that $\varepsilon \leq w \leq \varepsilon^{-1}$. This means the variables $r$ and $s$ are bounded below and above and the domain becomes

$$\mathcal{D}_Q^\varepsilon := \left\{V \in \mathcal{D}_Q : \varepsilon \leq r \leq \varepsilon^{-1}, \varepsilon \leq s \leq \varepsilon^{-1}\right\}.$$

LEMMA 2 (Existence and properties of the Bellman function).   *There exists a function $B(V) = B_Q$, that is, $\mathcal{C}^1$ on $\mathcal{D}_Q^\varepsilon$, and piecewise $\mathcal{C}^2$, with the estimate*

$$B(V) \lesssim \frac{|x|^2}{r} + \frac{|y|^2}{s}$$

*and on each sub-domain where it is $C^2$ there holds*

$$\mathrm{d}^2 B \geq \frac{2}{Q} |\mathrm{d}x||\mathrm{d}y|.$$

*Whenever $V$ and $V_0$ are in the domain, the function has the property*

$$B(V) - B(V_0) - \mathrm{d}B(V_0)(V - V_0) \geq \frac{2}{Q} |x - x_0||y - y_0|.$$

*Moreover, we have the estimates*

$$|(\partial_x^2 B \, \mathrm{d}x, \mathrm{d}x)| \lesssim \varepsilon^{-1} |\mathrm{d}x|^2, \qquad |(\partial_y^2 B \, \mathrm{d}y, \mathrm{d}y)| \lesssim \varepsilon^{-1} |\mathrm{d}y|^2$$

*with the implied constants independent of $V$ and $(\mathrm{d}x, \mathrm{d}y)$.*

We have an explicit expression of the function described in Lemma 2. This is, aside from Theorem 1, one of the main results of this paper.

## 3. Existence and properties of the Bellman function.

PROOF OF LEMMA 2 (Existence and properties of the Bellman function).    We give an explicit expression for such a function. Let $V = (x, y, r, s)$ and $W = (r, s)$. We first consider

$$B_1(x, y, r, s) = \frac{\langle x, x \rangle}{r} + \frac{\langle y, y \rangle}{s}.$$

Then trivially $0 \leq B_1 \leq \frac{\langle x, x \rangle}{r} + \frac{\langle y, y \rangle}{s}$ and

$$
\begin{aligned}
(\mathrm{d}^2 B_1 \, \mathrm{d}V, \mathrm{d}V) &= \frac{2}{r} \langle \mathrm{d}x, \mathrm{d}x \rangle + \frac{2\langle x, x \rangle}{r^3} (\mathrm{d}r)^2 - 4\frac{\langle x, \mathrm{d}x \rangle}{r^2} \mathrm{d}r \\
&\quad + \frac{2}{s} \langle \mathrm{d}y, \mathrm{d}y \rangle + \frac{2\langle y, y \rangle}{s^3} (\mathrm{d}s)^2 - 4\frac{\langle y, \mathrm{d}y \rangle}{s^2} \mathrm{d}s \\
&= \frac{2}{r} \Big\langle \mathrm{d}x - \frac{x}{r} \mathrm{d}r, \mathrm{d}x - \frac{x}{r} \mathrm{d}r \Big\rangle \\
&\quad + \frac{2}{s} \Big\langle \mathrm{d}y - \frac{y}{s} \mathrm{d}s, \mathrm{d}y - \frac{y}{s} \mathrm{d}s \Big\rangle \\
&\geq 0.
\end{aligned}
$$

Letting $V_0 = (x_0, y_0, r_0, s_0)$ and $V = (x, y, r, s)$ we also calculate

$$
\begin{aligned}
&-\big(B_1(V_0) - B_1(V) + \mathrm{d}B_1(V_0)(V - V_0)\big) \\
&= -\left( \frac{x_0^2}{r_0} - \frac{x^2}{r} + \frac{2x_0}{r_0}(x - x_0) - \frac{x_0^2}{r_0^2}(r - r_0) \right)
\end{aligned}
$$

$$-\left(\frac{y_0^2}{s_0} - \frac{y^2}{s} + \frac{2y_0}{s_0}(y - y_0) - \frac{y_0^2}{s_0^2}(s - s_0)\right)$$

$$= r\left\langle \frac{x}{r} - \frac{x_0}{r_0}, \frac{x}{r} - \frac{x_0}{r_0}\right\rangle + s\left\langle \frac{y}{s} - \frac{y_0}{s_0}, \frac{y}{s} - \frac{y_0}{s_0}\right\rangle.$$

We now consider the two functions from [31],

$$K(r, s) = \frac{\sqrt{rs}}{\sqrt{Q}}\left(1 - \frac{\sqrt{rs}}{8\sqrt{Q}}\right),$$

$$N(r, s) = \frac{\sqrt{rs}}{\sqrt{Q}}\left(1 - \frac{(rs)^{3/2}}{128Q^{3/2}}\right)$$

in the domain $1 \le rs \le Q$. We have

$$0 \le K \le \left(1 - \frac{1}{8\sqrt{Q}}\right)\frac{\sqrt{rs}}{\sqrt{Q}} < \frac{\sqrt{rs}}{\sqrt{Q}} \le 1,$$

$$0 \le N \le \left(1 - \frac{1}{128Q^{3/2}}\right)\frac{\sqrt{rs}}{\sqrt{Q}} < \frac{\sqrt{rs}}{\sqrt{Q}} \le 1,$$

so in particular $rs \ge rs - K^2 > rs(1 - \frac{1}{Q})$. One calculates that

$$-(\mathrm{d}^2 K \, \mathrm{d}W, \mathrm{d}W) \ge \frac{1}{8Q}|\mathrm{d}r||\mathrm{d}s|,$$

$$-(\mathrm{d}^2 N \, \mathrm{d}W, \mathrm{d}W) \gtrsim \frac{1}{Q^2}s^2(\mathrm{d}r)^2,$$

$$-(\mathrm{d}^2 N \, \mathrm{d}W, \mathrm{d}W) \gtrsim \frac{1}{Q^2}r^2(\mathrm{d}s)^2.$$

Furthermore, if $W$ and $W_0$ belong to the domain of $K$ or $N$, then one also has

$$K(W_0) - K(W) + \mathrm{d}K(W_0)(W - W_0) \gtrsim \frac{1}{Q}|r - r_0||s - s_0|,$$

$$N(W_0) - N(W) + \mathrm{d}N(W_0)(W - W_0) \gtrsim \frac{1}{Q^2}s_0s|r - r_0|^2,$$

$$N(W_0) - N(W) + \mathrm{d}N(W_0)(W - W_0) \gtrsim \frac{1}{Q^2}r_0r|s - s_0|^2.$$

These remarkable one-leg concavity properties were proven in [31].

Let now

$$B_2 = \frac{\langle x, x\rangle}{2r - \frac{1}{s(N(r,s)+1)}} + \frac{\langle y, y\rangle}{s} = \frac{\langle x, x\rangle}{r + M(r, s)} + \frac{\langle y, y\rangle}{s},$$

where

$$M(r, s) = r - \frac{1}{s(N(r, s) + 1)}.$$

We will need an appropriate lower bound for the Hessian of $B_2$. One easily checks that the functions

$$F(x, r, M) = \frac{\langle x, x \rangle}{r + M},$$

$$G(r, s, N) = \frac{1}{s(N + 1)}$$

are convex, directly from the analysis of their Hessians. In order to estimate the Hessian of $B_2$ from below, one merely requires estimates of derivatives:

$$-\partial_M F = \frac{\langle x, x \rangle}{(r + M)^2} \geq \frac{\langle x, x \rangle}{4r^2} \quad \text{and} \quad -\partial_N G = \frac{1}{s(N + 1)^2} \geq \frac{1}{4s}.$$

Since $0 \leq r - \frac{1}{s(N(r,s)+1)} \leq r$, we know $0 \leq B_2 \leq \frac{|x|^2}{r} + \frac{|y|^2}{s}$. Therefore,

$$\left( d^2 B_2 \, dV, dV \right)$$

$$\gtrsim \frac{\langle x, x \rangle}{4r^2} \frac{1}{s(N + 1)^2} \frac{1}{Q^2} |dr|^2 s^2 + \frac{2}{s} \left\langle dy - \frac{y}{s} ds, dy - \frac{y}{s} ds \right\rangle$$

$$\gtrsim \frac{|x|^2 s}{Q^2 r^2} |dr|^2 + \frac{2}{s} \left\langle dy - \frac{y}{s} ds, dy - \frac{y}{s} ds \right\rangle$$

$$\gtrsim \frac{|x|}{Qr} |dr| \left| dy - \frac{y}{s} ds \right|.$$

The function $B_2$ has the additional property

$$-\left( B_2(V_0) - B_2(V) + dB_2(V_0)(V - V_0) \right)$$

$$\gtrsim \frac{\langle x_0, x_0 \rangle}{Q^2 r_0^2} s(r - r_0)^2 + s \left\langle \frac{y}{s} - \frac{y_0}{s_0}, \frac{y}{s} - \frac{y_0}{s_0} \right\rangle.$$

Indeed, write

$$\frac{\langle x, x \rangle}{2r - \frac{1}{s(N(r,s)+1)}} = H(x, r, s, N(r, s)) \quad \text{with } H(x, r, s, N) = \frac{\langle x, x \rangle}{2r - \frac{1}{s(N+1)}}.$$

Observe that $H$ is convex and that

$$-\partial_N H \gtrsim \frac{\langle x, x \rangle}{Q^2 r^2 s}.$$

Setting $P_0 = (x_0, r_0, s_0, N_0)$ and $P = (x, r, s, N)$, we have by convexity of $H$ that $H(P) \geq H(P_0) + dH(P)(P - P_0)$. Equivalently,

$$H(P) - H(P_0)$$
$$- \partial_x H(P_0)(x - x_0) - \partial_r H(P_0)(r - r_0) - \partial_s H(P_0)(s - s_0)$$
$$\geq -\partial_N H(P_0)(N_0 - N).$$

Recall the one-leg concavity property of $N$:

$$N(r_0, s_0) - N(r, s) + \partial_r N(r_0, s_0)(r - r_0) + \partial_s N(r_0, s_0)(s - s_0)$$
$$\gtrsim \frac{1}{Q^2} s_0 s |r - r_0|^2.$$

Setting $N_0 = N(r_0, s_0)$, $N = N(r, s)$ and using the lower derivative estimate of $H$ and the chain rule, we obtain the following one-leg convexity for $B_2$:

$$B_2(V) - B_2(V_0) - dB_2(V_0)(V - V_0)$$
$$\gtrsim \frac{\langle x_0, x_0 \rangle}{Q^2 r_0^2} s |r - r_0|^2 + s \left\langle \frac{y}{s} - \frac{y_0}{s_0}, \frac{y}{s} - \frac{y_0}{s_0} \right\rangle.$$

