

## BELLMAN FUNCTIONS AND $L^p$ ESTIMATES FOR PARAPRODUCTS\*

BY

VJEKOSLAV KOVAČ (ZAGREB) AND KRISTINA ANA ŠKREB (ZAGREB)

*Abstract.* We give an explicit formula for one possible Bellman function associated with the  $L^p$  boundedness of dyadic paraproducts regarded as bilinear operators or trilinear forms. Then we apply the same Bellman function in various other settings, to give self-contained alternative proofs of the estimates for several classical operators. These include the martingale paraproducts of Bañuelos and Bennett and the paraproducts with respect to the heat flows.

**2010 AMS Mathematics Subject Classification:** Primary: 60G46; Secondary: 42B15.

**Key words and phrases:** Bellman function, martingale, paraproduct.

### 1. INTRODUCTION

According to Janson and Peetre [14] the name “paraproduct” denotes an idea rather than a unique object. Various types of paraproducts appear in the literature on analysis or probability and in each case certain boundedness properties (i.e. continuity) are crucial for their applications. An interested reader can find the historical overview and further references in the short expository paper [4]. In this paper we will focus mostly on martingale paraproducts and revisit the  $L^p$  estimates, which they are well known to satisfy.

We start with the dyadic paraproduct as a motivation for the forthcoming Bellman function that we construct. For two functions  $f$  and  $g$  from an appropriate space of real-valued test functions on  $\mathbb{R}$  we can define the *dyadic paraproduct* as a bilinear operator in the following way:

$$(1.1) \quad \Pi_\epsilon(f, g) := \sum_{I \in \mathcal{D}} \epsilon_I |I|^{-2} \langle f, \mathbb{1}_I \rangle \langle g, \mathbb{h}_I \rangle \mathbb{h}_I.$$

Here  $\mathcal{D}$  denotes the family of dyadic intervals in  $\mathbb{R}$ ,  $\mathbb{1}_I$  is the indicator function of an interval  $I$ ,  $\mathbb{h}_I := \mathbb{1}_{I_{\text{left}}} - \mathbb{1}_{I_{\text{right}}}$  is the  $L^\infty$ -normalized Haar function, while

---

\* This work has been supported by the Croatian Science Foundation under the project 3526.

$I_{\text{left}}$  and  $I_{\text{right}}$  are respectively the left half and the right half of  $I$ . Moreover,  $\langle \cdot, \cdot \rangle$  denotes the standard inner product with respect to the Lebesgue measure and  $\epsilon = (\epsilon_I)_{I \in \mathcal{D}}$  is a collection of real numbers such that  $|\epsilon_I| \leq 1$  for each  $I \in \mathcal{D}$ . (If we choose  $\epsilon_I \in \{-1, 1\}$ , then they simply represent  $-$  and  $+$  signs.) A convenient choice for the test functions are the so-called *dyadic step functions*, i.e. finite linear combinations of the indicator functions of dyadic intervals.

Typically, such an object is viewed as a linear operator in  $g$  with  $f$  fixed, when it becomes a particular instance of *Burkholder's martingale transform* [5]. Alternatively, one can fix  $g$  and consider it as a linear operator in  $f$ , in which case it is known as the linear paraproduct. In this text we prefer to look at  $\Pi_\epsilon$  symmetrically and discuss its properties as a bilinear operator. This is partly motivated by the multilinear harmonic analysis, where more singular operators of this type are studied; see the book [24].

Equivalently, we can define the dyadic paraproduct as a trilinear form. We take the third test function  $h$ , and dualize (1.1) to get

$$(1.2) \quad \begin{aligned} \Lambda_\epsilon(f, g, h) &:= \int_{\mathbb{R}} \Pi_\epsilon(f, g)h = \sum_{I \in \mathcal{D}} \epsilon_I |I|^{-2} \langle f, \mathbb{1}_I \rangle \langle g, \mathbb{1}_I \rangle \langle h, \mathbb{1}_I \rangle \\ &= \sum_{I \in \mathcal{D}} \epsilon_I |I| [f]_I \frac{[g]_{I_{\text{left}}} - [g]_{I_{\text{right}}}}{2} \frac{[h]_{I_{\text{left}}} - [h]_{I_{\text{right}}}}{2}. \end{aligned}$$

Here  $[f]_I$  denotes the average of a function  $f$  on a dyadic interval  $I$ .

It is well known that (1.2) satisfies certain  $L^p$  estimates, i.e. there exists a finite constant  $C_{p,q,r} \geq 0$  depending only on three exponents  $p, q, r$  such that

$$(1.3) \quad |\Lambda_\epsilon(f, g, h)| \leq C_{p,q,r} \|f\|_{L^p(\mathbb{R})} \|g\|_{L^q(\mathbb{R})} \|h\|_{L^r(\mathbb{R})}$$

holds whenever  $1 < p, q, r \leq \infty$  and  $1/p + 1/q + 1/r = 1$ . By  $\|\cdot\|_{L^p(\mathbb{R})}$  we have denoted the  $L^p$  norm on  $\mathbb{R}$  with respect to the Lebesgue measure.

The easiest proof of (1.3) when  $q, r < \infty$  uses boundedness of the dyadic maximal function and the dyadic square function. We simply apply the Cauchy–Schwarz and Hölder inequalities to get

$$|\Lambda_\epsilon(f, g, h)| \leq \int_{\mathbb{R}} (Mf)(Sg)(Sh) \leq \|Mf\|_{L^p(\mathbb{R})} \|Sg\|_{L^q(\mathbb{R})} \|Sh\|_{L^r(\mathbb{R})},$$

where

$$Mf := \sup_{I \in \mathcal{D}} |I|^{-1} |\langle f, \mathbb{1}_I \rangle| \mathbb{1}_I \quad \text{and} \quad Sf := \left( \sum_{I \in \mathcal{D}} |I|^{-2} |\langle f, \mathbb{1}_I \rangle|^2 \mathbb{1}_I \right)^{1/2}$$

are the *dyadic maximal function* and the *dyadic square function*. Now the well-known  $L^p$  estimates for  $Mf$  and  $Sf$  give us the desired estimate (1.3).

On the side  $p = \infty$  of the triangle in Figure 1, without loss of generality we can assume that  $f \equiv 1$ . The sharp constant in (1.3) was found by Burkholder in [6] and it equals  $C_{\infty,q,r} = \max\{q - 1, r - 1\}$ .

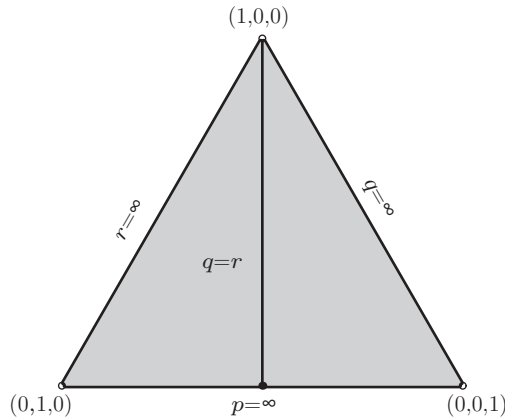


FIGURE 1. The Banach triangle with barycentric coordinates  $(\frac{1}{p}, \frac{1}{q}, \frac{1}{r})$ .

On the other hand, on the sides  $q = \infty$  and  $r = \infty$ , instead of the  $L^p$  estimates it is more natural to consider the BMO estimates, which will not be discussed in this paper. On the altitude  $q = r$  of the triangle in Figure 1, the  $L^p$  estimates for the trilinear form (1.2) reduce to the  $L^p$  estimates for the dyadic square function, since

$$\int_{\mathbb{R}} f(Sg)^2 = \Lambda_\epsilon(f, g, g) \quad \text{if } \epsilon_I = 1 \text{ for each } I \in \mathcal{D}.$$

This implies  $\|S\|_{L^q(\mathbb{R}) \rightarrow L^q(\mathbb{R})} \leq \sqrt{C_{p,q,q}}$ . Actually, if the constant  $C_{p,q,q}$  is sharp, the last inequality turns into an equality. That sharp constant was found by Davis in [11] and it equals  $C_{p,q,q} = (z_q^*)^{-2}$ , where  $z_q^*$  is the smallest positive zero of the confluent hypergeometric function (see [1]).

The special cases listed above are well studied and even the appropriate Bellman functions are found. For  $p = \infty$  one can find them in the papers by Burkholder [6], Nazarov and Treil [17], Vasyunin and Volberg [25], Bañuelos and Osękowski [3], while for  $q = r$  the reader can consult the book by Osękowski [20]. Therefore, because of the symmetry, throughout this paper we restrict our attention to the triples of exponents  $(p, q, r)$  satisfying

$$(1.4) \quad 1 < p, q, r < \infty, \quad q > r, \quad \frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1,$$

which correspond to the right half of the Banach triangle depicted in Figure 1.

