SHARP L^P–BOUNDS FOR A SMALL PERTURBATION OF BURKHOLDER'S MARTINGALE TRANSFORM

NICHOLAS BOROS, PRABHU JANAKIRAMAN, AND ALEXANDER VOLBERG

ABSTRACT. Let $\{d_k\}_{k\geq 0}$ be a complex martingale difference in $L^p[0,1]$, where $1<p<\infty$, and $\{\varepsilon_k\}_{k\geq 0}$ a sequence in $\{\pm 1\}$. We obtain the following generalization of Burkholder's famous result. If $\tau\in[-\frac{1}{2},\frac{1}{2}]$ and $n\in\mathbb{Z}_+$ then

$$\left\|\sum_{k=0}^{n}\begin{pmatrix}\varepsilon_k\\ \tau\end{pmatrix}d_k\right\|_{L^p([0,1),\mathbb{C}^2)}\leq((p^*-1)^2+\tau^2)^{\frac{1}{2}}\left\|\sum_{k=0}^{n}d_k\right\|_{L^p([0,1),\mathbb{C})},$$

where $((p^*-1)^2+\tau^2)^{\frac{1}{2}}$ is sharp and $p^*-1=\max\{p-1,\frac{1}{p-1}\}$. For $2\leq p<\infty$ the result is also true with sharp constant for $|\tau|\leq 1$.

1. Introduction

Determining the L^p operator norm of singular integrals is quite difficult, in many cases. While the operator norm of the Hilbert transform can be computed by means of classical techniques, see Pichorides [17], the Ahlfors–Beurling operator (the two dimensional analog), denoted T, cannot. It was shown in 1965, by Lehto [15], that $\|T\|_{p\to p}\geq p^*-1$. Iwaniec conjectured in 1982, see [14], that $\|T\|_{p\to p}=p^*-1$. The only progress toward showing the validity of this conjecture (see Nazarov, Volberg [16] and Bañuelos, Janakiraman [3] for two of the large steps toward this) has been using the fact that Burkholder computed the L^p operator norm of the martingale transform in [7]. But, the estimates of martingale transform can also be used for determining lower bounds, for example, for $\Re T$ and $\Im T$ in Geiss, Montgomery-Smith, Saksman [13]. Operator T itself is a linear combination (of course) of $\Re T$ and $\Im T$. So we are interested in linear combinations of all second order Riesz transforms. One can start with investigation of linear combination $aR_1^2+bR_2^2$ of Riesz transforms (leaving R_1R_2 alone for a while).

Actually Geiss, Montgomery-Smith, Saksman [13] shows that if one wants to estimate any linear combination $aR_1^2+bR_2^2$ of Riesz transforms, then one needs to estimate a corresponding linear combination of Burkholder's martingale transform and the identity operator. "Corresponding" here means the following. Notice that

$aR_1^2 + bR_2^2 = \frac{a-b}{2}(R_1^2 - R_2^2) + \frac{a+b}{2} Id$. Essentially we come to the need to estimate (compute) the norm of $(R_1^2 - R_2^2) + \tau \cdot Id$, where τ is an arbitrary constant. Now Geiss, Montgomery-Smith, Saksman [13] proved that the norm $(R_1^2 - R_2^2) + \tau \cdot Id$ in $L^p(\mathbb{C})$ is bounded below by $MT + \tau \cdot Id$ in $L^p[0,1]$ (see (1.1) below for a definition of MT). This is not formulated directly in [13], but it is easy to extract this claim from [13].

The problem of computing the norm of $MT + \tau \cdot Id$ in L^p seems to be *very difficult.* It is done in two cases: 1) $\tau = 0$ in [7], 2) $\tau = \pm 1$, Choi [12]. For all other τ's it is still open, and even though we have some approach to it, it seems interesting that another type of perturbation of MT, namely *the quadratic perturbation* considered in the present paper (and also in [1], [2]), seems to be relatively easy to handle. This article will focus on setting up the Bellman function and using it to solve the problem, but showing very little detail in the actual computation of the Bellman function. For full details of the computation of the Bellman function, refer to [1].

Therefore, if we can determine the operator norm of a quadratic perturbation of the martingale transform then we can also determine quadratic perturbations of singular integrals as an application.

To prove the main result we are going to take a slightly indirect approach. Burkholder (see [7]) defined the martingale transform, MT_ε, as

$$MT_\varepsilon \left(\sum_{k=1}^n d_k \right) := \sum_{k=1}^n \varepsilon_k d_k. \tag{1.1}$$

Then the main result can be stated as

$$\sup_{\vec{\varepsilon}} \left\| \begin{pmatrix} MT_{\vec{\varepsilon}} \\ \tau I \end{pmatrix} \right\|_{L^p(\mathbb{C}) \to L^p(\mathbb{C}^2)} = \sup_{\vec{\varepsilon}} \frac{\left\| \sum_{k=1}^n \begin{pmatrix} \varepsilon_k \\ \tau \end{pmatrix} d_k \right\|_p}{\| \sum_{k=1}^n d_k \|_p} = ((p^* - 1)^2 + \tau^2)^{\frac{1}{2}},$$

where I is the identity transformation and τ is "small". However, rather than working with this martingale transform in terms of the martingale differences, in a probabilistic setting, we will define another martingale transform in terms of the Haar expansion of $L^p[0,1]$ functions and set up a Bellman function in that context. Burkholder showed, in [11], that these two different martingale transforms have the same L^p operator norm, for $\tau = 0$, so we expected a perturbation of these to act similarly, and they do. For convenience, we will work with the martingale transform in the Haar setting. Using the Bellman function technique will turn the problem of finding the sharp constant of the above estimate into solving a second order partial differential equation. The beauty of this approach is that it gets right to the heart of the problem with very little advanced techniques needed in the process. In fact, the only background material that is needed for the Bellman function technique approach, is some basic knowledge of partial differential equations and some elementary analysis.

Observe that for $2 \leq p < \infty$, the estimate from above in the main result is just an application of Minkowski's inequality on $L^{\frac{p}{2}}$ and Burkholder's original result. But, this argument doesn't address sharpness, even though the constant obtained turns out to be the sharp constant for small τ. For $1 < p < 2$, Minkowski's inequality (in $l^{\frac{2}{p}}$) also plays a role, but to a lesser extent and cannot give the sharp constant, as we will see Proposition 17. We will now rigorously develop some background ideas needed to set up the Bellman function.