Analogously,

$$B_3 = \frac{\langle x, x \rangle}{r} + \frac{\langle y, y \rangle}{2s - \frac{1}{r(N(r,s)+1)}}$$

has the same size estimates as well as

$$(d^2 B_3 \, dV, dV) \gtrsim \frac{|y|}{Qs} |ds| \left| dx - \frac{x}{r} dr \right|$$

and one-leg convexity

$$B_3(V) - B_3(V_0) - dB_3(V_0)(V - V_0)$$
$$\gtrsim \frac{\langle y_0, y_0 \rangle}{Q^2 s_0^2} r |s - s_0|^2 + r \left\langle \frac{x}{r} - \frac{x_0}{r_0}, \frac{x}{r} - \frac{x_0}{r_0} \right\rangle.$$

Let us now consider

$$H_4(x, y, r, s, K) = \sup_{a>0} \beta(a, x, y, r, s, K) = \sup_{a>0} \left( \frac{\langle x, x \rangle}{r + aK} + \frac{\langle y, y \rangle}{s + a^{-1}K} \right).$$

Testing for critical points gives

$$\partial_a \beta = -\frac{\langle x, x \rangle K}{(r + aK)^2} + \frac{\langle y, y \rangle K}{(as + K)^2}.$$

So $\partial_a \beta = 0$ if and only if

$$a = a' = \frac{|y| r - |x| K}{|x| s - |y| K}.$$

Since only $a > 0$ are admissible, we require that $|y|r - |x|K$ and $|x|s - |y|K$ have the same sign. To determine sign change of $\partial_a \beta$ at $a'$, consider

$$-\frac{|x|}{r + aK} + \frac{|y|}{as + K} = \frac{(|y|r - |x|K) - a(|x|s - |y|K)}{(r + aK)(as + K)}.$$

If the signs are negative, then the sign change is from negative to positive otherwise from positive the negative. For a maximum to be attained at $a' > 0$, we require that both numerator and denominator be positive. Then, if $K$ is relatively small, meaning $|y|r - |x|K$ and $|x|s - |y|K$ are positive, we have

$$H_4(x, y, r, s, K) = \beta(a', x, y, r, s)$$

$$= \frac{\langle x, x \rangle (|x|s - |y|K)}{r(|x|s - |y|K) + (|y|r - |x|K)K}$$

$$+ \frac{\langle y, y \rangle (|y|r - |x|K)}{s(|y|r - |x|K) + (|x|s - |y|K)K}$$

$$= \frac{\langle x, x \rangle s - 2|x||y|K + \langle y, y \rangle r}{rs - K^2}.$$

Observe that by the above considerations on $K$, the denominator is never 0. The case $|x| = 0$ or $|y| = 0$ corresponds to other parts of the domain, so when $K$ is small in the above sense, this function is in $\mathcal{C}^2$.

When $|y|r - |x|K \leq 0$ or $|x|s - |y|K \leq 0$, the supremum is attained at the boundary and $H_4 = \frac{\langle y, y \rangle}{s}$ or $H_4 = \frac{\langle x, x \rangle}{r}$. Thanks to the size restrictions on $K$, we never have both $|x|s - |y|K \leq 0$ and $|y|r - |x|K \leq 0$ unless $x, y = 0$. Indeed

$$|x|(|x|s - |y|K) + |y|(|y|r - |x|K)$$

$$= \frac{|x|^2}{r} - 2\frac{|x||y|}{rs}K + \frac{|y|^2}{s}$$

$$= \left(\frac{|x|}{\sqrt{r}} - \frac{|y|}{\sqrt{s}}\right)^2 + \frac{2|x||y|}{\sqrt{rs}}\left(1 - \frac{K}{\sqrt{rs}}\right).$$

With $1 - \frac{K}{\sqrt{rs}} > 0$, we see that the above is never negative and the vanishing of the quantity implies $x = y = 0$. If $|x|s - |y|K \leq 0$ and $|y|r - |x|K > 0$, then $\frac{\langle x, x \rangle}{r} < \frac{\langle y, y \rangle}{s}$ and $H_4 = \frac{\langle y, y \rangle}{s}$, if $|y|r - |x|K \leq 0$ and $|x|s - |y|K > 0$ then $H_4 = \frac{\langle x, x \rangle}{r}$.

Notice that when $\frac{\langle x, x \rangle}{r} = \frac{\langle y, y \rangle}{s}$ and $x, y \neq 0$ then $|y|r - |x|K > 0$ and $|x|s - |y|K > 0$. Indeed, we have seen $\frac{|x|^2}{r} - 2\frac{|x||y|}{rs}K + \frac{|y|^2}{s} > 0$. Thus $\frac{\langle x, x \rangle}{r} = \frac{\langle y, y \rangle}{s} > \frac{|x||y|}{rs}K$ and $|y|r - |x|K > 0$ and $|x|s - |y|K > 0$.

Thus $H_4 \in \mathcal{C}^2$ for these parts of the domain. We also see from these considerations that in order to see $H_4 \in \mathcal{C}^1$ we only need to check the cuts $|x|s - |y|K = 0$ and $|y|r - |x|K \geq 0$ as well as $|y|r - |x|K = 0$ and $|x|s - |y|K \geq 0$.

When $|y|r - |x|K > 0$ and $|x|s - |y|K > 0$ (we call this part of the domain $R_1$),

$$(\partial_x H_4, dx) = 2\frac{\langle dx, x \rangle}{|x|}\frac{|x|s - |y|K}{rs - K^2},$$

$$\partial_r H_4 = -\frac{(|x|s - |y|K)^2}{(rs - K^2)^2},$$

$$\partial_K H_4 = -2\frac{(|x|s - |y|K)(|y|r - |x|K)}{(rs - K^2)^2}.$$

We first prove that $\partial_x H_4$ is continuous throughout. Recall that we have to treat three regions: $R_2$, where $|y|r - |x|K > 0$ and $|x|s - |y|K \leq 0$, $R_3$, where $|x|s - |y|K > 0$ and $|y|r - |x|K \leq 0$ and $R_1$. Inside $R_2$, we have $H_4 = \frac{\langle y, y \rangle}{s}$, and thus $\partial_x H_4 = 0$. Inside $R_3$, we have $H_4 = \frac{\langle x, x \rangle}{r}$ and thus, $\partial_x H_4 = \frac{2\langle x, dx \rangle}{r}$. Inside $R_1$, we have

$$\partial_x H_4 = 2\frac{\langle dx, x \rangle}{|x|}\frac{|x|s - |y|K}{rs - K^2}$$

$$= 2\langle x, dx \rangle \frac{|x|s - |y|K}{r(|x|s - |y|K) + (|y|r - |x|K)K}.$$

Consider three cases, first, let us first approach a boundary point of $R_1$ from within $R_1$ such that $|y|r - |x|K = a > 0$ and $|x|s - |y|K = 0$. Let $0 < \varepsilon < a/2$ and assume $a - \varepsilon < |y|r - |x|K < a + \varepsilon$ and $0 < |x|s - |y|K < \varepsilon$. There holds $|\langle \partial_x H_4, dx \rangle| \leq 2|dx|\frac{\varepsilon}{rs - K^2} \lesssim \varepsilon|dx|$ since $rs - K^2$ is bounded below. Letting $\varepsilon \to 0$ shows continuity at this point. Second, let us approach a boundary point such that $|x|s - |y|K = a > 0$ and $|y|r - |x|K = 0$ from within $R_1$. Assume therefore again $0 < \varepsilon < a/2$, $a - \varepsilon < |x|s - |y|K < a + \varepsilon$ and $0 < |y|r - |x|K < \varepsilon$. We show there holds $(\partial_x H_4, dx) \lesssim \frac{\varepsilon}{a}|dx|$. Since

$$\frac{1}{r} - (|y|r - |x|K)K\frac{|x|s - |y|K}{r^2(|x|s - |y|K)^2}$$

$$\leq \frac{(|x|s - |y|K)}{r(|x|s - |y|K) + (|y|r - |x|K)K}$$

$$\leq \frac{1}{r},$$

we have

$$\left|\frac{2\langle x, dx \rangle(|x|s - |y|K)}{r(|x|s - |y|K) + (|y|r - |x|K)K} - \frac{2\langle x, dx \rangle}{r}\right|$$

$$\leq 2|\langle x, dx \rangle|\frac{(|y|r - |x|K)K}{r^2(|x|s - |y|K)}$$

$$\lesssim |x||dx|\frac{\varepsilon}{a}.$$

Since $0 < |y|r - |x|K < \varepsilon$ and $s, r, K$ controlled, one can deduce from $|x|s - |y|K \sim a$ that $|x| \sim a$. Last, let us approach $|y|r - |x|K = 0$ and $|x|s - |y|K = 0$. To this end, one can see that if $0 < |y|r - |x|K < \varepsilon$ and $0 < |x|s - |y|K < \varepsilon$ then $|x|, |y| \lesssim \varepsilon$, establishing continuity in the third case.

The derivative $\partial_r H_4$ is similar since the term $\frac{|x|s-|y|K}{rs-K^2}$ reappears as a square and in $R_3$ notice that $H_4 = \frac{\langle x,x \rangle}{r}$ so $\partial_r H_4 = -\frac{\langle x,x \rangle}{r^2}$. It is easy to see that the derivative $\partial_K H_4$ is zero in $R_2$ and $R_3$ as well as when approaching the boundary of $R_1$.