Our goal is to give a direct proof of (1.3) using the Bellman function method. Such proofs typically give a better quantitative control and the same Bellman function can often be applied in various other settings.

First, we may assume that  $f, g, h$  are non-negative, as otherwise we split them into positive and negative parts. Furthermore, we observe that it turns out to be more practical to apply Young’s inequality on the right-hand side of (1.3), but the newly obtained inequality is actually equivalent to the old one, because of

the homogeneity of the left-hand side. Therefore, it is enough to prove a non-homogeneous estimate

$$\sum_{I \in \mathcal{D}} |I| [f]_I \left| \frac{[g]_{I_{\text{left}}} - [g]_{I_{\text{right}}}}{2} \right| \left| \frac{[h]_{I_{\text{left}}} - [h]_{I_{\text{right}}}}{2} \right| \leq C_{p,q,r} \left( \frac{1}{p} \|f\|_{L^p(\mathbb{R})}^p + \frac{1}{q} \|g\|_{L^q(\mathbb{R})}^q + \frac{1}{r} \|h\|_{L^r(\mathbb{R})}^r \right).$$

If we want to recover (1.3), we just have to homogenize the above inequality and use the assumed bound on  $\epsilon_I$ .

For an arbitrary dyadic interval  $I$  we define a scale-invariant expression

$$\Phi_I(f, g, h) := \frac{1}{|I|} \sum_{\substack{J \in \mathcal{D} \\ J \subseteq I}} |J| [f]_J \frac{|[g]_{J_{\text{left}}} - [g]_{J_{\text{right}}}|}{2} \frac{|[h]_{J_{\text{left}}} - [h]_{J_{\text{right}}}|}{2},$$

so that we can normalize the desired estimate and rewrite it as

$$(1.5) \quad \Phi_I(f, g, h) \leq C_{p,q,r} \left( \frac{1}{p} [f^p]_I + \frac{1}{q} [g^q]_I + \frac{1}{r} [h^r]_I \right).$$

This is easily seen multiplying (1.5) by  $|I|$  and letting  $I$  exhaust the positive and the negative half-axis. Splitting  $\sum_{J \subseteq I}$  into  $\sum_{J \subseteq I_{\text{left}}}$ ,  $\sum_{J \subseteq I_{\text{right}}}$ , and  $J = I$  gives us the following scaling identity:

$$(1.6) \quad \Phi_I(f, g, h) = \frac{1}{2} \Phi_{I_{\text{left}}}(f, g, h) + \frac{1}{2} \Phi_{I_{\text{right}}}(f, g, h) + [f]_I \frac{|[g]_{I_{\text{left}}} - [g]_{I_{\text{right}}}|}{2} \frac{|[h]_{I_{\text{left}}} - [h]_{I_{\text{right}}}|}{2}.$$

We can define the *abstract Bellman function*

$$\mathbb{B}(u, v, w, U, V, W) := \sup_{f,g,h} \Phi_I(f, g, h),$$

where the supremum is taken over all non-negative functions  $f, g, h$  such that  $[f]_I = u$ ,  $[g]_I = v$ ,  $[h]_I = w$ ,  $[f^p]_I = U$ ,  $[g^q]_I = V$ ,  $[h^r]_I = W$ . Note that the above supremum does not depend on the choice of the “base” interval  $I$ .

Now we list some properties of that function.

(B1) *Domain:* The function  $\mathbb{B}$  is defined on the set

$$\mathbb{D} := \{(u, v, w, U, V, W) \in [0, \infty)^6 : u^p \leq U, v^q \leq V, w^r \leq W\}.$$

The upper bounds simply follow from Jensen’s inequality.

(B2) Range:

$$0 \leq \mathbb{B}(u, v, w, U, V, W) \leq C_{p,q,r} \left( \frac{1}{p}U + \frac{1}{q}V + \frac{1}{r}W \right),$$

where on the right-hand side we assume that the estimate (1.5) holds.

(B3) The main inequality:

$$\mathbb{B}(\mathbf{x}) \geq \frac{1}{2}\mathbb{B}(\mathbf{x}_1) + \frac{1}{2}\mathbb{B}(\mathbf{x}_2) + u \frac{|v_1 - v_2|}{2} \frac{|w_1 - w_2|}{2},$$

whenever the six-tuples  $\mathbf{x} = (u, v, w, U, V, W)$  and  $\mathbf{x}_i = (u_i, v_i, w_i, U_i, V_i, W_i)$ ,  $i = 1, 2$ , belong to the domain and satisfy  $\mathbf{x} = \frac{1}{2}\mathbf{x}_1 + \frac{1}{2}\mathbf{x}_2$ . This can be easily seen by taking the supremum in the scaling identity (1.6) over all non-negative functions  $f, g, h$  such that  $[f]_{I_{\text{left}}} = u_1, [f^p]_{I_{\text{left}}} = U_1$ , etc.

Conversely, suppose that we have already found a function  $\mathcal{B}$  with properties (B1)–(B3). We will show how its existence implies the estimate (1.3). Applying (B3)  $n$  times with a fixed choice of the functions  $f, g, h \geq 0$  and a fixed base interval  $I$  gives us

$$\begin{aligned} & |I| \mathcal{B}([f]_I, [g]_I, [h]_I, [f^p]_I, [g^q]_I, [h^r]_I) \\ & \geq \sum_{\substack{J \subseteq I \\ |J|=2^{-n}|I|}} |J| \mathcal{B}([f]_J, [g]_J, [h]_J, [f^p]_J, [g^q]_J, [h^r]_J) \\ & \quad + \sum_{\substack{J \subseteq I \\ |J|>2^{-n}|I|}} |J| [f]_J \frac{|[g]_{J_{\text{left}}} - [g]_{J_{\text{right}}}|}{2} \frac{|[h]_{J_{\text{left}}} - [h]_{J_{\text{right}}}|}{2}. \end{aligned}$$

Since, by (B2), the first sum is non-negative and

$$\mathcal{B}([f]_I, [g]_I, [h]_I, [f^p]_I, [g^q]_I, [h^r]_I) \leq C_{p,q,r} \left( \frac{1}{p}[f^p]_I + \frac{1}{q}[g^q]_I + \frac{1}{r}[h^r]_I \right),$$

letting  $n \rightarrow \infty$  leads us to the estimate (1.5) and then in turn also to (1.3).

It will be convenient to find a function  $\mathcal{B}$  that also satisfies the following condition:

$$(B4) \quad \mathcal{B}(\mathbf{x}) + (d\mathcal{B})(\mathbf{x})(\mathbf{x}_1 - \mathbf{x}) \geq \mathcal{B}(\mathbf{x}_1) + \frac{2}{3}u|v_1 - v||w_1 - w|,$$

whenever the six-tuples  $\mathbf{x} = (u, v, w, U, V, W)$  and  $\mathbf{x}_1 = (u_1, v_1, w_1, U_1, V_1, W_1)$  belong to the domain (B1). Here  $d\mathcal{B}$  denotes the differential of  $\mathcal{B}$ , which is a linear form, and we consider it at the point  $\mathbf{x}$  and apply it to the vector  $\mathbf{x}_1 - \mathbf{x}$ . Condition (B4) is required by an application considered in Subsection 3.1.

Now we want to find an explicit formula for one possible function  $\mathcal{B}$ . We define the function  $\mathcal{B}: \mathbb{D} \rightarrow \mathbb{R}$  as

$$(1.7) \quad \mathcal{B}(u, v, w, U, V, W) := C_{p,q,r} \left( \frac{1}{p}U + \frac{1}{q}V + \frac{1}{r}W \right) - \mathcal{A}(u, v, w),$$

where  $\mathcal{A}: [0, \infty)^3 \rightarrow \mathbb{R}$  is given by

$$\mathcal{A}(u, v, w) := \begin{cases} Au^p + Bv^q + Cw^r, & u^p \leq w^r \leq v^q, \\ \frac{A(p-1)-C}{p-1}u^p + Bv^q + \frac{Cp}{p-1}uw^{r-r/p}, & w^r \leq u^p \leq v^q, \\ \frac{A(p-1)-(B+C)}{p-1}u^p + \frac{Bp}{p-1}uv^{q-q/p} + \frac{Cp}{p-1}uw^{r-r/p}, & w^r \leq v^q \leq u^p, \\ \frac{A(p-1)-(B+C)}{p-1}u^p + \frac{Bq}{2}uv^2w^{1-r/q} + \frac{2Cpr-Bp(q-r)}{2r(p-1)}uw^{r-r/p}, & v^q \leq w^r \leq u^p, \\ \frac{2Ar(p-1)-B(q+r)}{2r(p-1)}u^p + \frac{Bq^2}{2p(q-2)}u^{p-2p/q}v^2 + \frac{Bq(q-r)}{2r(q-2)}v^2w^{r-2r/q} \\ \quad + \frac{2Cr-B(q-r)}{2r}w^r, & v^q \leq u^p \leq w^r, \\ Au^p + \frac{Bq}{p(q-2)}v^q + \frac{Bq(q-r)}{2r(q-2)}v^2w^{r-2r/q} + \frac{2Cr-B(q-r)}{2r}w^r, & u^p \leq v^q \leq w^r. \end{cases}$$