1.1. **Motivation of the Bellman function.** Let I be an interval and $\alpha^{\pm} \in \mathbb{R}^{+}$ such that $\alpha^{+} + \alpha^{-} = 1$. These $\alpha^{\pm}$ generate two subintervals $I^{\pm}$ such that $|I^{\pm}| = \alpha^{\pm}|I|$ and $I = I^{-} \cup I^{+}$. We can continue this decomposition indefinitely as follows. For any sequence $\{\alpha_{n,m} : 0 < \alpha_{n,m} < 1, 0 \leq m < 2^{n}, 0 < n < \infty, \alpha_{n,2k} + \alpha_{n,2k+1} = 1\}$, we can generate the sequence $\mathcal{I} := \{I_{n,m} : 0 \leq m < 2^{n}, 0 < n < \infty\}$, where $I_{n,m} = I^{-}_{n,m} \cup I^{+}_{n,m} = I_{n+1,2m+1} \cup I_{n+1,2m+1}$ and $\alpha^{-} = \alpha_{n+1,2m}, \alpha^{+} = \alpha_{n+1,2m+1}$. Note that $I_{0,0} = I$.

For any $J \in \mathcal{I}$ we define the Haar function $h_J := -\sqrt{\frac{\alpha^{+}}{\alpha^{-}|J|}}\chi_{J^{-}} + \sqrt{\frac{\alpha^{-}}{\alpha^{+}|J|}}\chi_{J^{+}}$. If $\max\{|I_{n,m}| : 0 \leq m < 2^{n}\} \to 0$ as $n \to \infty$ then $\{h_J\}_{J\in\mathcal{I}}$ is an orthonormal basis for $L^2_0(I) := \{f \in L^2(I) : \int_I f = 0\}$. However, if we add one extra function then Haar functions form an orthonormal basis in $L^2[0,1]$ function. Fix $I_0 = [0,1]$ and $\mathcal{I} = \mathcal{D}$ as the dyadic subintervals of I_0. Let $\mathcal{D}_j = \{I \in \mathcal{D} : |I| = 2^{-j}\}$. We use the notation $\langle f \rangle_J$ to represent the average integral of f over the interval $J \in \mathcal{D}$ and $\Delta_j f = \langle f \rangle_{I_{j+1}} - \langle f \rangle_{I_j}$, where $I_j \in \mathcal{D}_j$ and $I_{j+1} \in \mathcal{D}_{j+1}$. For any $f \in L^1(I_0)$ we have $\Delta_j f = \sum_{I\in\mathcal{D}_j} (f, h_I)h_I$. Then

$$\sum_{j=0}^{\infty} \Delta_j f = \lim_{N\to\infty} \langle f \rangle_{I_{N+1}} - \langle f \rangle_{I_0} \tag{1.2}$$

By Lebesgue differentiation, the limit in (1.2) converges to f almost everywhere as $N \to \infty$. So any $f \in L^p(I_0) \subset L^1(I_0)$ can be decomposed in terms of the Haar system as

$$f = \langle f \rangle_{(I_0)} \chi_{(I_0)} + \sum_{I \in \mathcal{D}} (f, h_I) h_I.$$

In terms of the expansion in the Haar system we define the martingale transform, g of f, as

$$g := \langle g \rangle_{(I_0)} \chi_{(I_0)} + \sum_{I \in \mathcal{D}} \varepsilon_I (f, h_I) h_I,$$

where $\varepsilon_I \in \{\pm 1\}$. Requiring that $|(g, h_J)| = |(f, h_J)|$ for all $J \in \mathcal{D}$ is equivalent to g being the martingale transform of f, for $f, g \in L^p(I_0)$.

Now we define the Bellman function as $\mathcal{B}(x_1, x_2, x_3) :=$

$$\sup_{f,g} \{ \langle (g^2 + \tau^2 f^2)^{\frac{p}{2}} \rangle_I : x_1 = \langle f \rangle_I, x_2 = \langle g \rangle_I, x_3 = \langle |f|^p \rangle_I, |(f, h_J)| = |(g, h_J)|, \ \forall J \in \mathcal{D} \}$$

on the domain $\Omega = \{x \in \mathbb{R}^3 : x_3 \geq 0, |x_1|^p \leq x_3\}$. The Bellman function is defined in this way, since we would like to know the value of the supremum of $\left\| \begin{pmatrix} g \\ \tau f \end{pmatrix} \right\|_p$, where g is the martingale transform of f. Note that $|x_1|^p \leq x_3$ is just Hölder's inequality. Even though the Bellman function is only being defined for real-valued functions, we can vectorize it to work for complex-valued (and even Hilbert-valued) functions. Slightly abusing the language, we can call $\langle (|g|^2 + \tau^2 |f|^2)^{\frac{p}{2}} \rangle_I^{\frac{1}{p}}$ a "quadratic perturbation" of the martingale transform's norm $\langle |g|^p \rangle_I^{\frac{1}{p}}$.

Theorem 1. *Let $\{d_k\}_{k \geq 1}$ be a complex martingale difference in $L^p[0,1]$, where $1 < p < \infty$, and $\{\varepsilon_k\}_{k \geq 1}$ a sequence in $\{\pm 1\}$. If $\tau \in [-\frac{1}{2}, \frac{1}{2}]$ and $n \in \mathbb{Z}_+$ then*

$$\left\| \sum_{k=1}^{n} \begin{pmatrix} \varepsilon_k \\ \tau \end{pmatrix} d_k \right\|_{L^p([0,1], \mathbb{C}^2)} \leq ((p^* - 1)^2 + \tau^2)^{\frac{1}{2}} \left\| \sum_{k=1}^{n} d_k \right\|_{L^p([0,1], \mathbb{C})},$$

where $((p^* - 1)^2 + \tau^2)^{\frac{p}{2}}$ is sharp.

Note that when $\tau = 0$ we get Burkholder's famous result [7].

Now that we have the problem formalized, notice that $\mathcal{B}$ is independent of the initial choice of I_0 (or I as we will denote it from here on) and $\{\alpha_{n,m}\}_{n,m}$, so we return to having them arbitrary. The use of generalized dyadics will come into play in Lemma 4. Finding $\mathcal{B}$ when $p = 2$ is easy, so we will do this first.