By symmetry in the variables, the function is in $\mathcal{C}^1$. As a consequence,

$$B_4(x, y, r, s) = H_4(x, y, r, s, K(r, s)) \in \mathcal{C}^1.$$

Function $H_4$ is a supremum of convex functions and, therefore, convex. It has been shown indirectly in [24] that $-\partial_K H_4 \geq 0$ everywhere and that in $R'_1 \subset R_1$ where $|y|r - 2|x|K > 0$ and $|x|s - 2|y|K > 0$ we have $-\partial_K H_4 \gtrsim \frac{|x||y|}{rs}$. We present an easier argument. Recall that

$$-\partial_K H_4$$
$$= 2\frac{(|x|s - |y|K)(|y|r - |x|K)}{(rs - K^2)^2}$$
$$= 2\frac{(|x|s - |y|K)(|y|r - |x|K)|x||y|}{(r(|x|s - |y|K) + K(|y|r - |x|K))(s(|y|r - |x|K) + K(|x|s - |y|K))}.$$

So $-\partial_K H_4 \geq c\frac{|x||y|}{rs}$ if

$$\frac{rs}{c} \geq K^2 + rs + \frac{Kr(|x|s - |y|K)}{|y|r - |x|K} + \frac{Ks(|y|r - |x|K)}{|x|s - |y|K}.$$

Now $K^2 \leq 1 \leq rs$ and when $|y|r - 2|x|K \geq 0$ then $|y|r - |x|K \geq |x|K$. Similarly, $|x|s - |y|K \geq |y|K$. So the last two terms are bounded by $\frac{Kr|x|s}{|x|K} + \frac{Ks|y|r}{|y|K} = 2rs$. So $c = 1/4$ works. In $R'_1$,

$$(\mathrm{d}^2 B_4 \,\mathrm{d}V, \mathrm{d}V) \geq 4\frac{|x||y|}{8rsQ}|\mathrm{d}r||\mathrm{d}s| = \frac{|x||y|}{2rsQ}|\mathrm{d}r||\mathrm{d}s|.$$

We need to add more functions with the good concavity for other $K$. Let

$$B_5 = \frac{\langle x, x \rangle}{2r - \frac{1}{s(K(r,s)+1)}} + \frac{\langle y, y \rangle}{s}.$$

Since $0 \leq r - \frac{1}{s(K(r,s)+1)} \leq r$, we know $0 \leq B_5 \leq \frac{|x|^2}{r} + \frac{|y|^2}{s}$. Now the Hessian estimate becomes

$$(\mathrm{d}^2 B_5 \,\mathrm{d}V, \mathrm{d}V) \geq \frac{\langle x, x \rangle}{4r^2}\frac{1}{s(K+1)^2}\frac{1}{8Q}|\mathrm{d}r||\mathrm{d}s| \geq \frac{|x|^2}{128Qsr^2}|\mathrm{d}r||\mathrm{d}s|.$$

So the function $B_5$ is convex and when $2|x|K \geq |y|r$, then

$$\left(\mathrm{d}^2 B_5 \, \mathrm{d}V, \mathrm{d}V\right) \geq \frac{|x||y|}{256K \, Qsr} |dr||ds| \geq \frac{|x||y|}{256Qsr} |dr||ds|.$$

If we set

$$B_6 = \frac{\langle x, x \rangle}{r} + \frac{\langle y, y \rangle}{2s - \frac{1}{r(K(r,s)+1)}},$$

then $B_6$ is convex, $0 \leq B_6 \leq \frac{|x|^2}{r} + \frac{|y|^2}{s}$ and when $2|y|K \geq |x|s$ then

$$\left(\mathrm{d}^2 B_6 \, \mathrm{d}V, \mathrm{d}V\right) \geq \frac{|x||y|}{256Qsr} |dr||ds|.$$

Putting the above facts together, we see that the function $B_7 = B_4 + B_5 + B_6$ satisfies

$$\left(\mathrm{d}^2 B_7 \, \mathrm{d}V, \mathrm{d}V\right) \gtrsim \frac{|x||y|}{Qsr} |dr||ds|.$$

Through similar considerations as above, we have discrete one-leg convexity

$$B_7(V) - B_7(V_0) - \mathrm{d}B_7(V_0)(V - V_0) \gtrsim \frac{|x_0||y_0|}{Qs_0 r_0} |r - r_0||s - s_0|.$$

Letting for appropriate fixed $c_i$,

(3.1) $$B = c_1 B_1 + c_2 B_2 + c_3 B_3 + c_7 B_7,$$

we obtain $0 \leq B \lesssim \frac{|x|^2}{r} + \frac{|y|^2}{s}$ and $\mathrm{d}^2 B \geq \frac{2}{Q} |dx||dy|$ in the regions where $B \in \mathcal{C}^2$. Indeed,

$$\left(\mathrm{d}^2 B_1 \, \mathrm{d}V, \mathrm{d}V\right) \geq \frac{4}{Q} |dx||dy| + \frac{4|x||y|}{Qrs} |dr||ds| - \frac{4|y|}{Qs} |dx||ds| - \frac{4|x|}{Qr} |dy||dr|,$$

$$\left(\mathrm{d}^2 B_2 \, \mathrm{d}V, \mathrm{d}V\right) \geq \frac{\sqrt{3}|x|}{2Qr} |dy||dr| - \frac{\sqrt{3}|x||y|}{2Qrs} |dr||ds|,$$

$$\left(\mathrm{d}^2 B_3 \, \mathrm{d}V, \mathrm{d}V\right) \geq \frac{\sqrt{3}|y|}{2Qs} |dx||ds| - \frac{\sqrt{3}|x||y|}{2Qrs} |dr||ds|,$$

$$\left(\mathrm{d}^2 B_7 \, \mathrm{d}V, \mathrm{d}V\right) \geq \frac{|x||y|}{256Qrs} |dr||ds|,$$

where the last inequality holds in the regions where the function $B_4 \in \mathcal{C}^2$. The weighted sum of these inequalities according to (3.1) yields the desired inequality on convexity. Now,

$$B_1(V) - B_1(V_0) - \mathrm{d}B_1(V_0)(V - V_0)$$

$$\gtrsim \frac{rs}{Q} \left| \frac{x}{r} - \frac{x_0}{r_0} \right| \left| \frac{y}{s} - \frac{y_0}{s_0} \right| \geq \frac{rs}{Q} \left| \left\langle \frac{x}{r} - \frac{x_0}{r_0}, \frac{y}{s} - \frac{y_0}{s_0} \right\rangle \right|,$$

$$B_2(V) - B_2(V_0) - \mathrm{d}B_2(V_0)(V - V_0)$$

$$\gtrsim \frac{s}{Q}\frac{|x_0|}{r_0}|r - r_0|\left|\frac{y}{s} - \frac{y_0}{s_0}\right| \geq \frac{s}{Q}\left|\left\langle\frac{x_0}{r_0}, \frac{y}{s} - \frac{y_0}{s_0}\right\rangle\right||r - r_0|,$$

$$B_3(V) - B_3(V_0) - \mathrm{d}B_3(V_0)(V - V_0)$$

$$\gtrsim \frac{r}{Q}\frac{|y_0|}{s_0}|s - s_0|\left|\frac{x}{r} - \frac{x_0}{r_0}\right| \geq \frac{r}{Q}\left|\left\langle\frac{x}{r} - \frac{x_0}{r_0}, \frac{y_0}{s_0}\right\rangle\right||s - s_0|,$$

$$B_7(V) - B_7(V_0) - \mathrm{d}B_7(V_0)(V - V_0)$$

$$\gtrsim \frac{1}{Q}|x_0||y_0||r - r_0||s - s_0| \geq \frac{1}{Q}|\langle x_0, y_0\rangle||r - r_0||s - s_0|.$$

Notice that the last inequalities also remain true when we replace $x$ by $\Theta x$ and $x_0$ by $\Theta x_0$ where the rotation $\Theta$ is chosen so that $\Theta(x - x_0)$ and $y - y_0$ have the same direction, and thus we may assume that $\langle x - x_0, y - y_0\rangle = |x - x_0||y - y_0|$.

Summing the above inequalities give

$$Q\big(B(V) - B(V_0) - \mathrm{d}B(V_0)(V - V_0)\big)$$

$$\gtrsim \left\langle\left(\frac{x}{r} - \frac{x_0}{r_0}\right)r, \left(\frac{y}{s} - \frac{y_0}{s_0}\right)s + \frac{y_0}{s_0}(s - s_0)\right\rangle$$

$$+ \left\langle\frac{x_0}{r_0}(r - r_0), \left(\frac{y}{s} - \frac{y_0}{s_0}\right)s + \frac{y_0}{s_0}(s - s_0)\right\rangle$$

$$= \left\langle\left(\frac{x}{r} - \frac{x_0}{r_0}\right)r, y - y_0\right\rangle + \left\langle\frac{x_0}{r_0}(r - r_0), y - y_0\right\rangle$$

$$= \langle x - x_0, y - y_0\rangle = |x - x_0||y - y_0|$$

and we have proved the one-leg convexity. It remains to bound the second derivatives in $x$ and $y$. Let $\varepsilon$ be the cut off of the weights so that $\varepsilon \leq r, s \leq \varepsilon^{-1}$. We calculate

$$\left(\partial_x^2 \frac{\langle x, x\rangle}{r}\mathrm{d}x, \mathrm{d}x\right) = \frac{2\langle \mathrm{d}x, \mathrm{d}x\rangle}{r} \lesssim \varepsilon^{-1}\langle \mathrm{d}x, \mathrm{d}x\rangle;$$

$$\left(\partial_x^2 \frac{\langle x, x\rangle}{r + M(r, s)}\mathrm{d}x, \mathrm{d}x\right) = \frac{2\langle \mathrm{d}x, \mathrm{d}x\rangle}{r + M(r, s)} \leq \frac{2\langle \mathrm{d}x, \mathrm{d}x\rangle}{r} \lesssim \varepsilon^{-1}\langle \mathrm{d}x, \mathrm{d}x\rangle;$$

$$\left(\partial_x^2 \frac{\langle x, x\rangle s - 2|x||y|K + \langle y, y\rangle r}{rs - K^2}\mathrm{d}x, \mathrm{d}x\right)$$

$$= \frac{2\langle \mathrm{d}x, \mathrm{d}x\rangle s}{rs - K^2} - \frac{2|y|K}{rs - K^2}\left(\frac{\langle \mathrm{d}x, \mathrm{d}x\rangle}{|x|} - \frac{\langle x, \mathrm{d}x\rangle^2}{|x|^3}\right) \lesssim \varepsilon^{-1}\langle \mathrm{d}x, \mathrm{d}x\rangle,$$

where the lower bound for $rs - K^2 > rs(1 - \frac{1}{Q}) \geq 1 - \frac{1}{Q}$ is used in the last implied constant. We also used $\langle x, \mathrm{d}x\rangle^2 \leq \langle x, x\rangle\langle \mathrm{d}x, \mathrm{d}x\rangle$ which yields $\frac{\langle \mathrm{d}x, \mathrm{d}x\rangle}{|x|} - \frac{\langle x, \mathrm{d}x\rangle^2}{|x|^3} \geq 0$.

These imply that for $V \in \mathcal{D}_{Q,\varepsilon}$ we have

(3.2) $$(\partial_x^2 B(V)\, dx, dx) \lesssim \varepsilon^{-1}|dx|^2.$$

This concludes the proof of Lemma 2.   $\square$

Convexities of the form $d^2 B(V) \geq 2|dx||dy|$ can be self-improved using the following interesting lemma.