The coefficients  $A, B, C > 0$  will be appropriately chosen depending only on the exponents  $p, q, r$  and then one will be able to take  $C_{p,q,r} = \max\{Ap, Bq, Cr\}$ . We see that the function  $\mathcal{A}$  has a similar form to the one constructed by Nazarov and Treil [17], which can in our notation be written as

$$\mathcal{NT}(v, w) = A(v^q + w^r) + B \begin{cases} \frac{2}{q}v^q + \left(\frac{2}{r} - 1\right)w^r, & v^q \geq w^r, \\ v^2w^{2-r}, & v^q \leq w^r. \end{cases}$$

It corresponds to the endpoint case  $p = \infty, 1 < r < 2 < q < \infty$ . Instead of one critical curve  $v^q = w^r$  for  $\mathcal{NT}$ , we have three critical surfaces:

$$(1.8) \quad u^p = v^q, \quad u^p = w^r, \quad v^q = w^r.$$

Finally, we are ready to state our main result.

**THEOREM 1.1.** *For the exponents  $p, q, r$  satisfying (1.4) it is possible to choose the coefficients  $A, B, C$  such that the function  $\mathcal{B}$  defined by (1.7) is of class  $C^1$  on the whole domain  $\mathbb{D}$  and satisfies the conditions (B2) (with  $C_{p,q,r} = \max\{Ap, Bq, Cr\}$ ), (B3), and (B4). One possible choice of the coefficients is*

$$A = \frac{88q^4r}{(p-1)(r-1)(q-r)}, \quad B = 1, \quad \text{and} \quad C = \frac{11q^3r}{(r-1)(q-r)},$$

which yields

$$C_{p,q,r} = \frac{88pq^4r}{(p-1)(r-1)(q-r)}.$$

The claim that  $\mathcal{B}$  is of class  $C^1$  on  $\mathbb{D}$  should be understood in the sense that the function  $\mathcal{A}$  is continuous on  $[0, \infty)^3$ ,  $\mathcal{A}$  is continuously differentiable on  $(0, \infty)^3$ , and the partial derivatives of  $\mathcal{A}$  can be continuously extended to  $[0, \infty)^3$ . At a boundary point the differential  $d\mathcal{B}$  in (B4) is interpreted as the linear form whose coefficients are the aforementioned continuous extensions of partial derivatives to that point.

The motivation behind finding the explicit Bellman function (instead of just using the abstract one) is that in some contexts the explicit formula could be useful. For example, Carbonaro and Dragičević in [7] and [8] made use of the fact that the explicit Bellman function  $\mathcal{NT}$  involves powers. Another source of motivation is that we would also like to find a direct proof (without stopping time arguments) of the estimates for the “twisted” para-product considered by one of the authors in [15] or the “twisted” quadrilinear form considered by Durcik in [12] and [13]. This could also extend the range of exponents for a non-adapted stochastic integral considered by the authors in [16] or for the norm-variation of ergodic averages with respect to two commuting transformations [23]. So far we can only say that the Bellman function that has to be constructed for any of the mentioned problems should necessarily encode some structure of the function from Theorem 1.1, as dyadic para-products are the simplest and prototypical multilinear multipliers.

The Bellman function that we construct certainly does not give the best possible constants  $\mathcal{C}_{p,q,r}$  in (1.3). Indeed, the sharp constant for any triple of exponents from the generic range (1.4) has not yet been determined to the best of our knowledge. Search for the abstract Bellman function  $\mathbb{B}$  would lead us to the equations

$$(1.9) \quad \det \begin{bmatrix} \partial_u^2 \mathbb{B} & \partial_u \partial_v \mathbb{B} & \partial_u \partial_w \mathbb{B} & \partial_u \partial_U \mathbb{B} & \partial_u \partial_V \mathbb{B} & \partial_u \partial_W \mathbb{B} \\ \partial_u \partial_v \mathbb{B} & \partial_v^2 \mathbb{B} & \partial_v \partial_w \mathbb{B} \pm u & \partial_v \partial_U \mathbb{B} & \partial_v \partial_V \mathbb{B} & \partial_v \partial_W \mathbb{B} \\ \partial_u \partial_w \mathbb{B} & \partial_v \partial_w \mathbb{B} \pm u & \partial_w^2 \mathbb{B} & \partial_w \partial_U \mathbb{B} & \partial_w \partial_V \mathbb{B} & \partial_w \partial_W \mathbb{B} \\ \partial_u \partial_U \mathbb{B} & \partial_v \partial_U \mathbb{B} & \partial_w \partial_U \mathbb{B} & \partial_U^2 \mathbb{B} & \partial_U \partial_V \mathbb{B} & \partial_U \partial_W \mathbb{B} \\ \partial_u \partial_V \mathbb{B} & \partial_v \partial_V \mathbb{B} & \partial_w \partial_V \mathbb{B} & \partial_U \partial_V \mathbb{B} & \partial_V^2 \mathbb{B} & \partial_V \partial_W \mathbb{B} \\ \partial_u \partial_W \mathbb{B} & \partial_v \partial_W \mathbb{B} & \partial_w \partial_W \mathbb{B} & \partial_U \partial_W \mathbb{B} & \partial_V \partial_W \mathbb{B} & \partial_W^2 \mathbb{B} \end{bmatrix} = 0.$$

One way of simplifying (1.9) is to consider the non-homogeneous function  $\mathcal{B}$  of the form (1.7). The function  $\mathcal{B}$  is now a supersolution of the equation for the true Bellman function  $\mathbb{B}$ , but a function of that form can still yield the optimal (unknown) constant. This way (1.9) reduces to

$$(1.10) \quad \det \mathbb{A}_\pm = 0,$$

where  $\mathbb{A}_\pm$  are the matrices defined in (2.3) below. Alternatively, one can use the homogeneities of  $\mathbb{B}$  to reduce the dimension in (1.9). Equations like (1.10) can sometimes be turned into the Monge–Ampère equation by an appropriate change of variables, which does not seem to be the case here. At the moment, we do not know how to solve (1.10), so we impose slightly weaker conditions on our function  $\mathcal{B}$  that result in a constant  $\mathcal{C}_{p,q,r}$  which is not optimal. It would be interesting to find

a Bellman function  $\mathcal{B}$  that yields the optimal constant, or perhaps even the exact abstract Bellman function  $\mathbb{B}$ . Let us remark once again that this was achieved by Bañuelos and Osękowski [3] in the endpoint case  $p = \infty$ ,  $f \equiv 1$ .

We have organized the remainder of the paper as follows. In the next section we present the proof of Theorem 1.1. In Section 3 we apply Theorem 1.1 to reprove the well-known  $L^p$  estimates for martingale paraproducts and the heat flow paraproducts.

## 2. PROOF OF THEOREM 1.1

The continuity of  $\mathcal{A}$  on  $[0, \infty)^3$  is obvious. Indeed, observe that all exponents appearing in the definition of  $\mathcal{A}$  are positive. Thus,  $\mathcal{A}$  is clearly well-defined and continuous on each of the six closed regions determined by the inequalities for  $u, v, w$  and it is straightforward to verify that the six formulas are compatible on the common boundaries.

To see that  $\mathcal{A}$  is continuously differentiable on the open octant  $(0, \infty)^3$ , we just calculate the first order partial derivatives in the interior of each of the previously mentioned regions. The formula for each of these derivatives inside any of the regions continuously extends to the whole open octant. Moreover, these formulas coincide on the boundaries of each of the two adjacent regions, so we can deduce that  $\mathcal{A}$  really is of class  $C^1$  on  $(0, \infty)^3$ . For instance, both formulas for  $\frac{\partial \mathcal{A}}{\partial u}(u, v, w)$  at the common boundary of the two adjacent open regions  $v^q < u^p < w^r$  and  $u^p < v^q < w^r$ , which is a subset of  $u^p = v^q$ , simplify to  $Apu^{p-1}$ . All other cases are treated in the same manner.