Proposition 2. *If $p = 2$ then $\mathcal{B}(x) = x_2^2 - x_1^2 + (1 + \tau^2) x_3$.*

Proof. Since $f \in L^2(I)$ then $f = \langle f\rangle_I \chi_I + \sum_{J\in\mathcal{D}}(f,h_J)h_J$ implies

$$\begin{aligned}\langle |f|^2\rangle_I &= \frac{1}{|I|}\int_I |f|^2\\ &= \langle f\rangle_I^2 + 2\langle f\rangle_I \sum_{J\in\mathcal{D}}(f,h_J)\frac{1}{|I|}\int_I h_J + \frac{1}{|I|}\int_I \sum_{J,K\in\mathcal{D}}(f,h_J)(f,h_K)h_J h_K\\ &= \langle f\rangle_I^2 + \frac{1}{|I|}\sum_{J\in\mathcal{D}}|(f,h_J)|^2.\end{aligned}$$

So $\|f\|_2^2 = |I|x_3 = |I|x_1^2 + \sum_{J\in\mathcal{D}}|(f,h_J)|^2$ and similarly

$$\|g\|_2^2 = |I|x_2^2 + \sum_{J\in\mathcal{D}}|(g,h_J)|^2 = |I|x_2^2 + \sum_{J\in\mathcal{D}}|(f,h_J)|^2.$$

Now we can compute $\mathcal{B}$ explicitly, (when $p=2$)

$$\begin{aligned}\langle (g^2+\tau^2 f^2)^{\frac{p}{2}}\rangle_I &= \langle |g|^2\rangle_I + \tau^2\langle |f|^2\rangle_I = x_2^2+\tau^2 x_1^2 + (1+\tau^2)\frac{1}{|I|}\sum_{J\in\mathcal{D}}|(f,h_J)|^2\\ &= x_2^2+\tau^2x_1^2+(1+\tau^2)(x_3-x_1^2). \quad \square\end{aligned}$$

1.2. **Outline of Argument to Prove Main Result.** Computing the Bellman function, $\mathcal{B}$, for $p \neq 2$, is much more difficult, so more machinery is needed. In Section 1.3 we will derive properties of the Bellman function, the most notable of which is concavity under certain conditions. Finding a $\mathcal{B}$ to satisfy the concavity will amount to solving a partial differential equation, after adding an assumption. This PDE has a solution on characteristics that is well known, so we just need to find an explicit solution from this, using the Bellman function properties. How the characteristics behave in the domain of definition for the Bellman function will give us several cases to consider. In Section 2 we get the Bellman function for $1 < p < \infty$ by putting together several cases. Once we have what we think is the Bellman function, we need to show that it has the necessary smoothness and that Assumption 7 was not too restrictive to give us the Bellman function. This is covered in Section 3. Finally the main result is shown in Section 4.

1.3. **Properties of the Bellman function.** One of the properties we nearly always have (or impose) for any Bellman function, is concavity (or convexity). It is not true that $\mathcal{B}$ is globally concave, on all of Ω, but under certain conditions it is concave. The needed condition is that g is the martingale transform of f, or $|x_1^+ - x_1^-| = |x_2^+ - x_2^-|$ in terms of the variables in Ω.

Definition 3. We say that a function B is restrictively concave if $x^{\pm} \in \Omega$ such that $x = \alpha^+ x^+ + \alpha^- x^-, \alpha^+ + \alpha^- = 1$ and if $|x_1^+ - x_1^-| = |x_2^+ - x_2^-|$ implies $\mathcal{B}(x) \geq \alpha^+ \mathcal{B}(x^+) + \alpha^- \mathcal{B}(x^-)$.

Proposition 4. *$\mathcal{B}$ is restrictively concave in the $x-$variables.*

Proof. Let $\varepsilon > 0$ be given and $x^{\pm} \in \Omega$. By the definition of $\mathcal{B}$, there exists $f^{\pm}, g^{\pm}$ on $I^{\pm}$ such that $\langle f \rangle_{I^{\pm}} = x_1^{\pm}, \langle g \rangle_{I^{\pm}} = x_2^{\pm}, \langle |f^{\pm}|^p \rangle_{I^{\pm}} = x_3^{\pm}$ and

$$\mathcal{B}(x^{\pm}) - \langle [(g^{\pm})^2 + \tau^2 (f^{\pm})^2]^{\frac{p}{2}} \rangle_{I^{\pm}} \leq \varepsilon$$

On $I = I^+ \cup I^-$ we define f and g as $f := f^+ \chi_{I^+} + f^- \chi_{I^-}, g := g^+ \chi_{I^+} + g^- \chi_{I^-}$. So,

$$\begin{aligned} |x_1^+ - x_1^-| &= |\langle f \rangle_{I^+} - \langle f \rangle_{I^-}| = \left| \frac{1}{|I^+|} \int_{|I^-|} f - \frac{1}{|I^-|} \int_{I^-} f \right| \\ &= \left| \frac{1}{\alpha^+ |I|} \int_{|I^-|} f - \frac{1}{\alpha^- |I|} \int_{I^-} f \right| = \frac{1}{|I|} \left| \int f \left(\frac{1}{\alpha^+} \chi_{I^+} - \frac{1}{\alpha^-} \chi_{I^-} \right) \right| \\ &= \sqrt{\frac{|I|}{\alpha^+ \alpha^-}} \left| \int f h_I \right| =: \sqrt{\frac{|I|}{\alpha^+ \alpha^-}} \left| (f, h_I) \right|. \end{aligned}$$

Similarly,$|x_2^+ - x_2^-| = \sqrt{\frac{|I|}{\alpha^+ \alpha^-}} \left| (g, h_I) \right|$. So our assumption $|x_1^+ - x_1^-| = |x_2^+ - x_2^-|$ is equivalent to $|(f, h_I)| = |(g, h_I)|$. Since $x_1 = \langle f \rangle_I, x_2 = \langle g \rangle_I$ and $x_3 = \langle |f|^p \rangle_I$ then f and g are test functions and so

$$\begin{aligned} \mathcal{B}(x) &\geq \langle (g^2 + \tau^2 f^2)^{\frac{p}{2}} \rangle_I \\ &= \alpha^+ \langle [(g^+)^2 + \tau^2 (f^+)^2]^{\frac{p}{2}} \rangle_{I^+} + \alpha^- \langle [(g^-)^2 + \tau^2 (f^-)^2]^{\frac{p}{2}} \rangle_{I^-} \\ &\geq \alpha^+ \mathcal{B}(x^+) + \alpha^- \mathcal{B}(x^-) - \varepsilon. \quad \square \end{aligned}$$

At this point we do not quite have concavity of $\mathcal{B}$ on Ω since there is the restriction $|x_1^+ - x_1^-| = |x_2^+ - x_2^-|$ needed. To make this condition more manageable, we will make a change of coordinates. Let $y_1 := \frac{x_2 + x_1}{2}, y_2 := \frac{x_2 - x_1}{2}$ and $y_3 := x_3$. We will also change notation for the Bellman function and corresponding domain in the new variable y. Let $\mathcal{M}(y_1, y_2, y_3) := \mathcal{B}(x_1, x_2, x_3) = \mathcal{B}(y_1 - y_2, y_1 + y_2, y_3)$. Then the domain of definition for $\mathcal{M}$ will be $\Xi := \{y \in \mathbb{R}^3 : y_3 \geq 0, |y_1 - y_2|^p \leq y_3\}$.