LEMMA 3 (Ellipse lemma, Dragicevic–Treil–Volberg [15]).   *Let $\mathbb{H}$ be a Hilbert space with $A, B$ two positive definite operators on $\mathbb{H}$. Let $T$ be a self-adjoint operator on $\mathbb{H}$ such that*

$$(Th, h) \geq 2(Ah, h)^{1/2}(Bh, h)^{1/2}$$

*for all $h \in \mathbb{H}$. Then there exists $\tau > 0$ satisfying*

$$(Th, h) \geq \tau(Ah, h) + \tau^{-1}(Bh, h)$$

*for all $h \in \mathbb{H}$.*

For our specific Bellman function, we will need a quantitative version.

LEMMA 4 (Quantitative ellipse lemma for $B$).   *Let $V \in \mathcal{D}_Q^{\varepsilon}$. Assume moreover that $B$ is $C^2$ at $V$. Then there exists $\tau(V) > 0$ such that*

$$Q\, d_V^2 B(V) \geq \tau(V)|dx|^2 + (\tau(V))^{-1}|dy|^2.$$

*Moreover, we have the bound*

$$Q^{-1}\varepsilon \lesssim \tau(V) \lesssim Q\varepsilon^{-1}.$$

PROOF (Quantitative ellipse lemma for $B$).   Let $V \in \mathcal{D}_Q^{\varepsilon}$. We have already seen in Lemma 2 that

$$d_V^2 B(V) \geq \frac{2}{Q}|dx||dy|.$$

The ellipse lemma [15] implies the existence of $\tau(V)$ such that for all vectors $dx$ and $dy$ there holds

$$Q\, d_V^2 B(V) \geq \tau(V)|dx|^2 + (\tau(V))^{-1}|dy|^2.$$

We can estimate $\tau(V)$ by testing the Hessian on any $dV$ of the form $dV = (dx, 0, 0, 0)$,

$$\tau(V)|dx|^2 \leq Q(d_V^2 B(V)\, dV, dV) = Q(\partial_x^2 B(V)\, dx, dx) \lesssim Q\varepsilon^{-1}|dx|^2,$$

K. DOMELEVO AND S. PETERMICHL

where the last inequality follows from (3.2). Hence $\tau(V) \lesssim Q\varepsilon^{-1}$ as claimed. The same bound holds for $(\tau(V)^{-1})$ by testing against $dV = (0, dy, 0, 0)$. Finally, we have proved that for all $V \in \mathcal{D}_Q^\varepsilon$,

$$Q^{-1}\varepsilon \lesssim \tau(V) \lesssim Q\varepsilon^{-1}. \qquad \square$$

We now address the lack of smoothness of $B$. All functions aside from $H_4$ that appear are at least in $\mathcal{C}^2$. We apply a standard mollifying procedure via convolution with $\varphi_\ell$ directly on $H_4(x, y, r, s, K)$, now only taking *real* variables with $x$, $y$ positive, $1 < rs < Q$ and $0 < K < 1$. Here, $\varphi$ denotes a standard mollifying kernel in the five real variables $(x, y, r, s, K) \in \mathbb{R}^5$ with support in the corresponding unit ball, whereas $\varphi_\ell(\cdot) := \ell^{-5}\varphi(\cdot/\ell)$ denotes its scaled version with support of size $\ell$. By slightly changing the constructions, the upper and lower estimate on the product $rs$ can be modified at the cost of a multiplicative constant in the final estimate of the Bellman function. Also take into account that the weights are cut, therefore bounded above and below. Further, we will assume that the positive variables $x$ and $y$ be bounded below. These considerations give us enough room to smooth the function $H_4$. It is important that $H_4$ is at least in $\mathcal{C}^1$ and its second-order partial derivatives exist almost everywhere. So we have $d^2(H_4 * \varphi_\ell) = (d^2 H_4) * \varphi_\ell$. Last, we are observing that as long as the norms of vectors $|x|$ and $|y|$ are bounded away from 0, our function $H_4 * \varphi_\ell$, mollified in $\mathbb{R}^5$ remains smooth when taking vector variables (observe that the final Bellman function only depends upon $|x|$ and $|y|$). It is important that the smoothing happens before the function is composed with $K$, we therefore preserve fine convexity properties, in particular also the much needed one-leg convexity. Size estimates change slightly, but are recovered when the mollifying parameter goes to 0. These details are either standard and have appeared in numerous articles on Bellman functions, or an easy consequence of the above construction.

LEMMA 5 (Regularised Bellman function and its properties). *Let $\varepsilon > 0$ given. Let $0 < \ell \leq \varepsilon/2$. There exists a function $B_\ell(x, y, r, s)$ defined on the domain*

$$\mathcal{D}_Q^{\varepsilon, \ell} := \left\{V \in \mathcal{D}_Q^\varepsilon; |x| \geq \ell, |y| \geq \ell\right\} \subset \mathcal{D}_Q^\varepsilon$$

*such that for all $V_0, V \in \mathcal{D}_Q^{\varepsilon, \ell}$, we have*

$$B_\ell \lesssim (1 + \ell)\left(\frac{|x|^2}{r} + \frac{|y|^2}{s}\right),$$

(3.3)
$$d_V^2 B_\ell(V) \geq \frac{2}{Q}|dx||dy|,$$

$$B_\ell(V) - B_\ell(V_0) - d_V B_\ell(V_0)(V - V_0) \geq \frac{1}{Q}|\Delta x||\Delta y|$$
(3.4)
$$= \frac{1}{Q}|x - x_0||y - y_0|.$$

*Moreover, the quantitative ellipse lemma now holds in the form*

$$Q \, \mathrm{d}_V^2 B_\ell(V) \geq \tau_\ell(V)|\mathrm{d}x|^2 + (\tau_\ell(V))^{-1}|\mathrm{d}y|^2,$$

*where $\tau_\ell := \tau_\ell(V)$ is a continuous function of its arguments, and where*

$$Q^{-1}\varepsilon \lesssim \tau_\ell(V) \lesssim Q\varepsilon^{-1}.$$

**4. Dissipation estimates.** Let $V := (X, Z, u, w)$ a càdlàg adapted martingale with values in $\mathcal{D}_Q^\varepsilon$. In order to bound the $\mathbb{H}$-valued martingale $X := (X^1, X^2, \ldots)$ away from 0, it is classical to introduce the $\mathbb{R} \times \mathbb{H}$-valued martingales $X^a := (a, X^1, X^2, \ldots)$ where $a > 0$. It follows that $\|X^a\|^2 = \|X\|^2 + a^2$ and $\|X^a\| \geq a$, and the same construction holds for $Z$. We note $V^a := (X^a, Z^a, u, w)$. Given a smoothing parameter $\ell > 0$, take $a \geq \ell$ then it follows that

$$V \in \mathcal{D}_Q^\varepsilon \quad \Rightarrow \quad V^a \in \mathcal{D}_Q^{\varepsilon,\ell}.$$

The main result of this section is the following dissipation estimate.

PROPOSITION 2 (Dissipation estimate). *Let $\varepsilon > 0$, $\ell > 0$ as defined above. Let $V$ a càdlàg adapted martingale with $V \in \mathcal{D}_Q^\varepsilon$. Let finally $F_t := \mathbb{E}(|X_\infty|^2 w_\infty^\varepsilon|\mathcal{F}_t)$ and $G_t := \mathbb{E}(|Z_\infty|^2 u_\infty^\varepsilon|\mathcal{F}_t)$. Let finally $a \geq \ell$. We have*

$$Q(1+\ell)\big(\mathbb{E}F_t + \mathbb{E}G_t + 2a^2\varepsilon^{-1}\big)$$
$$\gtrsim \frac{1}{2}\mathbb{E}\int_0^t \tau_\ell(V_{s-}) \, \mathrm{d}[X,X]_s^c + (\tau_\ell(V_{s-}))^{-1} \, \mathrm{d}[Z,Z]_s^c + \mathbb{E}\sum_{0<s\leq t}|\Delta X_s||\Delta Z_s|.$$

We need the preliminary lemma.

LEMMA 6 (Comparison of quadratic forms in stochastic integrals). *Let $\mathcal{Q}$ denote the set of quadratic forms from $\mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ and $A := (A_{\alpha\beta})_{1\leq\alpha,\beta\leq m}$ and $B := (B_{\alpha\beta})_{1\leq\alpha,\beta\leq m}$ two $\mathcal{Q}$-valued càdlàg processes. Assume for all $t \geq 0$ and a.s. that $A(t) \geq B(t)$ [resp., $A(t) \geq |B(t)|$], in the sense that*

$$(A \, \mathrm{d}V, \mathrm{d}V) \geq (B \, \mathrm{d}V, \mathrm{d}V) \qquad [resp., (A \, \mathrm{d}V, \mathrm{d}V) \geq |(B \, \mathrm{d}V, \mathrm{d}V)|]$$

*for all $\mathrm{d}V \in \mathbb{R}^m$. Abbreviating $A_{s-} : \mathrm{d}[V,V]_s = \sum_{\alpha,\beta}(A_{\alpha\beta})_{s-} \, \mathrm{d}[V_\alpha, V_\beta]_s$ we have for all $t \geq 0$,*

$$\mathbb{E}\int_0^t A_{s-} : \mathrm{d}[V,V]_s \geq \mathbb{E}\int_0^t B_{s-} : \mathrm{d}[V,V]_s$$
$$\left(resp., \mathbb{E}\int_0^t A_{s-} : \mathrm{d}[V,V]_s \geq \mathbb{E}\int_0^t |B_{s-} : \mathrm{d}[V,V]_s|\right).$$