Also, it is easy to see that the partial derivatives have limits at each point of the boundary of  $[0, \infty)^3$  and hence they can be continuously extended to  $[0, \infty)^3$ . For example, if  $0 < v^q \leq w^r \leq u^p$ , then the partial derivative of  $\mathcal{A}$  with respect to  $w$  equals

$$\frac{\partial \mathcal{A}}{\partial w}(u, v, w) = \frac{B(q-r)}{2} uv^2 w^{-r/q} + \frac{2Cr - B(q-r)}{2} uw^{r/q}.$$

Obviously, the only problematic points are the ones on the part of the boundary lying on the plane  $w = 0$ , but since  $v^q/w^r \leq 1$ , the limit as  $w \rightarrow 0$  still exists and equals zero. The existence of the other limits can be shown in a similar way.

The estimate (B2) follows directly from the definitions of the functions  $\mathcal{A}$  and  $\mathcal{B}$ , since

$$(A2) \quad 0 \leq \mathcal{A}(u, v, w) \leq Au^p + Bv^q + Cw^r$$

as long as  $A, B, C \geq 0$ . This is easily seen by using Young's inequality. The non-negativity of  $\mathcal{B}$  on  $\mathbb{D}$  is guaranteed if  $C_{p,q,r} \geq Ap, Bq, Cr$ .

Observe that (B3) is equivalent to

$$(A3) \quad \frac{1}{2} \mathcal{A}(u_1, v_1, w_1) + \frac{1}{2} \mathcal{A}(u_2, v_2, w_2) - \mathcal{A}(u, v, w) \geq u \frac{|v_1 - v_2|}{2} \frac{|w_1 - w_2|}{2},$$



where  $(u, v, w)$ ,  $(u_1, v_1, w_1)$ , and  $(u_2, v_2, w_2)$  are in  $[0, \infty)^3$  and such that

$$(2.1) \quad (u, v, w) = \frac{1}{2}(u_1, v_1, w_1) + \frac{1}{2}(u_2, v_2, w_2),$$

while  $(\mathcal{B}4)$  is equivalent to

$$(A4) \quad \mathcal{A}(u_1, v_1, w_1) \geq \mathcal{A}(u, v, w) + (d\mathcal{A})(u, v, w)(u_1 - u, v_1 - v, w_1 - w) + \frac{2}{3}u|v_1 - v||w_1 - w|,$$

where  $(u, v, w)$  and  $(u_1, v_1, w_1)$  are in  $[0, \infty)^3$ . Instead of proving  $(A3)$  and  $(A4)$  directly, we will reduce them conveniently to an inequality for quadratic forms.

Let  $(u, v, w) \in (0, \infty)^3$  be a point that does not lie on any of the three critical surfaces (1.8). This means that  $\mathcal{A}$  is of class  $C^2$  on an open ball around that point. If we take  $(u_1, v_1, w_1), (u_2, v_2, w_2)$  from that open ball such that (2.1) holds, then substituting  $u = (u_1 + u_2)/2, \Delta u = (u_1 - u_2)/2$ , etc., and adding Taylor's formulas at  $(u, v, w)$  for  $\mathcal{A}(u \pm \Delta u, v \pm \Delta v, w \pm \Delta w)$  gives us the infinitesimal version of  $(A3)$ :

$$(A3') \quad (d^2\mathcal{A})(u, v, w)(\Delta u, \Delta v, \Delta w) \geq 2u|\Delta v||\Delta w|.$$

Here  $d^2\mathcal{A}$  denotes the second differential of  $\mathcal{A}$  as a quadratic form, which we consider at the point  $(u, v, w)$  and apply to the vector  $(\Delta u, \Delta v, \Delta w)$ . Notice that  $(A3')$  does not hold on the whole domain of the function  $\mathcal{A}$ , which is  $[0, \infty)^3$ , but it does hold on the interior of each of the six regions into which the three surfaces divide  $(0, \infty)^3$ .

Conversely,  $(A3')$  implies  $(A3)$ , i.e. the two inequalities are equivalent for continuously differentiable functions, which is enabled by the convexity of the domain. To show the converse, first take a point  $(u, v, w) \in (0, \infty)^3$  and a vector  $(\Delta u, \Delta v, \Delta w) \in \mathbb{R}^3$  such that also  $(u \pm \Delta u, v \pm \Delta v, w \pm \Delta w) \in (0, \infty)^3$ . Now define the function  $\alpha: [-1, 1] \rightarrow \mathbb{R}$  as

$$(2.2) \quad \alpha(t) := \mathcal{A}(u + t\Delta u, v + t\Delta v, w + t\Delta w).$$

This function is continuously differentiable on  $[-1, 1]$  since  $\mathcal{A}$  is of class  $C^1$  on  $(0, \infty)^3$ . Also,  $\alpha$  is piecewise  $C^2$  on  $[-1, 1]$ . This follows from the facts that  $\mathcal{A}$  is of class  $C^2$  on  $(0, \infty)^3$  outside the surfaces (1.8), it has bounded second derivatives away from the coordinate planes  $u = 0, v = 0$ , and  $w = 0$ , and the segment  $\{(u + t\Delta u, v + t\Delta v, w + t\Delta w) : t \in [-1, 1]\}$  intersects the three critical surfaces at finitely many points. Using the integration by parts and the fundamental theorem of calculus (both in the versions for absolutely continuous functions; see [9]) gives us the equality

$$\frac{1}{2}\alpha(1) + \frac{1}{2}\alpha(-1) - \alpha(0) = \frac{1}{2} \int_{-1}^1 (1 - |t|)\alpha''(t)dt.$$

From the above identity we deduce

$$\begin{aligned} & \frac{1}{2}\mathcal{A}(u + \Delta u, v + \Delta v, w + \Delta w) + \frac{1}{2}\mathcal{A}(u - \Delta u, v - \Delta v, w - \Delta w) - \mathcal{A}(u, v, w) \\ &= \frac{1}{2} \int_{-1}^1 (1 - |t|)(d^2\mathcal{A})(u + t\Delta u, v + t\Delta v, w + t\Delta w)(\Delta u, \Delta v, \Delta w)dt. \end{aligned}$$

Finally, by  $(\mathcal{A}3')$  applied at all but finitely many points, the last expression is at least

$$\frac{1}{2} \int_{-1}^1 (1 - |t|)2(u + t\Delta u)|\Delta v||\Delta w|dt = u|\Delta v||\Delta w|,$$

which gives exactly  $(\mathcal{A}3)$ .

Moreover,  $(\mathcal{A}3')$  implies  $(\mathcal{A}4)$ . To verify this, we also take  $(u, v, w) \in (0, \infty)^3$  and  $(\Delta u, \Delta v, \Delta w) \in \mathbb{R}^3$  such that  $(u + \Delta u, v + \Delta v, w + \Delta w) \in (0, \infty)^3$ . We define  $\alpha: [0, 1] \rightarrow \mathbb{R}$  again by the formula (2.2). Integration by parts, the fundamental theorem of calculus, and  $(\mathcal{A}3')$  this time give

$$\alpha(1) = \alpha(0) + \alpha'(0) + \int_0^1 (1 - t)\alpha''(t)dt,$$

and therefore,

$$\begin{aligned} \mathcal{A}(u + \Delta u, v + \Delta v, w + \Delta w) &\geq \mathcal{A}(u, v, w) + (d\mathcal{A})(u, v, w)(\Delta u, \Delta v, \Delta w) \\ &\quad + \int_0^1 (1 - t)2(u + t\Delta u)|\Delta v||\Delta w|dt. \end{aligned}$$

Since  $u + t\Delta u = (1 - t)u + t(u + \Delta u) \geq (1 - t)u$ , the integral in  $t$  on the right-hand side is at least  $(2/3)u|\Delta v||\Delta w|$ , which establishes  $(\mathcal{A}4)$ .

This way we proved that  $(\mathcal{A}3')$  implies  $(\mathcal{A}3)$  and  $(\mathcal{A}4)$ , but only on  $(0, \infty)^3$ . To see that these two also hold on  $[0, \infty)^3$ , we just have to extend the obtained inequalities by the continuity of  $\mathcal{A}$  and  $d\mathcal{A}$ . We have commented in the introduction how we interpret  $d\mathcal{A}$  at the boundary of the domain.

Now we are left with proving  $(\mathcal{A}3')$ , which is equivalent to showing that the two matrices

$$(2.3) \quad \mathbb{A}_{\pm} = \begin{bmatrix} \partial_u^2 \mathcal{A} & \partial_u \partial_v \mathcal{A} & \partial_u \partial_w \mathcal{A} \\ \partial_u \partial_v \mathcal{A} & \partial_v^2 \mathcal{A} & \partial_v \partial_w \mathcal{A} \pm u \\ \partial_u \partial_w \mathcal{A} & \partial_v \partial_w \mathcal{A} \pm u & \partial_w^2 \mathcal{A} \end{bmatrix}$$

are positive semi-definite on each of the six open regions into which the surfaces (1.8) split  $(0, \infty)^3$ . To do so, we will use Sylvester's criterion and verify that all three principal minors are positive. More precisely, we will prove that the constants  $A, B, C$  can be chosen so that this is fulfilled.