If we consider $x^{\pm} \in \Omega$ such that $|x_1^+ - x_1^-| = |x_2^+ - x_2^-|$, then the corresponding points $y^{\pm} \in \Xi$ satisfy either $y_1^+ = y_1^-$ or $y_2^+ = y_2^-$. This implies that fixing y_1 as

$y_1^+ = y_1^-$ or y_2 as $y_2^+ = y_2^-$ will make $\mathcal{M}$ concave with respect to y_2, y_3 under fixed y_1 and with respect to y_1, y_3 under y_2 fixed.

Rather than using Proposition 4 to check the concavity of the Bellman function we can just check it in the following way, assuming $\mathcal{M}$ is C^2. Let $j \neq i \in \{1,2\}$ and fix y_i as $y_i^+ = y_i^-$. Then $\mathcal{M}$ as a function of y_j, y_3 is concave if

$$\begin{pmatrix} \mathcal{M}_{y_jy_j} & \mathcal{M}_{y_jy_3} \\ \mathcal{M}_{y_3y_j} & \mathcal{M}_{y_3y_3} \end{pmatrix} \leq 0,$$

which is equivalent to

$$\mathcal{M}_{y_jy_j} \leq 0, \mathcal{M}_{y_3y_3} \leq 0, D_j = \mathcal{M}_{y_jy_j}\mathcal{M}_{y_3y_3} - \mathcal{M}_{y_3y_j}\mathcal{M}_{y_jy_3} \geq 0.$$

Proposition 5. *(Restrictive Concavity in $y-$variables) Let $j \neq i \in \{1,2\}$ and fix y_i as $y_i^+ = y_i^-$. If $\mathcal{M}_{y_jy_j} \leq 0, \mathcal{M}_{y_3y_3} \leq 0$ and $D_j = \mathcal{M}_{y_jy_j}\mathcal{M}_{y_3y_3} - (\mathcal{M}_{y_jy_3})^2 \geq 0$ for $j = 1$ and $j = 2$ then $\mathcal{M}$ is restrictively concave.*

The Bellman function, as it turns out, has many other nice properties.

Proposition 6. *Suppose that $\mathcal{M}$ is $C^1(\mathbb{R}^3)$, then $\mathcal{M}$ has the following properties.*

(i) Symmetry: $\mathcal{M}(y_1, y_2, y_3) = \mathcal{M}(y_2, y_1, y_3) = \mathcal{M}(-y_1, -y_2, y_3)$

(ii) Dirichlet boundary data: $\mathcal{M}(y_1, y_2, (y_1 - y_2)^p) = ((y_1 + y_2)^2 + \tau^2(y_1 - y_2)^2)^{\frac{p}{2}}$

(iii) Neumann conditions: $\mathcal{M}_{y_1} = \mathcal{M}_{y_2}$ on $y_1 = y_2$ and $\mathcal{M}_{y_1} = -\mathcal{M}_{y_2}$ on $y_1 = -y_2$

(iv) Homogeneity: $\mathcal{M}(ry_1, ry_2, r^py_3) = r^p\mathcal{M}(y_1, y_2, y_3), \forall r > 0$

(v) Homogeneity relation: $y_1\mathcal{M}_{y_1} + y_2\mathcal{M}_{y_2} + py_3\mathcal{M}_{y_3} = p\mathcal{M}$

Proof. (i) Note that we get $\mathcal{B}(x_1, x_2, x_3) = \mathcal{B}(-x_1, x_2, x_3) = \mathcal{B}(x_1, -x_2, x_3)$ by considering test functions $\widetilde{f} = -f$ and $\widetilde{g} = -g$. Change coordinates from x to y and the result follows.

(ii) On the boundary $\{x_3 = |x_1|^p\}$ of Ω we see that

$$\frac{1}{|I|}\int_I |f|^p = \langle |f|^p \rangle_I = x_3 = |x_1|^p = |\langle f \rangle_I|^p = \left|\frac{1}{|I|}\int_I f\right|^p$$

is only possible if $f \equiv$ const. (i.e. $f = x_1$). But, $|(f, h_J)| = |(g, h_J)|$ for all $J \in \mathcal{I}$, which implies that $g \equiv$ const. (i.e. $g = x_2$). Then $\mathcal{B}(x_1, x_2, |x_1|^p) = \langle (g^2 + \tau^2 f^2)^{\frac{p}{2}} \rangle_I = (x_2^2 + \tau^2 x_1^2)^{\frac{p}{2}}$. Changing coordinates gives the result.

(iii) This follows from from (i).

(iv) Consider the test functions $\widetilde{f} = rf, \widetilde{g} = rg$

(v) Differentiate (iv) with respect to r and evaluate it at $r = 1$. ☐

Now that we have all of the properties of the Bellman function we will turn our attention to finding it. Proposition 5 gives us a partial differential inequality to solve, which can be quite difficult. We can get a PDE instead to work with, by assuming that

$$\begin{pmatrix} \mathcal{M}_{y_jy_j} & \mathcal{M}_{y_jy_3} \\ \mathcal{M}_{y_3y_j} & \mathcal{M}_{y_3y_3} \end{pmatrix}$$

is degenerate (i.e. $D_j = \mathcal{M}_{y_jy_j}\mathcal{M}_{y_3y_3} - (\mathcal{M}_{y_3y_j})^2 = 0$). The PDE that we now have is the well known Monge–Ampère equation which has a solution. Let us make it clear that we have added an assumption.

Assumption 7. *If we fix y_i, then*

$$\begin{pmatrix} \mathcal{M}_{y_jy_j} & \mathcal{M}_{y_jy_3} \\ \mathcal{M}_{y_3y_j} & \mathcal{M}_{y_3y_3} \end{pmatrix}$$

is degenerate, where $i \neq j \in \{1, 2\}$.

Adding this assumption comes with a price. Any function that we construct satisfying all properties of the Bellman function (we call such functions Bellman function candidates), must still be shown to equal the Bellman function.

Proposition 8. *For $j = 1$ or 2, $\mathcal{M}_{y_jy_j}\mathcal{M}_{y_3y_3} - (\mathcal{M}_{y_3y_j})^2 = 0$ has the solution $M(y) = y_jt_j + y_3t_3 + t_0$ on the characteristics $y_jdt_j + y_3dt_3 + dt_0 = 0$, which are straight lines in the $y_j \times y_3$ plane. Furthermore, t_0, t_j, t_3 are constant on characteristics with the property $M_{y_j} = t_j, M_{y_3} = t_3$.*

This is a result of Pogorelov, see [18], [19]. Now that we have a solution M to the Monge–Ampère, we need get rid of t_0, t_j, t_3 so that we have an explicit form of M, without the characteristics. We note that a solution to the Monge–Ampère is not necessarily the Bellman function. It must satisfy the restrictive concavity of Proposition 5, be C^1-smooth, and satisfy the properties of Proposition 6. The restrictive concavity property is one of the key deciding factors of whether or not we have a Bellman function candidate in many cases. Even if the Monge–Ampère solution satisfies all of those conditions, it must still be shown to be equal to the Bellman function, because we added an additional assumption (Assumption 7) to get the Monge–Ampère solution. This will be considered rigorously in Section 3, after we obtain a solution to the Monge–Ampère equation, with the appropriate Bellman function properties (our candidate). So from this point on we will use M

and B to denote solutions to the Monge–Ampère equation, i.e. Bellman function candidates, and $\mathcal{M}$ and $\mathcal{B}$ to denote the true Bellman function.