PROOF (Comparison of quadratic forms in stochastic integrals).    Under the hypotheses above, let us consider the case $A(t) \geq B(t)$, the case $A(t) \geq |B(t)|$ being treated in the same manner. Given $t \geq 0$, assume that

$$\int_0^t A_{s-} : \mathrm{d}[V, V]_s = \sum_{\alpha, \beta} \int_0^t (A_{\alpha\beta})_{s-} \, \mathrm{d}[V_\alpha, V_\beta]_s < \infty,$$

otherwise the claim is proved. Given the process $V$, let $\sigma_n := (0 \leq T_0^n \leq T_1^n \leq \cdots \leq T_i^n \leq \cdots \leq T_{k_n}^n \leq t)$ denote a random partition of stopping times tending to the identity as $n$ tends to infinity. Given $\alpha$ and $\beta$, we have that $A_{\alpha\beta}$ is a $\mathbb{R}$-valued càdlàg process. It follows (see, e.g., Protter [29]) that the stochastic integral

$$(4.1) \qquad \int_0^t A_{\alpha\beta}(s-) \, \mathrm{d}[V_\alpha, V_\beta]_s$$

is the limit in *u.c.p.* (uniform convergence in probability) as $n$ tends to infinity of sums

$$S_{\alpha\beta}^A := \sum_{i=0}^{k_n-1} A_{\alpha\beta}(T_i^n)(V_\alpha^{T_{i+1}^n} - V_\alpha^{T_i^n})(V_\beta^{T_{i+1}^n} - V_\beta^{T_i^n})$$

involving the stopping times defined above. Since $A \geq B$, summing w.r.t. $\alpha, \beta$ yields, for any $s \in [0, t]$,

$$\left( \sum_{\alpha, \beta} S_{\alpha\beta}^A \right)(s) := \sum_{\alpha, \beta} \sum_{i=0}^{k_n-1} A_{\alpha\beta}(T_i^n)(V_{\alpha,s}^{T_{i+1}^n} - V_{\alpha,s}^{T_i^n})(V_{\beta,s}^{T_{i+1}^n} - V_{\beta,s}^{T_i^n})$$

$$= \sum_{i=0}^{k_n-1} \sum_{\alpha, \beta} A_{\alpha\beta}(T_i^n)(V_{\alpha,s}^{T_{i+1}^n} - V_{\alpha,s}^{T_i^n})(V_{\beta,s}^{T_{i+1}^n} - V_{\beta,s}^{T_i^n})$$

$$\geq \sum_{i=0}^{k_n-1} \sum_{\alpha, \beta} B_{\alpha\beta}(T_i^n)(V_{\alpha,s}^{T_{i+1}^n} - V_{\alpha,s}^{T_i^n})(V_{\beta,s}^{T_{i+1}^n} - V_{\beta,s}^{T_i^n})$$

$$\geq \left( \sum_{\alpha, \beta} S_{\alpha\beta}^B \right)(s)$$

with an obvious definition for $S_{\alpha\beta}^B$. Passing to the limit in the sums $\sum_{\alpha, \beta}$ gives the result. $\square$

PROOF OF PROPOSITION 2 (Dissipation estimates).
*Step 1.* We first pass to a finite dimensional case. Let $V$ a càdlàg adapted martingale with $V \in \mathcal{D}_Q^\varepsilon$. Then $V^a \in \mathcal{D}_Q^{\varepsilon, \ell}$. We note $X^{a,m}$ the projection of $X^a \in \mathbb{R} \times \mathbb{H}$ onto $\mathbb{R} \times \mathbb{R}^m$, and introduce accordingly $Z^{a,m}$ and $V^{a,m}$. Notice

that $[X^a, X^a] = a^2 + [X, X]$ and similarly $[X^{a,m}, X^{a,m}] = a^2 + [X^m, X^m]$. Since $V^{a,m} \in \mathcal{D}_Q^{\varepsilon,\ell}$ where $B_\ell$ is $\mathcal{C}^2$ and we can apply Itô's formula and obtain, for all $t > 0$, almost sure paths,

$$
\begin{aligned}
B_\ell(V_t^{a,m}) = {} & B_\ell(V_0^{a,m}) \\
& + \int_{0+}^t \mathrm{d}_V B(V_{s-}^{a,m}) \, \mathrm{d}V_s^m + \frac{1}{2} \int_{0+}^t \mathrm{d}_V^2 B_\ell(V_{s-}^{a,m}) : \mathrm{d}[V^m, V^m]_s^c \\
& + \sum_{0 < s \leq t} \left\{ B_\ell(V_s^{a,m}) - B_\ell(V_{s-}^{a,m}) - \mathrm{d}_V B_\ell(V_{s-}^{a,m}) \Delta V_s^m \right\}.
\end{aligned}
$$

Thanks to Lemma 4 and Lemma 6, the concavity properties (3.3) of $B_\ell$ imply for the continuous part

$$
\begin{aligned}
& \frac{1}{2} \int_{0+}^t \mathrm{d}_V^2 B_\ell(V_{s-}^{a,m}) : \mathrm{d}[V^m, V^m]_s^c \\
& \quad \geq \frac{1}{2Q} \int_{0+}^t \tau_\ell(V_{s-}^{a,m}) \, \mathrm{d}[X^m, X^m]_s^c + \left( \tau_\ell(V_{s-}^{a,m}) \right)^{-1} \mathrm{d}[Z^m, Z^m]_s^c.
\end{aligned}
$$

Also, the concavity properties (3.4) of $B_\ell$ for the jump part

$$
B_\ell(V_s^{a,m}) - B_\ell(V_{s-}^{a,m}) - \mathrm{d}_V B_\ell(V_{s-}^{a,m}) \Delta V_s^m \geq \frac{1}{Q} |\Delta X_s^m| |\Delta Z_s^m|.
$$

Plugging the continuous and jump dissipation estimates into Itô's formula yields for all times, almost sure paths,

$$
\begin{aligned}
B_\ell(V_t^{a,m}) \geq {} & B_\ell(V_0^{a,m}) + \int_{0+}^t \mathrm{d}_V B_\ell(V_{s-}^{a,m}) \, \mathrm{d}V_s^m \\
& + \frac{1}{2Q} \int_0^t \tau_\ell(V_{s-}^{a,m}) \, \mathrm{d}[X^m, X^m]_s^c + \left( \tau_\ell(V_{s-}^m) \right)^{-1} \mathrm{d}[Z^m, Z^m]_s^c \\
& + \frac{1}{Q} \sum_{0 < s \leq t} |\Delta X_s^m| |\Delta Z_s^m|.
\end{aligned}
$$

*Step 2.* For technical reasons in the proof, we work with bounded martingales that we obtain through a usual stopping procedure. Recall that $V$ is a càdlàg adapted martingale with $V \in \mathcal{D}_Q^\varepsilon$ and $V^a \in \mathcal{D}_Q^{\varepsilon,\ell}$. For all $M \in \mathbb{N}$, define the stopping time $T_M := \inf\{t > 0; |V^a|_t^2 + [V^a, V^a]_t > M^2\}$, so that $T_M$ is a stopping time that tends to infinity as $M$ goes to infinity. It follows that $V^{a,T_M}$ is a local martingale, and that $V^{a,T_M-}$ and $[V^a, V^a]^{T_M-}$ are bounded semimartingales. Let $m \in \mathbb{N}^\star$ and $V^{a,m}$ the projection of $V^a$ onto $\mathbb{R}^m \subset \mathbb{H}$. For each $M$, there exists a sequence $\{T_{M,k}\}_{k \geq 1}$ of stopping times such that $T_{M,k} \uparrow T_M$ as $k \uparrow \infty$, and such that $(V^{a,m})^{T_{M,k}}$ is a martingale. Since $|V^{a,m}| \leq |V^a|$, it follows that $(V^{a,m})^{T_{M,k}-}$ is a bounded semimartingale, to which we can apply the dissipation estimate of

Step 1 above and obtain

$$
B_\ell\big(V^{a,m}_{t\wedge T_{M,k}-}\big)
$$

$$
\geq B_\ell\big(V^{a,m}_0\big) + \int_{0+}^{t\wedge T_{M,k}-} \mathrm{d}_V B_\ell\big(V^{a,m}_{s-}\big)\, \mathrm{d}V^m_s
$$

$$
+ \frac{1}{2Q}\int_0^{t\wedge T_{M,k}-} \tau_\ell\big(V^{a,m}_{s-}\big)\,\mathrm{d}[X^m, X^m]^c_s + \big(\tau_\ell\big(V^{a,m}_{s-}\big)\big)^{-1}\,\mathrm{d}[Z^m, Z^m]^c_s
$$

$$
+ \frac{1}{Q}\sum_{0<s<t\wedge T_{M,k}} |\Delta X^m_s||\Delta Z^m_s|
$$

$$
= B_\ell\big(V^{a,m}_0\big) + \int_{0+}^{t\wedge T_{M,k}} \mathrm{d}_V B_\ell\big(V^{a,m}_{s-}\big)\, \mathrm{d}V^m_s
$$

$$
+ \frac{1}{2Q}\int_0^{t\wedge T_{M,k}-} \tau_\ell\big(V^{a,m}_{s-}\big)\,\mathrm{d}[X^m, X^m]^c_s + \big(\tau_\ell\big(V^{a,m}_{s-}\big)\big)^{-1}\,\mathrm{d}[Z^m, Z^m]^c_s
$$

$$
+ \frac{1}{Q}\sum_{0<s<t\wedge T_{M,k}} |\Delta X^m_s||\Delta Z^m_s| - \mathrm{d}_V B_\ell\big(V^{a,m}_{t\wedge T_{M,k}-}\big)\Delta V^m_{t\wedge T_{M,k}}.
$$

Taking expectation and then letting $k \to \infty$, the dominated convergence theorem yields

$$
\mathbb{E}B_\ell\big(V^{a,m}_{t\wedge T_M-}\big)
$$

$$
\geq \mathbb{E}B_\ell\big(V^{a,m}_0\big)
$$

(4.2)
$$
+ \frac{1}{2Q}\mathbb{E}\int_0^{t\wedge T_M} \tau_\ell\big(V^{a,m}_{s-}\big)\,\mathrm{d}[X^m, X^m]^c_s + \big(\tau_\ell\big(V^{a,m}_{s-}\big)\big)^{-1}\,\mathrm{d}[Z^m, Z^m]^c_s
$$

$$
+ \frac{1}{Q}\mathbb{E}\sum_{0<s<t\wedge T_M} |\Delta X^m_s||\Delta Z^m_s| - \mathbb{E}\{\mathrm{d}_V B_\ell\big(V^{a,m}_{t\wedge T_M-}\big)\Delta V^m_{t\wedge T_M}\}.
$$

Observe that we used size properties of $B_\ell$, the definition of the stopping time $T_{M,k}$, the estimate of the $\tau_\ell$ provided by Lemma 5 and the size control of the weights.