We can simplify the calculations a bit by substituting  $t = v^q/u^p$ ,  $s = w^r/u^p$  and noting that

$$(2.4) \quad \mathcal{A}(u, v, w) = u^p \gamma(t, s),$$

where  $\gamma: (0, \infty)^2 \rightarrow \mathbb{R}$  is given by

$$\gamma(t, s) = \begin{cases} A + Bt + Cs, & 1 \leq s \leq t, \\ \frac{A(p-1)-C}{p-1} + Bt + \frac{Cp}{p-1} s^{1-1/p}, & s \leq 1 \leq t, \\ \frac{A(p-1)-(B+C)}{p-1} + \frac{Bp}{p-1} t^{1-1/p} + \frac{Cp}{p-1} s^{1-1/p}, & s \leq t \leq 1, \\ \frac{A(p-1)-(B+C)}{p-1} + \frac{Bq}{2} t^{2/q} s^{1/r-1/q} + \frac{2Cpr-Bp(q-r)}{2r(p-1)} s^{1-1/p}, & t \leq s \leq 1, \\ \frac{2Ar(p-1)-B(q+r)}{2r(p-1)} + \frac{Bq^2}{2p(q-2)} t^{2/q} + \frac{Bq(q-r)}{2r(q-2)} t^{2/q} s^{1-2/q} + \frac{2Cr-B(q-r)}{2r} s, & t \leq 1 \leq s, \\ A + \frac{Bq}{p(q-2)} t + \frac{Bq(q-r)}{2r(q-2)} t^{2/q} s^{1-2/q} + \frac{2Cr-B(q-r)}{2r} s, & 1 \leq t \leq s. \end{cases}$$

After plugging (2.4) into (2.3) and multiplying from both sides with the diagonal matrix  $\text{diag}(u^{1-p/2}, u^{p/q-p/2}, u^{p/r-p/2})$ , we obtain the matrices  $M = [m_{ij}]$ , where

$$\begin{aligned} m_{11} &= p(p-1)\gamma(t, s) - p(p-1)t\partial_t\gamma(t, s) - p(p-1)s\partial_s\gamma(t, s) \\ &\quad + 2p^2ts\partial_t\partial_s\gamma(t, s) + p^2t^2\partial_t^2\gamma(t, s) + p^2s^2\partial_s^2\gamma(t, s), \\ m_{12} &= m_{21} = -pqt^{1-1/q}s\partial_t\partial_s\gamma(t, s) - pqt^{2-1/q}\partial_t^2\gamma(t, s), \\ m_{13} &= m_{31} = -prts^{1-1/r}\partial_t\partial_s\gamma(t, s) - prs^{2-1/r}\partial_s^2\gamma(t, s), \\ m_{22} &= q(q-1)t^{1-2/q}\partial_t\gamma(t, s) + q^2t^{2-2/q}\partial_t^2\gamma(t, s), \\ m_{23} &= m_{32} = qrt^{1-1/q}s^{1-1/r}\partial_t\partial_s\gamma(t, s) \pm 1, \\ m_{33} &= r(r-1)s^{1-2/r}\partial_s\gamma(t, s) + r^2s^{2-2/r}\partial_s^2\gamma(t, s), \end{aligned}$$

and the problem is reduced to verifying that these matrices are positive definite on the interior of each of the six regions determined by the inequalities for  $t$  and  $s$ . First, we will calculate the three principal minors of the above matrices for each region, and then we will explain why we can choose the constants  $A, B, C$  such that all of them are positive.

The following expressions were calculated using *Mathematica* [26].

Region I:  $1 < s < t$ . Minor  $1 \times 1$ :  $\boxed{Ap(p-1)}$

Minor  $2 \times 2$ :  $\boxed{ABp(p-1)q(q-1)t^{1-2/q}}$

Determinants (with  $\pm$ ):

$$\boxed{ABCp(p-1)q(q-1)r(r-1)t^{1-2/q}s^{1-2/r}} - Ap(p-1)$$

Region 2:  $s < 1 < t$ . Minor  $1 \times 1$ :  $p(A(p-1) - C)$

Minor  $2 \times 2$ :  $Bp(A(p-1) - C)q(q-1)t^{1-2/q}$

Determinants (with  $\pm$ ):

$$BCp(A(p-1) - C)(q-1)r^2t^{1-2/q}s^{1/q-1/r} - BC^2q(q-1)r^2t^{1-2/q}s^{2/q} - p(A(p-1) - C)$$

Region 3:  $s < t < 1$ . Minor  $1 \times 1$ :  $p(A(p-1) - B - C)$

Minor  $2 \times 2$ :  $\frac{Bp(A(p-1) - B - C)q^2}{r}t^{1/r-1/q} - B^2q^2t^{2/r}$

Determinants (with  $\pm$ ):

$$BCp(Ap(q+r) - qr(B+C))t^{1/r-1/q}s^{1/q-1/r} - B^2Cqr^2t^{2/r}s^{1/q-1/r} - BC^2q^2rt^{1/r-1/q}s^{2/q} - p(A(p-1) - B - C) \pm 2BCqrt^{1/r}s^{1/q}$$

Region 4:  $t < s < 1$ . Minor  $1 \times 1$ :  $p(A(p-1) - B - C)$

Minor  $2 \times 2$ :  $Bp(A(p-1) - B - C)qs^{1/r-1/q} - B^2q^2t^{2/q}s^{2/r-2/q}$

Determinants (with  $\pm$ ):

$$\frac{(A(p-1) - B - C)(Bpr(2Cr - B(q-r)) - 2p)}{2} \pm B^2q(q-r)t^{3/q}s^{1/r-2/q} \mp 2Bp(A(p-1) - B - C)(q-r)t^{1/q}s^{-1/q} + \frac{B^3q(q-r)(3q-r)}{4}t^{4/q}s^{1/r-3/q} - \frac{B^2p(A(p-1) - B - C)(q-r)(2q-r)}{2}t^{2/q}s^{-2/q} \pm Bq(2Cr - B(q-r))t^{1/q}s^{1/r} - \frac{Bq(2Cr - B(q-r))^2}{4}s^{1/q+1/r} + \frac{B^2q(2Cr - B(q-r))(q-2r)}{2}t^{2/q}s^{1/r-1/q}$$

Region 5:  $t < 1 < s$ . Minor  $1 \times 1$ :

$$\frac{p(2Ar(p-1) - B(q+r))}{2r} + \frac{Bp(q-r)}{2r}t^{2/q}$$

Minor  $2 \times 2$ :

$$\boxed{\frac{Bpq(q-r)(2Ar(p-1) - B(q+r))}{2r^2(q-2)} s^{1-2/q}} - \frac{B^2q^2(pq+q-2p)t^{2/q}}{2p(q-2)} + \frac{B^2pq(q-r)^2}{2r^2(q-2)} t^{2/q} s^{1-2/q} + \boxed{\frac{Bq^2(2Ar(p-1) - B(q+r))}{2r(q-2)}}$$

Determinants (with  $\pm$ ):

$$\boxed{\frac{B(q-r)(p+q)(2Ar(p-1) - B(q+r))(2Cr - B(q-r))}{4r(q-2)} s^{2/p}} + \boxed{\frac{B(2Ar(p-1) - B(q+r))(2Cr - B(q-r))q^2(r-1)}{4r(q-2)} s^{1-2/r}} - \frac{p(2Ar(p-1) - B(q+r))}{2r} + \frac{B^2pq(q-r)^2(r-1)(2Cr - B(q-r))}{4r^2(q-2)} t^{2/q} s^{2/p} - \frac{B^2q^2(pq-2p+q)(r-1)(2Cr - B(q-r))}{4p(q-2)} t^{2/q} s^{1-2/r} - \frac{Bp(q-r)}{2r} t^{2/q} \mp \frac{B^2p(q-r)^2}{r} t^{3/q} s^{1/p-1/q} + \frac{B^2q(q-r)(q-p)(2Ar(p-1) - B(q+r))}{4p(q-2)} t^{2/q} s^{2/p-1} - \frac{B^3p(q-r)^3(qr-2r+q)}{4r^2(q-2)} t^{4/q} s^{2/p-2/q} - \frac{B^3qr(pq-2p+q)(q-r)(q-p)}{4p^2(q-2)} t^{4/q} s^{2/p-1} \mp \frac{Bp(q-r)(2Ar(p-1) - B(q+r))}{r} t^{1/q} s^{1/p-1/q} - \frac{B^2(q-r)^2(2Ar(p-1) - B(q+r))(2pq-3p-q)}{4r(q-2)} t^{2/q} s^{2/p-2/q}$$