2. **Finding the Monge–Ampère solution which is the Bellman function candidate**

In this section we impose some of the Bellman function properties on the Monge–Ampère solution to get a Bellman function candidate. We omit the details only giving a rough outline for how to do this. For a detailed long computation of the Bellman function, see [1].

Due to the symmetry property of $\mathcal{M}$, from Proposition 6, we only need to consider the domain $\Xi_+ := \{y : -y_1 \leq y_2 \leq y_1, y_3 \geq 0, (y_1 - y_2)^p \leq y_3\}$ rather than Ξ. Since the characteristics are straight lines, then (for at least part of the domain) one end of each line must be on the boundary $\{y : (y_1 - y_2)^p = y_3\}$. Let U denote the point at which the characteristic touches the boundary. Furthermore, there are only four possibilities for the behavior of the characteristics in the plane.

(1) The characteristic goes from U to $\{y : y_1 = y_2\}$
(2) The characteristic goes from U to to infinity, running parallel to the y_3-axis
(3) The characteristic goes from U to $\{y : y_1 = -y_2\}$
(4) The characteristic goes from U to $\{y : (y_1 - y_2)^p = y_3\}$.

Note that these cases may not entirely describe how the characteristics fill the domain. For example if one is able to find a Bellman function candidate in Case (4) with y_1 fixed, then there must be another set of characteristics to fill the remaining part of the domain, which may not be one of the remaining cases listed here (for example take horizontal lines in Figure 1). So these cases should be thought of as a starting point. However, it turns out that Cases (1), (2) and (3) are enough to build the Bellman function for $\tau-$values small enough.

To find a Bellman function candidate we must first fix a variable (y_1 or y_2) and a case for the characteristics. Then we use the Bellman function properties to get rid of the characteristics. If the Monge–Ampère solution satisfies restrictive concavity, then it is a Bellman function candidate. Checking the restrictive concavity is quite difficult in many of the cases, since it amounts to doing second derivative estimates for an implicitly defined function. Let us now find our Bellman function candidate.

First of all, fix y_1. Then the Monge–Ampère solution from Case (1) is only valid on part of the domain Ξ_+. Furthermore, restrictive concavity is only valid for $2 < p < \infty$. Turning to Case (2) we obtain similar results. Restrictive concavity is only valid on part of the domain Ξ_+ for Case (2). But, as luck would have it, the two partial Bellman candidates are the missing halves of one another. So we can glue together the partial solutions from Cases (1) and (2) to get a whole solution in Ξ_+. The characteristics for this solution can be seen in Figure 1.

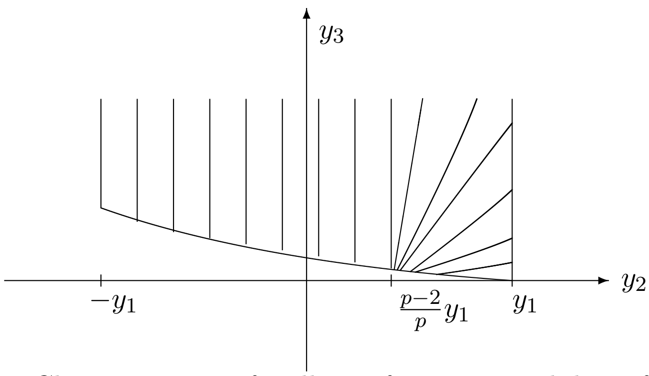

FIGURE 1. Characteristics of Bellman function candidate, for $2 < p < \infty$.

Proposition 9. *For $2 < p < \infty, |\tau| \le 1$ and $\gamma = \frac{1-\tau^2}{1+\tau^2}$, the Monge–Ampère solution given by the following is a Bellman function candidate.*

$$M(y) = (1+\tau^2)^{\frac{p}{2}}[y_1^2 + 2\gamma y_1 y_2 + y_2^2]^{\frac{p}{2}} + ((p-1)^2+\tau^2)^{\frac{p}{2}}[y_3 - (y_1-y_2)^p]$$

when $-y_1 < y_2 \le \frac{p-2}{p}y_1$ and is given implicitly by

$G(y_1+y_2, y_1-y_2) = y_3 G(\sqrt{\omega^2-\tau^2}, 1)$ when $\frac{p-2}{p}y_1 \le y_2 < y_1$, where $G(z_1,z_2) = (z_1+z_2)^{p-1}[z_1-(p-1)z_2]$ and $\omega = \left(\frac{M(y)}{y_3}\right)^{\frac{1}{p}}$. This solution satisfies all properties of the Bellman function, including restrictive concavity.

Keeping y_1 fixed we turn our attention to Case (3). The Monge–Ampère solution from Case (3) only gives a solution on part of the domain Ξ_+. Moreover, restrictive concavity is only valid for $1 < p < 2$. Just as for the dual $p-$values, Case (2) only provides a partial solution (due to restrictive concavity only being valid on part of the domain). So we can glue the two partial solutions together to get a Bellman function candidate for $1 < p < 2$ and $|\tau| \le \frac{1}{2}$. The $\tau-$values had to be restricted slightly more than for $2 < p < \infty$.

Proposition 10. *Let $1 < p < 2$. If $|\tau| \le \frac{1}{2}$ and $\gamma = \frac{1-\tau^2}{1+\tau^2}$, then the Monge–Ampère solution given by the following is a Bellman function candidate.*

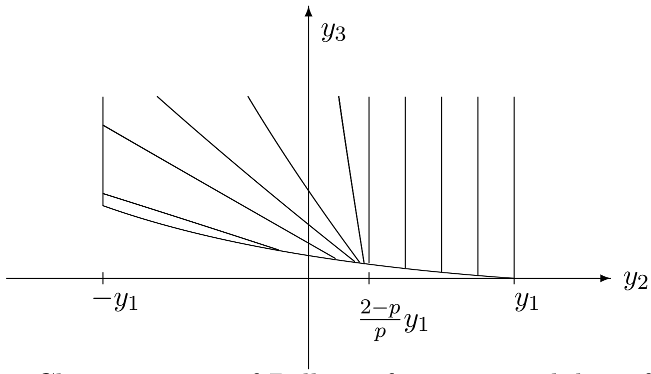

FIGURE 2. Characteristics of Bellman function candidate, for $1 < p < 2$.