*Step 3.* Now, we wish to return to the infinite dimensional case. First, recall that $0 \leq B_\ell(V) \lesssim (1+\ell)(X^2/u + Y^2/w)$. Let $F_t := \mathbb{E}(|X_\infty|^2 w_\infty|\mathcal{F}_t)$, $G_t := \mathbb{E}(|Z_\infty|^2 u_\infty|\mathcal{F}_t)$ and $F^a_t := \mathbb{E}(|X^a_\infty|^2 w_\infty|\mathcal{F}_t)$, $G^a_t := \mathbb{E}(|Z^a_\infty|^2 u_\infty|\mathcal{F}_t)$. Notice that $F^a_t = \mathbb{E}((|X_\infty|^2 + a^2)w_\infty|\mathcal{F}_t) \leq F_t + \mathbb{E}(a^2 w^\varepsilon_\infty|\mathcal{F}_t) \leq F_t + a^2\varepsilon^{-1}$. It follows, thanks to Jensen inequality, that

$$
B_\ell\big(V^a_t\big) \leq C_0(1+\ell)\big(F^a_t + G^a_t\big) \lesssim (1+\ell)\big(F_t + G_t + 2a^2\varepsilon^{-1}\big).
$$

A similar inequality holds for $V^{a,m}$ and in particular

$$
B_\ell\big(V^{a,m}_{t\wedge T_M-}\big) \lesssim (1+\ell)\big(F_{t\wedge T_M-} + G_{t\wedge T_M-} + 2a^2\varepsilon^{-1}\big).
$$

Hence, the dominated convergence theorem implies that $\mathbb{E}B_\ell(V^{a,m}_{t \wedge T_M -})$ converges when $m$ goes to infinity towards $\mathbb{E}B_\ell(V^a_{t \wedge T_M -})$.

Let us consider the first term in the last integral of step 2, the second term integral in inequality (4.2). We write

$$\mathbb{E}\int_0^{t \wedge T_M} \tau_\ell(V^{a,m}_{s-})\,d[X^m, X^m]^c_s = \mathbb{E}\int_0^{t \wedge T_M} \tau_\ell(V^a_{s-})\,d[X, X]^c_s$$

$$+ \mathbb{E}\int_0^{t \wedge T_M} \left(\tau_\ell(V^{a,m}_{s-}) - \tau_\ell(V^a_{s-})\right)d[X, X]^c_s$$

$$+ \mathbb{E}\int_0^{t \wedge T_M} \tau_\ell(V^{a,m}_{s-})\,d\left([X^m, X^m]^c - [X, X]^c\right)_s.$$

The uniform boundedness and continuity of $\tau_\ell$, the square integrability of $X$ and the Dominated convergence theorem imply that the second term of the right-hand side converges to zero. The last term can be bounded above using the estimates for $\tau_\ell$,

$$\left| \mathbb{E}\int_0^{t \wedge T_M} \tau_\ell(V^{a,m}_{s-})\,d\left([X^m, X^m]^c - [X, X]^c\right)_s \right|$$

$$\lesssim \frac{Q}{\varepsilon}\mathbb{E}\int_0^{t \wedge T_M} \left| d\left([X^m, X^m]^c - [X, X]^c\right)_s \right|$$

$$\lesssim \frac{Q}{\varepsilon}\mathbb{E}\int_0^{t \wedge T_M} d\left([X, X]^c - [X^m, X^m]^c\right)_s$$

$$\lesssim \frac{Q}{\varepsilon}\left(\mathbb{E}[X, X]^c_{t \wedge T_M} - \mathbb{E}[X^m, X^m]^c_{t \wedge T_M}\right),$$

where we used that for a fixed $m$, $[X, X]^c - [X^m, X^m]^c$ is a nonnegative nondecreasing process. The last expression in the last line tends to zero when $m \to \infty$ by the monotone convergence theorem. We prove in a similar manner the convergence

$$\mathbb{E}\sum_{0 < s < t \wedge T_M} |\Delta X^m_s||\Delta Z^m_s| \xrightarrow[m \to \infty]{} \mathbb{E}\sum_{0 < s < t \wedge T_M} |\Delta X_s||\Delta Z_s|.$$

Finally, since $|V^{a,m}_{t \wedge T_M -}| \leq |V^a_{t \wedge T_M -}|$ for all $m$, $d_V B_\ell$ is continuous and bounded on compacts, $|\Delta V^m_{t \wedge T_M}|^2 \leq |\Delta V_{t \wedge T_M}|^2 \leq [V, V]_t$ and $\mathbb{E}[V, V]_t = \mathbb{E}|V_t|^2 < \infty$, the dominated convergence theorem ensures that

$$-\mathbb{E}\{d_V B_\ell(V^{a,m}_{t \wedge T_M -})\Delta V^m_{t \wedge T_M}\} \to -\mathbb{E}\{d_V B_\ell(V^a_{t \wedge T_M -})\Delta V_{t \wedge T_M}\}.$$

Collecting all terms,

$$\mathbb{E}B_\ell(V^a_{t \wedge T_M -})$$

$$\geq \frac{1}{2Q}\mathbb{E}\int_0^{t \wedge T_M} \tau_\ell(V^a_{s-})\,d[X, X]^c_s + \left(\tau_\ell(V^a_{s-})\right)^{-1}d[Z, Z]^c_s$$

$$+ \frac{1}{Q} \mathbb{E} \sum_{0 < s < t \wedge T_M} |\Delta X_s||\Delta Z_s|$$

$$- \mathbb{E}\{d_V B_\ell(V^a_{t \wedge T_M-}) \Delta V_{t \wedge T_M}\}.$$

*Step 4.* Now we add the contribution of the possible jumps occurring at $T_M$. We have seen in Step 1 the dissipation estimate along one jump

$$B_\ell(V^a_{t \wedge T_M}) - B_\ell(V^a_{t \wedge T_M-}) - d_V B_\ell(V^a_{t \wedge T_M-}) \Delta V_{t \wedge T_M}$$

$$\geq \frac{1}{Q} |\Delta X_{t \wedge T_M}||\Delta Z_{t \wedge T_M}|.$$

Taking expectation and adding the contribution of Step 3 yields

(4.3)
$$\mathbb{E} B_\ell(V^a_{t \wedge T_M}) \geq \frac{1}{2Q} \mathbb{E} \int_0^{t \wedge T_M} \tau_\ell(V^a_{s-}) \, d[X, X]^c_s + (\tau_\ell(V^a_{s-}))^{-1} \, d[Z, Z]^c_s$$

$$+ \frac{1}{Q} \mathbb{E} \sum_{0 < s \leq t \wedge T_M} |\Delta X_s||\Delta Z_s|.$$

*Step 5.* We will pass to the limit $M \to \infty$. Recall again that $0 \leq B_\ell(V) \lesssim (1+\ell)(X^2/u + Y^2/w)$. Using Doob's inequality for square integrable martingales, we have for all $M$,

$$\mathbb{E} B_\ell(V^a_{t \wedge T_M}) \lesssim \varepsilon^{-1}(1+\ell)(\mathbb{E}X^2 + \mathbb{E}Y^2) < \infty.$$

So, by the dominated convergence theorem, $\mathbb{E} B_\ell(V^a_{t \wedge T_M}) \to \mathbb{E} B_\ell(V^a_t)$ as $M \to \infty$. The monotone convergence theorem for the integral in the right-hand side of the inequality (4.3) therefore yields

$$(1 + \ell)(\mathbb{E}F_t + \mathbb{E}G_t + 2a^2\varepsilon^{-1})$$

$$\geq \mathbb{E} B_\ell(V^a_t)$$

$$\gtrsim \frac{1}{2Q} \mathbb{E} \int_0^t \tau_\ell(V^a_{s-}) \, d[X, X]^c_s + (\tau_\ell(V^a_{s-}))^{-1} \, d[Z, Z]^c_s$$

$$+ \frac{1}{Q} \mathbb{E} \sum_{0 < s \leq t} |\Delta X_s||\Delta Z_s|.$$

This concludes the proof of Proposition 2.   □

## 5. Truncation of the weights.
Due to several technicalities in the proof, we have used weights bounded from above and away from 0. In order to pass to the general case, we cut a possibly unbounded weight above and below and show that this operation does not increase the characteristics of the weight. This is convenient and has been used in several places; here, we extend [30] to the martingale setting. We need the following preliminary lemmas.

LEMMA 7 (Truncation from above).   *For $a > 0$, let $M = \{w \leq a\}$, $H = \{w > a\}$ and set $w_{\bar{a}} = w\chi_M + a\chi_H$. Then $Q_2^{\mathcal{F}}[w_{\bar{a}}] \leq Q_2^{\mathcal{F}}[w]$.*

PROOF.   Let $\tau$ be a stopping time and let us decompose

$$\mathbb{E}(w|\mathcal{F}_\tau) = \mathbb{E}(w\chi_M|\mathcal{F}_\tau) + \mathbb{E}(w\chi_H|\mathcal{F}_\tau)$$
$$= \mathbb{E}(\chi_M|\mathcal{F}_\tau)\mathbb{E}_M(w|\mathcal{F}_\tau) + \mathbb{E}(\chi_H|\mathcal{F}_\tau)\mathbb{E}_H(w|\mathcal{F}_\tau),$$

where, for example, $\mathbb{E}_M(w|\mathcal{F}_\tau)$ means expectation is taken with respect to the measure $\chi_M \, d\mathbb{P}$. Write as usual $\mathbb{E}(\chi_M|\mathcal{F}_\tau) = (\chi_M)_\tau$. We have

$$\mathbb{E}(w|\mathcal{F}_\tau)\mathbb{E}(w^{-1}|\mathcal{F}_\tau) - \mathbb{E}(w_{\bar{a}}|\mathcal{F}_\tau)\mathbb{E}(w_{\bar{a}}^{-1}|\mathcal{F}_\tau)$$
$$= \big((\chi_M)_\tau\mathbb{E}_L(w|\mathcal{F}_\tau) + (\chi_H)_\tau\mathbb{E}_H(w|\mathcal{F}_\tau)\big)\big((\chi_M)_\tau\mathbb{E}_L(w^{-1}|\mathcal{F}_\tau)$$
$$+ (\chi_H)_\tau\mathbb{E}_H(w^{-1}|\mathcal{F}_\tau)\big) - \big((\chi_M)_\tau\mathbb{E}_L(w|\mathcal{F}_\tau)$$
$$+ (\chi_H)_\tau a\big)\big((\chi_M)_\tau\mathbb{E}_L(w^{-1}|\mathcal{F}_\tau) + (\chi_H)_\tau a^{-1}\big)$$
$$= (\chi_M)_\tau(\chi_H)_\tau\big(\mathbb{E}_M(w|\mathcal{F}_\tau)\mathbb{E}_H(w^{-1}|\mathcal{F}_\tau) + \mathbb{E}_M(w^{-1}|\mathcal{F}_\tau)\mathbb{E}_H(w|\mathcal{F}_\tau)$$
$$- \mathbb{E}_M(w|\mathcal{F}_\tau)a^{-1} - \mathbb{E}_M(w^{-1}|\mathcal{F}_\tau)a\big)$$
$$+ (\chi_H)_\tau^2\big(\mathbb{E}_H(w|\mathcal{F}_\tau)\mathbb{E}_H(w^{-1}|\mathcal{F}_\tau) - 1\big).$$