Region 6:  $1 < t < s$ . Minor  $1 \times 1$ :  $\boxed{Ap(p-1)}$

Minor  $2 \times 2$ :

$$\boxed{\frac{ABpq(p-1)(q-r)}{r(q-2)} s^{1-2/q}} + \boxed{\frac{ABq^2(p-1)(q-1)}{q-2} t^{1-2/q}}$$

Determinants (with  $\pm$ ):

$$\begin{aligned}
 & \boxed{\frac{AB(2Cr - B(q-r))(p-1)(q-r)(p+q)}{2(q-2)} s^{2/p}} \\
 & - \frac{AB^2qr(p-1)(q-1)(q-r)(p-q)}{2p(q-2)} t s^{2/p-1} \\
 & + \left( \frac{AB(2Cr - B(q-r))qr(p-1)(q-1)(p+q)}{2p(q-2)} t^{1-2/q} s^{1-2/r} \right) - Ap(p-1) \\
 & - \frac{AB^2(p-1)(q-r)^2(2pq-3p-q)}{2(q-2)} t^{2/q} s^{2/p-2/q} \\
 & \mp 2ABp(p-1)(q-r)t^{1/q}s^{1/p-1/q}
 \end{aligned}$$

In each of the expressions there is a unique dominant term (regarding the exponents of  $t$  and  $s$ ) and it is double framed. We choose  $B$  arbitrarily (say  $B = 1$ ), then take  $C$  large enough (depending on  $p, q, r, B$ ), and finally take  $A$  large enough (depending on  $p, q, r, B, C$ ). While doing so, we take care that the coefficient of the double framed term is greater than the sum of the absolute values of coefficients of the terms that are neither framed nor circled. We can do so because by taking  $C$  large enough the expression multiplying  $A$  in the coefficient of the dominant term can be made larger than the sum of the absolute values of the corresponding expressions in other non-circled terms that contain  $A$ . Consequently, the coefficient of the dominant term grows faster than the sum of the absolute values of the coefficients in the other terms as  $A$  tends to infinity. This means that we can take  $A$  large enough so that the dominant term actually dominates the sum of all other non-framed and non-circled terms in each expression. Another way of phrasing the argument that sufficiently large  $A$  and  $C$  make six considered determinantal expressions positive is to observe that each dominant term contains the product  $AC$ , as opposed to any other non-circled term.

The only problematic terms that we cannot dominate with the dominant term are the circled ones, because of their uncontrollable growth in  $A$ . However, just by taking

$$C \geq B(q-r)/(2r) \quad \text{and} \quad A \geq B(q+r)/(2r(p-1))$$

we make sure that all of them are non-negative, so they only contribute to the positivity of the expressions.

To explain how the values of the coefficients  $A, B$ , and  $C$  in Theorem 1.1 were obtained, let us consider Region 4 as a representative example. The other regions are treated similarly.

First, notice that the double framed term really is the dominant one, since  $t < s < 1$  implies

$$\frac{3}{t^q s^r} - \frac{2}{q}, \frac{1}{t^q s} - \frac{1}{q}, \frac{4}{t^q s^r} - \frac{3}{q}, \frac{2}{t^q s} - \frac{2}{q}, \frac{1}{t^q s^r}, s^{\frac{1}{q} + \frac{1}{r}}, \frac{2}{t^q s^r} - \frac{1}{q} < 1 = t^0 s^0.$$

We can choose  $B = 1$  and then take  $C$  large enough such that

$$r(2Cr - q + r) > \max\{28(q - r) + 2, 7(q - r)(2q - r) + 2\}.$$

Clearly,  $C = 11q^3r / ((r - 1)(q - r))$  satisfies the above condition. This way the expression multiplying  $A$  in the coefficient of the dominant term is seven times larger than the expressions multiplying  $A$  in the coefficients of the two non-framed terms that contain  $A$ . Now we just have to take  $A$  large enough such that

$$(A(p - 1) - C - 1)(pr(2Cr - q + r) - 2p)$$

is at least

$$\max \left\{ \frac{7}{2}q(2Cr - q + r)^2, 14q(2Cr - q + r), 7q(2Cr - q + r)|q - 2r|, \right. \\ \left. 14q(q - r), \frac{7}{2}q(q - r)(3q - r) \right\}.$$

It is easy to see that  $A = 88q^4r / ((p - 1)(r - 1)(q - r))$  is one possible choice. Now the dominant term is more than seven times larger than the absolute value of any other term, which means that the dominant term dominates the sum of all other terms.

This way we accomplish the positivity of each of the expressions, which is exactly what we needed and the proof of  $(\mathcal{A}3')$  is completed. This also completes the proof of Theorem 1.1.

In the next section, it will sometimes be more convenient to use the infinitesimal version of  $(\mathcal{B}3)$ :

$(\mathcal{B}3')$

$$- (d^2\mathcal{B})(u, v, w, U, V, W)(\Delta u, \Delta v, \Delta w, \Delta U, \Delta V, \Delta W) \geq 2u|\Delta v||\Delta w|.$$

Again,  $(\mathcal{B}3')$  holds only for points  $(u, v, w, U, V, W)$  at which the second differential of  $\mathcal{B}$  is well defined, i.e. for the points such that  $(u, v, w)$  does not lie on any of the three critical surfaces. The equivalence of  $(\mathcal{B}3')$  and  $(\mathcal{B}3)$  follows from the equivalence of  $(\mathcal{A}3')$  and  $(\mathcal{A}3)$ .

### 3. APPLICATIONS

Here we present several applications of the existence of the Bellman function from Theorem 1.1. We need to emphasize that the following problems are quite classical and can be solved using more standard tools. We only provide quite straightforward solutions based on Theorem 1.1. Moreover, only the existence of the Bellman function with properties  $(\mathcal{B}1)$ – $(\mathcal{B}3)$  is needed, even though  $(\mathcal{B}4)$  is quite convenient in Subsection 3.1. This existence can also follow if boundedness of the dyadic paraproduct is established in some other way, as commented in

the introduction. However, our goal is to illustrate how several classical problems become methodologically simple once we explicitly construct the function as in Theorem 1.1.

For two non-negative quantities  $A$  and  $B$  we will write  $A \lesssim_P B$  if there exists a finite constant  $C_P \geq 0$  depending on a set of parameters  $P$  such that  $A \leq C_P B$ .

**3.1. Discrete-time martingales.** Let us consider two martingales  $X = (X_n)_{n=0}^\infty$  and  $Y = (Y_n)_{n=0}^\infty$  with respect to the same filtration  $(\mathcal{F}_n)_{n=0}^\infty$ . Their *paraproduct* is a stochastic process  $((X \cdot Y)_n)_{n=0}^\infty$  defined as

$$(3.1) \quad (X \cdot Y)_0 := 0, \quad (X \cdot Y)_n := \sum_{k=1}^n X_{k-1}(Y_k - Y_{k-1}) \quad \text{for } n \geq 1.$$

This process can be regarded as a particular case of Burkholder's martingale transform [5] of the martingale  $Y$  with respect to the shifted adapted process  $X$ . We have also imposed the martingale property on  $X$ , since we want to treat  $X$  and  $Y$  symmetrically and since this is required by the existence of the  $L^p$  estimates in the interior of the Banach triangle in Figure 1. We want to prove that for the exponents  $p, q, r$  satisfying (1.4) the estimate

$$(3.2) \quad \|(X \cdot Y)_n\|_{L^{r'}} \lesssim_{p,q,r} \|X_n\|_{L^p} \|Y_n\|_{L^q}$$

holds uniformly in the positive integer  $n$ , where  $r'$  is the conjugate exponent of  $r$ . Instead of proving (3.2) directly, we will rather show the estimate for the dualized form, i.e. that for an arbitrary random variable  $Z \in L^r$  the inequality

$$(3.3) \quad |\mathbb{E}((X \cdot Y)_n Z)| \lesssim_{p,q,r} \|X_n\|_{L^p} \|Y_n\|_{L^q} \|Z\|_{L^r}$$

holds. This inequality is trivial unless all norms on the right-hand side are finite.

Let us introduce the third martingale  $(Z_n)_{n=0}^\infty$  with  $Z_n := \mathbb{E}(Z | \mathcal{F}_n)$ . By splitting  $Z = Z_{k-1} + (Z_k - Z_{k-1}) + (Z - Z_k)$  and using the martingale property in the form of  $\mathbb{E}(Y_k - Y_{k-1} | \mathcal{F}_{k-1}) = 0$  and  $\mathbb{E}(Z - Z_k | \mathcal{F}_k) = 0$ , we can write

$$\begin{aligned} \mathbb{E}((X \cdot Y)_n Z) &= \sum_{k=1}^n \mathbb{E}(X_{k-1}(Y_k - Y_{k-1})Z) \\ &= \sum_{k=1}^n \mathbb{E}(X_{k-1}(Y_k - Y_{k-1})(Z_k - Z_{k-1})). \end{aligned}$$

The estimate (3.3) is now a clear consequence of the Cauchy–Schwarz, Hölder, Doob and Burkholder–Gundy inequalities. Again, we will give a more direct proof using the Bellman function (1.7).