$M(y) = (1+\tau^2)^{\frac{p}{2}}[y_1^2 + 2\gamma y_1 y_2 + y_2^2]^{\frac{p}{2}} + \left(\frac{1}{(p-1)^2} + \tau^2\right)^{\frac{p}{2}}[y_3 - (y_1 - y_2)^p]$ when $\frac{2-p}{p} y_1 \le y_2 < y_1$ and is given implicitly by

$G(y_1 - y_2, y_1 + y_2) = y_3 G(1, \sqrt{\omega^2 - \tau^2})$ when $-y_1 < y_2 \le \frac{2-p}{p} y_1$, where $G(z_1, z_2) = (z_1 + z_2)^{p-1}[z_1 - (p-1)z_2]$ and $\omega = \left(\frac{M(y)}{y_3}\right)^{\frac{1}{p}}$. This function satisfies all of the properties of the Bellman function.

Most of the remaining cases do not yield a Bellman function candidate. If we fix y_2 then the Monge–Ampère solution from Cases (1) and (3) do not satisfy the restrictive concavity needed to be a Bellman function candidate. Propositions 9 and 10 are direct computations of these facts, see details in [1]. Case (2) yields the same partial solution, if we first fix y_1 or y_2, since restrictive concavity is only valid on part of the domain. All that remains is Case (4). However, we do not know whether or not Case (4) gives a Bellman function candidate. For $\tau = 0$, it was shown in [21] that Case (4) does not produce a Bellman function candidate, since otherwise some simple extremal functions give a contradiction to linearity of the Monge–Ampère solution on characteristics. However, for $\tau \neq 0$ it is much more difficult to show this. Case (4) could give a solution throughout Ξ_+ or could yield a partial solution that would work well with the characteristics from Case (2_1). But, we believe that Case (4) should not give us a Bellman candidate and actually it does not matter now, we will proceed in showing that our Bellman function candidate is actually the Bellman function now. This would have to be checked anyway and we got lucky that this is true without messing with Case (4).

3. The Monge–Ampère solution is the Bellman function

We will now show that the Monge–Ampère solution obtained in Proposition 9 and 10 is actually the Bellman function. To this end, let us revert back to the $x-$variables. We will denote the Bellman function candidate as B_τ and use $\mathcal{B}_\tau$ to denote the true Bellman function. Extending the function G to U_τ makes it possible to define the solution in terms of a single relation.

Definition 11. Denote $v(x,y) := v_{p,\tau}(x,y) = (\tau^2|x|^2+|y|^2)^{\frac{p}{2}} - ((p^*-1)^2+\tau^2)^{\frac{p}{2}}|x|^p$, $u(x,y) := u_{p,\tau}(x,y) = p(1-\frac{1}{p*})^{p-1}\left(1+\frac{\tau^2}{(p^*-1)^2}\right)^{\frac{p-2}{2}}(|x|+|y|)^{p-1}[|y|-(p^*-1)|x|]$ and

$$U(x,y) := U_{p,\tau}(x,y) = \begin{cases} v(x,y) & : |y| \geq (p^*-1)|x| \\ u(x,y) & : |y| \leq (p^*-1)|x| \end{cases}$$

for $1 < p < 2$. The two pieces of U are interchanged for $2 \leq p < \infty$.

Proposition 12. *For $1 < p < 2$ and $|\tau| \leq \frac{1}{2}$ or $2 \leq p < \infty$ and $|\tau| \leq 1$ the Bellman function candidate is the unique positive solution given by*

$$U(x_1,x_2) = U\left(x_3^{\frac{1}{p}}, \sqrt{B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}}\right).$$

Moreover, U is C^1-smooth on Ω.

Proof. First consider $2 < p < \infty$. It is clear that

$$U(x_1,x_2) = U\left(x_3^{\frac{1}{p}}, \sqrt{B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}}\right), \tag{3.1}$$

by comparing the solution obtained in Proposition 9 and using the symmetry property in Proposition 6. The constant $\alpha_{p,\tau} = p(1-\frac{1}{p^*})^{p-1}\left(1+\frac{\tau^2}{(p^*-1)^2}\right)^{\frac{p-2}{2}}$ was determined to make $U_x = U_y$ at $|y| = (p^*-1)|x|$. The partial derivatives are given by,

$$\begin{aligned}
u_x &= \alpha_{p,\tau}(p-1)x'(|x|+|y|)^{p-2}(|y|-(p^*-1)|x|) - \alpha_{p,\tau}(p^*-1)x'(|x|+|y|)^{p-1}, \\
v_x &= p\tau^2 x(\tau^2|x|^2+|y|^2)^{\frac{p-2}{2}} - px'((p^*-1)^2+\tau^2)^{\frac{p}{2}}|x|^{p-1}, \\
u_y &= \alpha_{p,\tau}(p-1)y'(|x|+|y|)^{p-2}(|y|-(p^*-1)|x|) + \alpha_{p,\tau}y'(|x|+|y|)^{p-1}, \\
v_y &= py(\tau^2|x|^2+|y|^2)^{\frac{p-2}{2}},
\end{aligned}$$

where $x' = \frac{x}{|x|}$ and $y' = \frac{y}{|y|}$. U is C^1-smooth except possibly at gluing and symmetry lines. It is easy to verify that u_x is continuous at $\{x = 0\}$, U_x and U_y are continuous at $\{|y| = (p^* - 1)|x|\}$ and v_y is continuous at $\{y = 0\}$. This proves that U is C^1-smooth on Ω.

Observe that $U_y > 0$ for $y \neq 0$ and $U_x < 0$ for $x \neq 0$. This is enough to show that B_τ is the unique positive solution to (3.1). Indeed, if $x \in \Omega$ such that $|x_1| = x_3^{\frac{1}{p}}$, then $\sqrt{B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}} = |x_2|$ by the Dirichlet boundary conditions. This gives us (3.1) uniquely at $B_\tau(x)$. Fix x_1, such that $|x_1| < x_3^{\frac{1}{p}}$, then $U\left(x_3^{\frac{1}{p}}, \sqrt{B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}}\right) < U\left(x_1, \sqrt{B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}}\right)$. Since x_1 is fixed, then $\sqrt{B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}} > |x_2|$, so $U\left(x_1, \sqrt{B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}}\right)$ strictly decreases to $U(x_1, x_2)$, as $\sqrt{B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}}$ decreases to $|x_2|$, giving us a unique $B_\tau(x)$ for which (3.1) holds.