The last term is positive thanks to Jensen inequality. Let us observe that also

$$\mathbb{E}_M(w|\mathcal{F}_\tau)\mathbb{E}_H(w^{-1}|\mathcal{F}_\tau) + \mathbb{E}_M(w^{-1}|\mathcal{F}_\tau)\mathbb{E}_H(w|\mathcal{F}_\tau)$$
$$- \mathbb{E}_M(w|\mathcal{F}_\tau)a^{-1} - \mathbb{E}_M(w^{-1}|\mathcal{F}_\tau)a$$
$$= \mathbb{E}_M(w|\mathcal{F}_\tau)\mathbb{E}_H(w^{-1} - a^{-1}|\mathcal{F}_\tau) + \mathbb{E}_M(w^{-1}|\mathcal{F}_\tau)\mathbb{E}_H(w - a|\mathcal{F}_\tau)$$
$$= \mathbb{E}_H\left(\frac{w-a}{wa}\big(wa\mathbb{E}_M(w^{-1}|\mathcal{F}_\tau) - \mathbb{E}_M(w|\mathcal{F}_\tau)\big)\Big|\mathcal{F}_\tau\right)$$
$$\geq 0.$$

Here, the last inequality uses $\mathbb{E}_M(w^{-1}|\mathcal{F}_\tau) \geq a^{-1}$ and $\mathbb{E}_M(w|\mathcal{F}_\tau) \leq a$ also $w - a \geq 0$ on $H$. This proves the lemma.   $\square$

LEMMA 8 (Two-sided truncation).   *For $a > 0$, let $M = \{a^{-1} \leq w \leq a\}$, $L = \{w < a^{-1}\}$, $H = \{w > a\}$ and set $w_a = a^{-1}\chi_L + w\chi_M + a\chi_H$. Then we have $Q_2^{\mathcal{F}}[w_a] \leq Q_2^{\mathcal{F}}[w]$.*

PROOF.   Let $w_{\bar{a}}$ be the weight obtained in the previous lemma. Apply now the previous lemma to $w_{\bar{a}}^{-1}$, truncating above by the same $a$.   $\square$

## 6. Proof of the main results.

PROOF OF PROPOSITION 1 (Bilinear estimate). Let $\lambda > 0$. If $Y$ differentially subordinate to $X$, then $\lambda Y$ is differentially subordinate to $\lambda X$. Let $w$ a weight in the $A_2$ class. Let $w^\varepsilon$ denote the $\varepsilon$-truncation of $w$. Use Proposition 2 with $V^{\varepsilon,\lambda} := (\lambda X, \lambda^{-1} Z, u^\varepsilon, w^\varepsilon)$ and $Q = Q_2^{\mathcal{F}}[w]$. Recalling that $Q_2^{\mathcal{F}}[w^\varepsilon] \leq Q_2^{\mathcal{F}}[w]$, noticing that $V^{\varepsilon,\lambda} \in \mathcal{D}_Q^{\varepsilon,\ell}$ and using the differential subordination of $\lambda Y$ w.r.t. $\lambda X$, we have for all $t > 0$,

$$Q_2^{\mathcal{F}}[w](1+\ell)\big(\mathbb{E}\lambda^2 F_t + \mathbb{E}\lambda^{-2} G_t + 2a^2\varepsilon^{-1}\big)$$

$$\gtrsim \frac{1}{2}\mathbb{E}\int_0^t \tau_\ell(V_{s-}^a)\,\mathrm{d}[\lambda X, \lambda X]_s^c + \big(\tau_\ell(V_{s-}^a)\big)^{-1}\,\mathrm{d}[\lambda^{-1}Z, \lambda^{-1}Z]_s^c$$

$$+ \mathbb{E}\sum_{0<s\leq t}|\lambda\Delta X_s||\lambda^{-1}\Delta Z_s|$$

$$\gtrsim \frac{1}{2}\mathbb{E}\int_0^t \tau_\ell(V_{s-}^a)\,\mathrm{d}[\lambda Y, \lambda Y]_s^c + \big(\tau_\ell(V_{s-})\big)^{-1}\,\mathrm{d}[\lambda^{-1}Z, \lambda^{-1}Z]_s^c$$

$$+ \mathbb{E}\sum_{0<s\leq t}|\Delta Y_s||\Delta Z_s|.$$

Since for any $0 < \kappa < \infty$ and any $x \in \mathbb{H}$, $y \in \mathbb{H}$, we have $\kappa x^2 + \kappa^{-1}y^2 \geq 2|\langle x, y\rangle|$; it follows easily that

$$\frac{1}{2}\mathbb{E}\int_0^t \tau_\ell(V_{s-})\,\mathrm{d}[\lambda Y, \lambda Y]_s^c + \big(\tau_\ell(V_{s-})\big)^{-1}\,\mathrm{d}[\lambda^{-1}Z, \lambda^{-1}Z]_s^c$$

$$+ \mathbb{E}\sum_{0<s\leq t}|\Delta Y_s||\Delta Z_s|$$

$$\geq \mathbb{E}\int_0^t \big|\mathrm{d}[\lambda Y, \lambda^{-1}Z]_s^c\big| + \mathbb{E}\sum_{0<s\leq t}|\Delta Y_s||\Delta Z_s|$$

$$\geq \mathbb{E}\int_0^t \big|\mathrm{d}[Y, Z]_s^c\big| + \mathbb{E}\sum_{0<s\leq t}|\Delta Y_s||\Delta Z_s|$$

$$\geq \mathbb{E}\int_0^t \big|\mathrm{d}[Y, Z]_s\big|,$$

where all integrals and sums converge. Hence for all $\lambda > 0$,

$$Q_2^{\mathcal{F}}[w](1+\ell)(\lambda^2\mathbb{E}F_t + \lambda^{-2}\mathbb{E}G_t + 2a^2\varepsilon^{-1}) \gtrsim \mathbb{E}\int_0^t \big|\mathrm{d}[Y, Z]_s\big|.$$

We let now successively $\ell \to 0$ then $a \to 0$. Also, choosing the specific value $\lambda^2 = (\mathbb{E}G_t)^{1/2}(\mathbb{E}F_t)^{-1/2}$, we can assume $\lambda > 0$ (otherwise the claim is trivial). We have

$$\mathbb{E}\int_0^t \big|\mathrm{d}[Y, Z]_s\big| \lesssim Q_2^{\mathcal{F}}[w](\mathbb{E}F_t)^{1/2}(\mathbb{E}G_t)^{1/2} \lesssim Q_2^{\mathcal{F}}[w]\|X\|_{2,w^\varepsilon}\|Z\|_{2,u^\varepsilon}.$$

The inequality above remains valid in the limit $t \to \infty$. Since the left-hand side does not depend on the truncation of the weight, it remains to observe that

$$\lim_{\varepsilon \to 0} \|X\|_{2,w^\varepsilon} = \|X\|_{2,w} \quad \text{and} \quad \lim_{\varepsilon \to 0} \|Z\|_{2,u^\varepsilon} = \|Z\|_{2,u}.$$

Indeed, since $X \in L^2(\Omega; d\mathbb{P}) \cap L^2(\Omega; d\mathbb{P}^w)$, we have for all $0 < \varepsilon < 1$, a.s. $X_\infty^2 w_\infty^\varepsilon \leq X_\infty^2 + X_\infty^2 w_\infty$ and the limits above are a consequence of the dominated convergence theorem. The same reasoning applied to $Z$ completes the proof of the bilinear embedding. $\square$

PROOF OF THEOREM 1 (Differential subordination under change of law). The proof of the main result is now straightforward since the proposition above allows us to estimate, for any test function $Z_\infty \in L^2(\Omega, d\mathbb{P}) \cap L^2(\Omega, d\mathbb{P}^u)$,

$$|(Y_\infty, Z_\infty)| = \left| \int_0^\infty d[Y, Z]_s \right| \leq \int_0^\infty |d[Y, Z]_s| \lesssim Q_2^{\mathcal{F}}[w] \|X\|_{2,w} \|Z\|_{2,u},$$

that is exactly

$$\|Y\|_{2,w} \lesssim Q_2^{\mathcal{F}}[w] \|X\|_{2,w}.$$

This concludes the proof of Theorem 1. $\square$

## 7. Sharpness and applications.

7.1. *Sharpness.* Sharpness means that the linear power in the martingale $A_2$ characteristic cannot be improved.

7.1.1. *Discrete time.* That the result is sharp in the dyadic, discrete-in-time filtration case is well known and follows from the sharpness of the linear estimate for the dyadic square function in this setting (see [17] for an explicit calculation). Notice that the norm of the square function is no larger than that of a predictable dyadic multiplier—given the dyadic square function is obtained by taking expectation of a $\sigma = \pm 1$ predictable multiplier $T_\sigma$. Indeed $Sf^2(t) = \mathbb{E}|T_\sigma f(t)|^2$; see, for example, [27].

7.1.2. *Continuous time.* To see an example with continuous-in-time filtration, we use the Hilbert transform as an intermediary. We briefly summarize the flow of the argument; see [13] for details.

Let $f(x)$ be a compactly supported integrable function, defined on $\mathbb{R}$ and $\tilde{f}(z) = \tilde{f}(x, y)$ its harmonic extension to the upper half-space. Let $W_t = (x_t, y_t)$ be background noise (see [16]).

Then the martingales

$$M_t^{\tilde{f}} = \tilde{f}(W_t) \quad \text{and} \quad M_t^{\widetilde{Hf}} = \widetilde{Hf}(W_t)$$

($H$ the Hilbert transform) are a pair of differentially subordinate martingales. To see this, use the formula by Gundy–Varopoulos [16] applied to the Hilbert transform,

$$M_t^{\widetilde{Hf}} = \int_{-\infty}^t A\nabla \tilde{f}(W_s)\,\mathrm{d}s,$$

where $A$ is the counter-clockwise rotation matrix by $\pi/2$, stemming from the Cauchy–Riemann relations. Further, it is easy to see that the deterministic Poisson $A_2$ characteristic and the martingale $A_2$ characteristic driven by background noise are comparable.

These facts enable us to obtain certain martingale estimates from Hilbert transform bounds and vice versa. But for the critical index $p = 2$ there is no explicit example that exhibits sharpness of the linear growth of the weighted bound for the Hilbert transform with respect to the Poisson characteristic.

However, for $1 < p < 2$, there are explicit examples that show the optimal behaviour for the Hilbert transform with respect to the Poisson $A_p$ characteristic. This allows one to pass to the exponent $p = 2$ through an extrapolation argument using again the martingale setting.

### 7.2. *Applications*.

#### 7.2.1. *Discrete-in-time predictable multipliers*.
The Bellman function in this paper and in particular its one-leg convexity can give a direct proof of the results in [21] and [31], a weighted estimate for predictable multipliers in the case of discrete in time filtrations.