It is enough to consider the times  $k = 0, 1, \dots, n$ , but we need to show the estimate that is uniform in  $n$ . We can assume that  $X_k, Y_k, Z_k \geq 0$  for  $0 \leq k \leq n$ ,



as otherwise we split the variables  $X_n, Y_n, Z_n$  into positive and negative parts, and introduce three new martingales (for a fixed  $n$ ):

$$U_k := \mathbb{E}(X_n^p | \mathcal{F}_k), \quad V_k := \mathbb{E}(Y_n^q | \mathcal{F}_k), \quad W_k := \mathbb{E}(Z_n^r | \mathcal{F}_k).$$

If we write  $\mathbf{X}_k = (X_k, Y_k, Z_k, U_k, V_k, W_k)$ , then property (B4) of the Bellman function  $\mathcal{B}$  gives us

$$\begin{aligned} \mathcal{B}(\mathbf{X}_{k-1}) + (d\mathcal{B})(\mathbf{X}_{k-1})(\mathbf{X}_k - \mathbf{X}_{k-1}) \\ \geq \mathcal{B}(\mathbf{X}_k) + \frac{2}{3} X_{k-1} |Y_k - Y_{k-1}| |Z_k - Z_{k-1}|, \end{aligned}$$

from which we deduce

$$\mathcal{B}(\mathbf{X}_{k-1}) \geq \mathbb{E}(\mathcal{B}(\mathbf{X}_k) | \mathcal{F}_{k-1}) + \frac{2}{3} \mathbb{E}(X_{k-1} |Y_k - Y_{k-1}| |Z_k - Z_{k-1}| | \mathcal{F}_{k-1}),$$

by taking the conditional expectation with respect to  $\mathcal{F}_{k-1}$  and using the martingale property. Finally, taking the expectation of the above inequality, summing over  $k = 1, \dots, n$ , telescoping, and using (B2) gives

$$\begin{aligned} \frac{2}{3} \sum_{k=1}^n \mathbb{E}(X_{k-1} |Y_k - Y_{k-1}| |Z_k - Z_{k-1}|) &\leq \mathbb{E}\mathcal{B}(\mathbf{X}_0) - \mathbb{E}\mathcal{B}(\mathbf{X}_n) \\ &\leq C_{p,q,r} \mathbb{E} \left( \frac{1}{p} U_0 + \frac{1}{q} V_0 + \frac{1}{r} W_0 \right) = C_{p,q,r} \left( \frac{1}{p} \|X_n\|_{L^p}^p + \frac{1}{q} \|Y_n\|_{L^q}^q + \frac{1}{r} \|Z_n\|_{L^r}^r \right). \end{aligned}$$

Homogenizing the above inequality, we get the desired estimate (3.3) and hence also (3.2).

**3.2. Continuous-time martingales.** Let  $X = (X_t)_{t \geq 0}$  and  $Y = (Y_t)_{t \geq 0}$  be two continuous-time càdlàg martingales with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  that satisfies the “usual hypotheses” [22]. In this case the *martingale paraproduct* is also a stochastic process  $((X \cdot Y)_t)_{t \geq 0}$ , but now defined via the stochastic integral

$$(3.4) \quad (X \cdot Y)_t := \int_{0+}^t X_{s-} dY_s.$$

Since we are allowed to choose dense subspaces on which the initial definition makes sense (and later extend by continuity), we can conveniently assume that  $X$  is bounded in  $L^\infty$  and  $Y$  is bounded in  $L^2$ . We want to prove that (3.4) satisfies the same  $L^p$  estimates as (3.1). To do so, we take  $(\pi_m)_{m=1}^\infty$  to be a refining sequence of partitions

$$0 = t_0^{(m)} < t_1^{(m)} < t_2^{(m)} < \dots < t_{n(m)}^{(m)} = t$$

such that  $\lim_{m \rightarrow \infty} \text{mesh}(\pi_m) = 0$ . We can calculate (3.4) as the limit of the Riemann sums in the following way:

$$(3.5) \quad \int_{0+}^t X_{s-} dY_s = \lim_{m \rightarrow \infty} \sum_{k=1}^{n(m)} X_{t_{k-1}^{(m)}} (Y_{t_k^{(m)}} - Y_{t_{k-1}^{(m)}}).$$

The above limit is interpreted as the convergence in probability; for more details see [22]. Notice that the right-hand side of (3.5) is actually a limit of discrete-time martingale paraproducts (3.1). By passing to an a.s. convergent subsequence, using Fatou's lemma, and applying (3.2), we get the desired estimate for (3.4):

$$\|(X \cdot Y)_t\|_{L^{r'}} \leq \sup_m \left\| \sum_{k=1}^{n(m)} X_{t_{k-1}^{(m)}} (Y_{t_k^{(m)}} - Y_{t_{k-1}^{(m)}}) \right\|_{L^{r'}} \lesssim_{p,q,r} \|X_t\|_{L^p} \|Y_t\|_{L^q}$$

for the exponents  $p, q, r$  satisfying (1.4).

As a special case we can consider martingales with respect to the augmented filtration of the one-dimensional Brownian motion  $(B_t)_{t \geq 0}$ . If we also assume that  $Y_0 = 0$ , then

$$(3.6) \quad (X \cdot Y)_t = \int_0^t X_s dY_s,$$

because  $(X_t)_{t \geq 0}$  and  $(Y_t)_{t \geq 0}$  now a.s. have continuous paths. We remark that (3.6) are the martingale paraproducts studied by Bañuelos and Bennett in [2] and they established  $L^p$ ,  $H^p$ , and BMO estimates for (3.6). Their proof of the  $L^p$  estimates uses Doob's inequality and the Burkholder–Gundy inequality.

Yet another short proof of the  $L^p$  estimates in this particular case can be given by applying Itô's formula in combination with (B2) and (B3'), instead of approximating by discrete-time processes. However, for that purpose our Bellman function should be of class  $C^2$  on the whole domain. This is achieved by shrinking the domain slightly and passing to  $\mathcal{B}_\varepsilon$  as in the next section; we omit the details.

**3.3. Heat flow paraproducts.** In order to be able to use the constructed Bellman function in relationship with the heat equation, we should first “smoothen it up”. Let us fix a non-negative even  $C^\infty$  function  $\varphi$  supported in  $(-1, 1)^3$  with integral one. For any  $\varepsilon > 0$  we define the function  $\mathcal{A}_\varepsilon: (\varepsilon, \infty)^3 \rightarrow \mathbb{R}$  by the formula

$$\mathcal{A}_\varepsilon(u, v, w) := \int_{(-\varepsilon, \varepsilon)^3} \varepsilon^{-3} \varphi(\varepsilon^{-1}a, \varepsilon^{-1}b, \varepsilon^{-1}c) \mathcal{A}(u-a, v-b, w-c) da db dc.$$

In words, the function  $\mathcal{A}_\varepsilon$  is the convolution of  $\mathcal{A}$  with the  $L^1$ -normalized dilation of  $\varphi$ . The newly obtained function is clearly of class  $C^\infty$ . We integrate (A3) translated by  $(a, b, c)$  and multiplied by  $\varepsilon^{-3} \varphi(\varepsilon^{-1}a, \varepsilon^{-1}b, \varepsilon^{-1}c)$ , and then “symmetrize” in  $(a, b, c)$  and use the fact that  $\varphi$  is even. That way we conclude that

$\mathcal{A}_\varepsilon$  still satisfies the condition (A3) and consequently also (A3') at every point of its domain. By the formula (1.7) with  $\mathcal{A}_\varepsilon$  in the place of  $\mathcal{A}$  we can define a  $C^\infty$  function  $\mathcal{B}_\varepsilon$  satisfying property (B3') for any  $u, v, w > \varepsilon$  and  $U \geq u^p, V \geq v^q, W \geq w^r$ . Moreover, property (A2) is retained up to an additional loss by the factor  $\max\{2^p, 2^q, 2^r\}$ , which in turn guarantees (B2) for some (sufficiently large) constant  $C_{p,q,r}$  independent of  $\varepsilon$ .

Now suppose that  $f, g, h$  are compactly supported  $C^\infty$  functions on  $\mathbb{R}$ . Also, let  $k(x, t) := \frac{1}{\sqrt{2\pi t}} \exp(-x^2/(2t))$  be the heat kernel on the real line and  $u$  be the heat extension of  $f$ :

$$u(x, t) := \int_{\mathbb{R}} f(y)k(x - y, t)dy.$$

Note that  $u$  is the solution of the heat equation  $\partial_t u = \frac{1}{2}\partial_x^2 u$  with the initial condition  $\lim_{t \rightarrow 0^+} u(x, t) = f(x)$ . Analogously we define  $v$  and  $w$  to be the heat extensions of  $g$  and  $h$ .