Now consider $1 < p < 2$. U is C^1-smooth on Ω, since v_x is continuous at $\{x = 0\}$, u_y is continuous at $\{y = 0\}$ and U_x and U_y are continuous at $\{|y| = (p^* - 1)|x|\}$. This is easily verified since the partial derivatives are computed above (just switch the two pieces of each function). Observe that for $x \neq 0$ and $y \neq 0, U_x < 0$ and for $y \neq 0, U_y > 0$. Then the argument above showing $U(x_1, x_2) = U\left(x_3^{\frac{1}{p}}, \sqrt{B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}}\right)$ uniquely determines B_τ also holds for this range of $p-$values as well, except maybe at $x_1 = x_2 = 0$. Suppose $U(0, 0) = U\left(x_3^{\frac{1}{p}}, \sqrt{B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}}\right)$, then $B_\tau(x) = ((p^* - 1)^2 + \tau^2)^{\frac{p}{2}} x_3$. So $B_\tau(x)$ is uniquely determined by the fixed $x-$value. □

Corollary 13. *B_τ is continuous in Ω.*

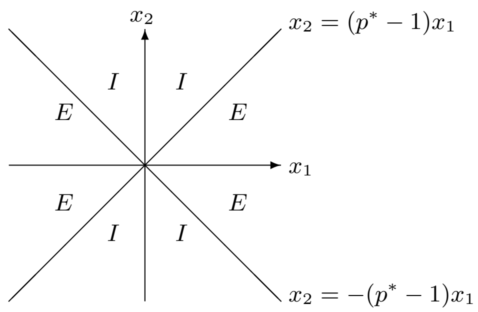

FIGURE 3. Location of Implicit (I) and Explicit (E) part of B_τ for $2 \leq p < \infty$.

Proof. We only consider $2 < p < \infty$ as the dual range is handled identically. By Proposition 12, we have that B_τ is the unique positive solution to 3.1. Since this is true for all $|\tau| \le 1$, then $B_0 = \left(B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}\right)^{\frac{p}{2}}$ on $|x_2| \ge (p^* - 1)|x_1|$, since $U_{p,\tau} = \left(1 + \frac{\tau^2}{(p^*-1)^2}\right)^{\frac{p-2}{2}} U_{p,0}$. Equivalently, we have

$$B_\tau = \left(B_0^{\frac{2}{p}} + \tau^2 x_3^{\frac{2}{p}}\right)^{\frac{p}{2}}. \tag{3.2}$$

Since B_0 was shown to be continuous in [21] (pg. 26) then B_τ is also continuous on $|x_2| \ge (p^* - 1)|x_1|$, using the relation. This takes care of the implicit part of B_τ. The explicit part of B_τ is clearly continuous on $|x_2| \le (p^* - 1)|x_1|$. □

Lemma 14. *Let $1 < p < \infty$. Then, $B_\tau\big|_L$ is C^1-smooth on Ω, where L is any line in Ω.*

Proof. Since $B_\tau\big|_L$ is C^2-smooth on Ω_+, all that remains to be checked is the smoothness at the gluing and symmetry lines, i.e. at $\{x_1 = 0\}, \{x_2 = 0\}$ and $\{|x_2| = (p^* - 1)|x_1|\}$. Let $L = L(t), t \in \mathbb{R}$, be any line in Ω passing through any of the planes in question, such that $L(0)$ is on the plane. Now plug $L(t)$ into (3.1) and differentiate with respect to t. Let $t \to 0^+$ and $t \to 0^-$ and equate the two relations. This gives

$$\frac{d}{dt} B_\tau(L(t))\big|_{t=0^-} = \frac{d}{dt} B_\tau(L(t))\big|_{t=0^+}. \quad \square$$

Proposition 15. *Let $1 < p < 2$ and $|\tau| \le \frac{1}{2}$ or $2 \le p < \infty$ and $|\tau| \le 1$. Then B_τ is restrictively concave.*

Proof. Recall that Propositions 9 and 10, together with the symmetry property of B_τ, establish this result everywhere, except at $\{x_1 = 0\}, \{x_2 = 0\}$ and $\{|x_2| = (p^* - 1)|x_1|\}$. Let $f(t) = B_\tau\big|_{L(t)}$, where L is any line in Ω, such that $L(0) \in \{x_1 = 0\}, \{x_2 = 0\}$ or $\{|x_2| = (p^* - 1)|x_1|\}$. Since $f'' < 0$ for $t < 0$ and $t > 0$ and f is C^1-smooth (by Lemma 14), then f is concave. □

Proposition 16. *Let $1 < p < \infty$. If a function $\widetilde{B}$ on Ω has restrictive concavity and $\widetilde{B}_\tau(x_1, x_2, |x_1|^p) \ge (\tau^2 x_1^2 + x_2^2)^{\frac{p}{2}}$, then $\widetilde{B}_\tau \ge \mathcal{B}_\tau$. In particular, $B_\tau \ge \mathcal{B}_\tau$.*

Proof. This was proven in [21] for B_0 (Lemma 2 on page 29). The same proof will apply here to B_τ. □

Proposition 17. *For $1 < p < \infty, B_\tau \le \mathcal{B}_\tau$.*

Proof. For $1 < p < 2$ there is a direct proof, which will be discussed first. By (3.2) we know that $B_0 = \left(B_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}\right)^{\frac{p}{2}}$ on $\{|x_2| \le (p^* - 1)|x_1|\}$. Consider, $\widetilde{\mathcal{B}}_0 = \left(\mathcal{B}_\tau^{\frac{2}{p}} - \tau^2 x_3^{\frac{2}{p}}\right)^{\frac{p}{2}}$. It suffices to show that $B_0 \le \widetilde{\mathcal{B}}_0$. But, Burkholder showed that $B_0 = \mathcal{B}_0$, so without the supremum's we can reduce to simply showing

$$\langle |g|^p\rangle_I^{\frac{2}{p}} + \tau^2 \langle |f|^p\rangle_I^{\frac{2}{p}} \le \langle (\tau^2|f|^2 + |g|^2)^{\frac{p}{2}}\rangle_I.$$

Apply Minkowski's inequality: $\left\|\int_I (A,B)\right\|_{l^{\frac{2}{p}}} \le \int_I \|(A,B)\|_{l^{\frac{2}{p}}}$. Choosing $A = |g|^p$ and $|\tau f|^p$ proves the result. So we have shown that $B_\tau \le \mathcal{B}_\tau$ on $\{|x_2| \le (p^*-1)|x_1|\}$.