#### 7.2.2. *Dimension-free weighted bounds on discrete operators*.
Through the recent stochastic integral formula for second-order Riesz transforms [1] on compact multiply-connected Lie groups $\mathbb{G}$, our result gives dimension-free weighted $L^2$ estimates in this setting, too, using the semi-discrete heat characteristic of the weight. The second-order Riesz transforms take the form

$$R_\alpha^2 = \sum_i \alpha_i R_i^2 + \sum_{j,k} \alpha_{ij} R_{ij}^2,$$

where the first diagonal sum are second-order Riesz transforms in discrete directions of the space and the second sum are continuous second-order Riesz transforms on the connected part; see [1] for more precise definitions. The process considered is deterministic in one variable and is Brownian in continuous directions together with a compound Poisson jump process in the other discontinuous directions. It was proved in [1] that $R_\alpha^2 f(z)$ can be written as the conditional expectation $\mathbb{E}(M_0^{\alpha,f}|Z_0 = z)$ where $M_t^{\alpha,f}$ is a martingale transform of $M_t^f$ associated to $f$ and $Z_t$ a suitable random walk. One obtains the estimate

(7.1)                     $$\|R_\alpha f\|_{L^2(w)} \lesssim Q_2(w)\|f\|_{L^2(w)}$$

with implied constant independent of dimension and where $Q_2(w)$ is the semi-discrete heat characteristic. An important special case are the second-order discrete Riesz transforms on products of integers. Notice that both the continuous-in-time filtrations and the consideration of jump processes are important to get this estimate.

It is also possible to get a deterministic proof of this application (7.1), using the Bellman function we construct in this paper in combination with part of the proof strategy in [12]. Notice though, that the trick used in [12] to overcome the difficulty of the jumps, does not work in the weighted setting, due to nonconvexity of the domain of the Bellman function. For a deterministic Bellman proof to give the weighted estimate (7.1), it is instrumental to have the one-leg convexity property proved here.

7.2.3. *Probabilistic proof for estimate of the weighted Beurling operator.* Our result gives a probabilistic proof of the weighted estimate for the well-known Beurling–Ahlfors transform that solved a famous borderline regularity problem in [28] previously proved by Bellman functions and other means. To see this, one invokes the stochastic integral identity formula [2] for the Beurling–Ahlfors operator using heat flow martingales. The comparability of heat flow $A_2$ characteristic and martingale characteristic obtained when using the filtration in [2] is not hard to see. In turn, in [28] it was seen that the heat flow characteristic compares linearly to the classical characteristic. The standard extrapolation result for sub-linear operators in [14] gives the sharp weighted result in $L^p$.

7.2.4. *Dimension-free weighted bound, Riemannian setting.* Dahmani [9] used the continuous properties of the Bellman function constructed in this paper to prove a dimensionless weighted bound for the Bakry Riesz vector. Her result gives an optimal estimate in terms of the Poisson characteristic. She considers a large class of manifolds with nonnegative Bakry–Emery curvature, such as, for example, the Gauss space. The explicit expression of the Bellman function is essential to her argument.

## REFERENCES

[1] ARCOZZI, N., DOMELEVO, K. and PETERMICHL, S. (2016). Second order Riesz transforms on multiply-connected Lie groups and processes with jumps. *Potential Anal.* **45** 777–794. MR3558360

[2] BAÑUELOS, R. and JANAKIRAMAN, P. (2008). $L^p$-bounds for the Beurling–Ahlfors transform. *Trans. Amer. Math. Soc.* **360** 3603–3612. MR2386238

[3] BAÑUELOS, R. and OSĘKOWSKI, A. (2014). On the Bellman function of Nazarov, Treil and Volberg. *Math. Z.* **278** 385–399. MR3267584

[4] BAÑUELOS, R. and OSĘKOWSKI, A. (2018). Weighted $L^2$ inequalities for square functions. *Trans. Amer. Math. Soc.* **370** 2391–2422. MR3748572

[5] BAÑUELOS, R. and WANG, G. (1995). Sharp inequalities for martingales with applications to the Beurling–Ahlfors and Riesz transforms. *Duke Math. J.* **80** 575–600. MR1370109

[6] BERNICOT, F., FREY, D. and PETERMICHL, S. (2016). Sharp weighted norm estimates beyond Calderón–Zygmund theory. *Anal. PDE* **9** 1079–1113. MR3531367

[7] BONAMI, A. and LÉPINGLE, D. (1979). Fonction maximale et variation quadratique des martingales en présence d'un poids. In *Séminaire de Probabilités*, *XIII* (*Univ. Strasbourg*, *Strasbourg*, 1977/78). *Lecture Notes in Math.* **721** 294–306. Springer, Berlin. MR0544802

[8] BURKHOLDER, D. L. (1984). Boundary value problems and sharp inequalities for martingale transforms. *Ann. Probab.* **12** 647–702. MR0744226

[9] DAHMANI, K. (2016). Sharp dimension-free weighted bounds for the Bakry–Riesz vector. In preparation. Available at arXiv:1611.07696.

[10] DELLACHERIE, C. and MEYER, P.-A. (1982). *Probabilities and Potential*. *B*: *Theory of Martingales*. *North-Holland Mathematics Studies* **72**. North-Holland, Amsterdam. Translated from the French by J. P. Wilson. MR0745449

[11] DOMELEVO, K. and PETERMICHL, S. (2014). Sharp $L^p$ estimates for discrete second order Riesz transforms. *Adv. Math.* **262** 932–952. MR3228446

[12] DOMELEVO, K. and PETERMICHL, S. (2014). Bilinear embeddings on graphs. Preprint.

[13] DOMELEVO, K., PETERMICHL, S. and WITTWER, J. (2017). A linear dimensionless bound for the weighted Riesz vector. *Bull. Sci. Math.* **141** 385–407. MR3667592

[14] DRAGIČEVIĆ, O., GRAFAKOS, L., PEREYRA, M. C. and PETERMICHL, S. (2005). Extrapolation and sharp norm estimates for classical operators on weighted Lebesgue spaces. *Publ. Mat.* **49** 73–91. MR2140200

[15] DRAGIČEVIĆ, O., TREIL, S. and VOLBERG, A. (2008). A theorem about three quadratic forms. *Int. Math. Res. Not. IMRN* Art. ID rnn 072. MR2439569

[16] GUNDY, R. F. and VAROPOULOS, N. TH. (1979). Les transformations de Riesz et les intégrales stochastiques. *C. R. Acad. Sci. Paris Sér. A-B* **289** A13–A16. MR0545671

[17] HUKOVIC, S., TREIL, S. and VOLBERG, A. (2000). The Bellman functions and sharp weighted inequalities for square functions. In *Complex Analysis*, *Operators*, *and Related Topics*. *Oper. Theory Adv. Appl.* **113** 97–113. Birkhäuser, Basel. MR1771755

[18] HUNT, R., MUCKENHOUPT, B. and WHEEDEN, R. (1973). Weighted norm inequalities for the conjugate function and Hilbert transform. *Trans. Amer. Math. Soc.* **176** 227–251. MR0312139

[19] HYTÖNEN, T. P. (2012). The sharp weighted bound for general Calderón–Zygmund operators. *Ann. of Math.* (2) **175** 1473–1506. MR2912709

[20] IZUMISAWA, M. and KAZAMAKI, N. (1977). Weighted norm inequalities for martingales. *Tohoku Math. J.* (2) **29** 115–124. MR0436313

[21] LACEY, M. T. (2017). An elementary proof of the $A_2$ bound. *Israel J. Math.* **217** 181–195. MR3625108

[22] LACEY, M. T., PETERMICHL, S. and REGUERA, M. C. (2010). Sharp $A_2$ inequality for Haar shift operators. *Math. Ann.* **348** 127–141. MR2657437

[23] LERNER, A. K. (2016). On pointwise estimates involving sparse operators. *New York J. Math.* **22** 341–349. MR3484688

[24] NAZAROV, F., TREIL, S. and VOLBERG, A. (1999). The Bellman functions and two-weight inequalities for Haar multipliers. *J. Amer. Math. Soc.* **12** 909–928. MR1685781

[25] OSĘKOWSKI, A. (2018). Sharp $L^p$-bounds for the martingale maximal function. *Tohoku Math. J.* (2) **70** 121–138. MR3772808

[26] PETERMICHL, S. (2007). The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical $A_p$ characteristic. *Amer. J. Math.* **129** 1355–1375. MR2354322

[27] PETERMICHL, S. and POTT, S. (2002). An estimate for weighted Hilbert transform via square functions. *Trans. Amer. Math. Soc.* **354** 1699–1703. MR1873024

[28] PETERMICHL, S. and VOLBERG, A. (2002). Heating of the Ahlfors–Beurling operator: Weakly quasiregular maps on the plane are quasiregular. *Duke Math. J.* **112** 281–305. MR1894362

[29] PROTTER, P. E. (2005). *Stochastic Integration and Differential Equations. Stochastic Modelling and Applied Probability* **21**. Second edition. Version 2.1. Corrected third printing. Springer, Berlin. MR2273672

[30] REZNIKOV, A., VASYUNIN, V. and VOLBERG, A. (2010). An observation: Cut-off of the weight $w$ does not increase the $a_{p_1, p_2}$-"norm" of $w$. Available at arXiv:1008.3635.

[31] THIELE, C., TREIL, S. and VOLBERG, A. (2015). Weighted martingale multipliers in the non-homogeneous setting and outer measure spaces. *Adv. Math.* **285** 1155–1188. MR3406523

[32] TREIL, S. (2013). Sharp $A_2$ estimates of Haar shifts via Bellman function. In *Recent Trends in Analysis. Theta Ser. Adv. Math.* **16** 187–208. Theta, Bucharest. MR3411052

[33] VASYUNIN, V. and VOLBERG, A. (2010). Burkholder's function via Monge–Ampère equation. *Illinois J. Math.* **54** 1393–1428. MR2981853

[34] WANG, G. (1995). Differential subordination and strong differential subordination for continuous-time martingales and related sharp inequalities. *Ann. Probab.* **23** 522–551. MR1334160

[35] WITTWER, J. (2000). A sharp estimate on the norm of the martingale transform. *Math. Res. Lett.* **7** 1–12. MR1748283

INSITUT DE MATHÉMATIQUES DE TOULOUSE
UNIVERSITÉ PAUL SABATIER
118, ROUTE DE NARBONNE
F-31062 TOULOUSE CEDEX 9
FRANCE
E-MAIL: komla.domelvo@math.univ-toulouse.fr
        stefanie.petermichl@math.univ-toulouse.fr