We can define the *heat paraproduct*, i.e. the paraproduct with respect to the heat semigroup as a trilinear form

$$(3.7) \quad \Lambda(f, g, h) := \int_{\mathbb{R}} \int_0^\infty u(x, t) \partial_x v(x, t) \partial_x w(x, t) dt dx.$$

If we define

$$\varphi_s(x) := k(x, s^2), \quad \psi_s(x) := -2^{1/2} s \partial_x k(x, s^2)$$

and substitute  $t = s^2$ , we get a more familiar expression:

$$(3.8) \quad \Lambda(f, g, h) = \int_{\mathbb{R}} \int_0^\infty (f * \varphi_s)(x) (g * \psi_s)(x) (h * \psi_s)(x) \frac{ds}{s} dx.$$

Smooth paraproducts like (3.8) appear naturally in the proof of the T1 theorem (see [10]), although one usually needs to be more flexible when choosing a bump function  $\varphi_s$  and a mean zero bump function  $\psi_s$ .

Again, we want to prove some  $L^p$  estimates for (3.7), i.e.

$$|\Lambda(f, g, h)| \lesssim_{p,q,r} \|f\|_{L^p(\mathbb{R})} \|g\|_{L^q(\mathbb{R})} \|h\|_{L^r(\mathbb{R})},$$

where  $p, q, r$  are exponents satisfying (1.4). To do so we will imitate the ‘‘heating’’ technique by Nazarov and Volberg [19] or Petermichl and Volberg [21].

Assume that  $f, g, h$  are non-negative and that none of them is identically zero. Fix  $R > 0, \delta > 0, T > 2\delta$ , and observe that  $u(x, t), v(x, t), w(x, t) > \varepsilon$  whenever  $x \in [-R, R], t \in [\delta, T - \delta]$  for some sufficiently small  $\varepsilon > 0$  depending on  $R, \delta, T$ , and the functions  $f, g, h$ . We introduce  $U, V, W$  as the heat extensions of  $f^p, g^q, h^r$  respectively and define

$$b(x, t) := \mathcal{B}_\varepsilon(u(x, t), v(x, t), w(x, t), U(x, t), V(x, t), W(x, t)),$$

where  $\mathcal{B}_\varepsilon$  is as above. It is easy to calculate that

$$\begin{aligned} (\partial_t - \frac{1}{2}\partial_x^2)b(x, t) &= (\nabla\mathcal{B}_\varepsilon)(u, v, w, U, V, W) \cdot (\partial_t - \frac{1}{2}\partial_x^2)(u, v, w, U, V, W) \\ &\quad - \frac{1}{2}(d^2\mathcal{B}_\varepsilon)(u, v, w, U, V, W)(\partial_x u, \partial_x v, \partial_x w, \partial_x U, \partial_x V, \partial_x W). \end{aligned}$$

(We have omitted writing the variables  $x, t$  on the right-hand side.) Since  $u, v, w, U, V, W$  all satisfy the heat equation, the first term on the right-hand side is zero and by  $(\mathcal{B}3')$  we get

$$(\partial_t - \frac{1}{2}\partial_x^2)b(x, t) \geq \pm u(x, t) \partial_x v(x, t) \partial_x w(x, t).$$

It remains to integrate this inequality over  $[-R, R] \times [\delta, T - \delta]$  with an appropriate weight, use Green's formula, and then let  $\delta \rightarrow 0, R, T \rightarrow \infty$ . We omit the details and refer to [19] and [21].

Let us emphasize once again that the previous trick of "smoothing" the Bellman function was already used in [19] and [21] and no explicit formula is needed for its application.

**Acknowledgments.** We would like to thank the anonymous referee for several useful comments and suggestions that improved the readability of this paper.

#### REFERENCES

- [1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, Inc., New York 1992.
- [2] R. Bañuelos and A. G. Bennett, *Paraproducts and commutators of martingale transforms*, Proc. Amer. Math. Soc. 103 (4) (1988), pp. 1226–1234.
- [3] R. Bañuelos and A. Osękowski, *On the Bellman function of Nazarov, Treil and Volberg*, Math. Z. 278 (2014), pp. 385–399.
- [4] Á. Bényi, D. Maldonado, and V. Naibo, *What is ... a paraproduct?*, Notices Amer. Math. Soc. 57 (7) (2010), pp. 858–860.
- [5] D. L. Burkholder, *Martingale transforms*, Ann. Math. Statist. 37 (1966), pp. 1494–1504.
- [6] D. L. Burkholder, *Boundary value problems and sharp inequalities for martingale transforms*, Ann. Probab. 12 (3) (1984), pp. 647–702.
- [7] A. Carbonaro and O. Dragičević, *Bellman function and linear dimension-free estimates in a theorem of Bakry*, J. Funct. Anal. 265 (7) (2013), pp. 1085–1104.
- [8] A. Carbonaro and O. Dragičević, *Functional calculus for generators of symmetric contraction semigroups*, Duke Math. J. 166 (5) (2017), pp. 937–974.
- [9] D. L. Cohn, *Measure Theory*, second edition, Birkhäuser/Springer, New York 2013.
- [10] G. David and J.-L. Journé, *A boundedness criterion for generalized Calderón–Zygmund operators*, Ann. of Math. (2) 120 (2) (1984), pp. 371–397.
- [11] B. Davis, *On the  $L^p$  norms of stochastic integrals and other martingales*, Duke Math. J. 43 (4) (1976), pp. 697–704.
- [12] P. Durcik, *An  $L^4$  estimate for a singular entangled quadrilinear form*, Math. Res. Lett. 22 (5) (2015), pp. 1317–1332.
- [13] P. Durcik,  *$L^p$  estimates for a singular entangled quadrilinear form*, Trans. Amer. Math. Soc. 369 (10) (2017), pp. 6935–6951.
- [14] S. Janson and J. Peetre, *Paracommutators—boundedness and Schatten–von Neumann properties*, Trans. Amer. Math. Soc. 305 (2) (1988), pp. 467–504.

- [15] V. Kovač, *Boundedness of the twisted paraproduct*, Rev. Mat. Iberoam. 28 (4) (2012), pp. 1143–1164.
- [16] V. Kovač and K. A. Škreb, *One modification of the martingale transform and its applications to paraproducts and stochastic integrals*, J. Math. Anal. Appl. 426 (2) (2015), pp. 1143–1163.
- [17] F. L. Nazarov and S. R. Treil, *The hunt for a Bellman function: Applications to estimates of singular integral operators and to other classical problems in harmonic analysis* (in Russian), Algebra i Analiz 8 (1996), pp. 32–162; English transl. in St. Petersburg Math. J. 8 (1997), pp. 721–824.
- [18] F. L. Nazarov and A. Volberg, *The Bellman function, the two-weight Hilbert transform, and embeddings of the model spaces  $K_\theta$* , J. Anal. Math. 87 (2002), pp. 385–414.
- [19] F. L. Nazarov and A. Volberg, *Heating of the Ahlfors–Beurling operator and estimates of its norm*, St. Petersburg Math. J. 15 (2004), pp. 563–573.
- [20] A. Osękowski, *Sharp Martingale and Semimartingale Inequalities*, Birkhäuser/Springer, Basel 2012.
- [21] S. Petermichl and A. Volberg, *Heating of the Ahlfors–Beurling operator: Weakly quasiregular maps on the plane are quasiregular*, Duke Math. J. 112 (2) (2002), pp. 281–305.
- [22] P. E. Protter, *Stochastic Integration and Differential Equations*, second edition, Springer, Berlin 2005.
- [23] T. Tao, *Norm convergence of multiple ergodic averages for commuting transformations*, Ergodic Theory Dynam. Systems 28 (2) (2008), pp. 657–688.
- [24] C. Thiele, *Wave Packet Analysis*, American Mathematical Society, Providence, RI, 2006.
- [25] V. Vasyunin and A. Volberg, *Bellster and others* (2008), preprint.
- [26] Wolfram Research, Inc., *Mathematica, Version 9.0*, Champaign, IL (2012).

Vjekoslav Kovač  
Department of Mathematics, Faculty of Science  
University of Zagreb  
Bijenička cesta 30  
10000 Zagreb, Croatia  
E-mail: vjekovac@math.hr

Kristina Ana Škreb  
Faculty of Civil Engineering  
University of Zagreb  
Fra Andrije Kačića Miošića 26  
10000 Zagreb, Croatia  
E-mail: kskreb@grad.hr

Received on 22.9.2016;  
revised version on 3.7.2017