Now we would like to show that $B_\tau \le \mathcal{B}_\tau$ on $\{|x_2| \ge (p^*-1)|x_1|\}$. Let $H_1(x_1,x_2,x_3) = B_\tau(x_1,x_2,x_3) - B_\tau(0,0,1)x_3$. Lemma 21, in the next section, proves that $H_1(x_1,x_2,\cdot)$ is an increasing function starting at $H_1(x_1,x_2,|x_1|^p) = v_\tau(x_1,x_2)$ and increasing to $\widetilde{U}_{p,\tau}(x,y) := \sup_{t\ge|x|^p}\{B_\tau(x,y,t) - B_\tau(0,0,1)t\}$. The same proof works for $H_2(x_1,x_2,x_3) = \mathcal{B}_\tau(x_1,x_2,x_3) - \mathcal{B}_\tau(0,0,1)x_3$. So

$$H_2(x_1,x_2,x_3) \ge v_\tau(x_1,x_2) = B_\tau(x_1,x_2,x_3) - B_\tau(0,0,1)x_3.$$

Since $B_\tau(0,0,1) \le \mathcal{B}_\tau(0,0,1)$, then $B_\tau \le \mathcal{B}_\tau$ on $\{|x_2| \ge (p^*-1)|x_1|\}$.

Now we consider $2 < p < \infty$. Let $\varepsilon > 0$ be arbitrarily small and consider the following extremal functions

$$f(x) = \begin{cases} -c & : 1 < x < \varepsilon \\ \gamma f\left(\frac{t-\varepsilon}{1-2\varepsilon}\right) & : \varepsilon < x < 1-\varepsilon \\ c & : 1-\varepsilon < x < 1, \end{cases}$$

$$g(x) = \begin{cases} d_- & : 1 < x < \varepsilon \\ \gamma g\left(\frac{t-\varepsilon}{1-2\varepsilon}\right) & : \varepsilon < x < 1-\varepsilon \\ d_+ & : 1-\varepsilon < x < 1, \end{cases}$$

where $c, d_\pm$ and γ are defined so that f and g are a pair of test functions at $(0,x_2,x_3)$. We can use f and g to show, just as in [21] (Lemma 3, pg. 30), that $B_\tau(0,x_2,x_3) \le \mathcal{B}_\tau(0,x_2,x_3)$.

Now we need to take care of the estimate when $x_1 \ne 0$. Making a change of coordinates from x to y we only need to consider $y \in \Xi_+$, by the symmetry property of the Bellman function and Bellman function candidate. So $M_\tau(y_1,y_1,y_3) \le \mathcal{M}_\tau(y_1,y_1,y_3)$. The Dirichlet boundary conditions give that $M(y_1,y_2,(y_1-y_2)^p) = \mathcal{M}(y_1,y_2,(y_1-y_2)^p)$. On any characteristic in $\{\frac{p-2}{p}y_1 \le y_2 \le y_1\}$, see Figure 1, M_τ is linear (since it is the Monge–Ampère solution) and $\mathcal{M}_\tau$ is concave (by Proposition

4). Therefore, $M_\tau(y_1, y_2, y_3) \le \mathcal{M}_\tau(y_1, y_2, y_3)$ on $\{\frac{p-2}{p} y_1 \le y_2 \le y_1\}$. On $\{\frac{p-2}{p} y_1 \le y_2 \le y_1\}$, we can use the same proof as for $1 < p < 2$, to get $M_\tau(y_1, y_2, y_3) \le \mathcal{M}_\tau(y_1, y_2, y_3)$ on $\{-y_1 \le y_2 \le \frac{p-2}{p} y_1\}$. □

Now that we have shown $B = \mathcal{B}$ we will derive another surprising relationship.

Definition 18. We define $\mathcal{B}^l = \mathcal{B}^l(x_1, x_2, x_3)$ as the least restrictively concave majorant of $(x_2^2 + \tau^2 x_1^2)^{\frac{p}{2}}$ in Ω.

Proposition 19. *For* $1 < p < \infty, B = \mathcal{B} = \mathcal{B}^l$.

Proof. First we will show $\mathcal{B} \ge \mathcal{B}^l$. By Minkowski's inequality,

$$\begin{aligned} B(x_1, x_2, x_3) &= \mathcal{B}(x_1, x_2, x_3) \ge \int (|g|^2 + \tau^2 |f|^2)^{\frac{p}{2}} = \int \left\| \begin{pmatrix} g \\ \tau f \end{pmatrix} \right\|_{l^2}^p \ge \left\| \int \begin{pmatrix} g \\ \tau f \end{pmatrix} \right\|_{l^2}^p \\ &= (|\langle g \rangle|^2 + \tau^2 |\langle f \rangle|^2)^{\frac{p}{2}} = (x_2^2 + \tau^2 x_1^2)^{\frac{p}{2}}, \end{aligned}$$

proving the estimate.

Conversely, one can show that $\mathcal{B} \le \mathcal{B}^l$ just as $\mathcal{B} \le B$ was shown in Proposition 16 (simply apply the same argument to $\mathcal{B}^l$ using the restrictive concavity of this function). □

4. **Proving the main result**

Now that we have the Bellman function, the main result can be proven without too much difficulty. But first, we will find another relationship between U and v. Quite surprisingly, U is the least zigzag-biconcave majorant of v.

Definition 20. We denote any function of (x, y) as zigzag-biconcave if it is biconcave in $(x + y, x - y)$.

Lemma 21. *Let* $1 < p < \infty$ *and* $\widetilde{U}_{p,\tau}(x, y) = \sup_{t \ge |x|^p} \{B_\tau(x, y, t) - B_\tau(0, 0, 1)t\}$. *Fix* (x, y). *The function* $H(x, y, t) = B_\tau(x, y, t) - B_\tau(0, 0, 1)t$ *is increasing in* t *from* $H(x, y, |x|^p) = v(x, y) := (\tau^2 x^2 + y^2)^{\frac{p}{2}} - ((p^* - 1)^2 + \tau^2)^{\frac{p}{2}} |x|^p$ *to* $\widetilde{U}_{p,\tau}(x, y)$.

Proof. Recall that B_τ is continuous in Ω and for (x, y) fixed, $B_\tau(x, y, \cdot)$ is concave. Then $H(x, y, \cdot)$ is also concave. Since $\widetilde{U}_{p,\tau}(x, y) = \sup_{t \ge |x|^p} \{B_\tau(x, y, t) - B_\tau(0, 0, 1)t\}$, then $H(x, y, \cdot)$ either increases to $\widetilde{U}_{p,\tau}(x, y)$, or there exists t_0 such that $H(x, y, t_0) = \widetilde{U}_{p,\tau}(x, y)$ and H is decreasing for $t > t_0$. If H is decreasing for

$t > t_0$, then $H \longrightarrow -\infty$ as $t \longrightarrow \infty$ by concavity. Then there exists $\varepsilon > 0$ and $t' > t_0$ such that $H(x,y,t') < \varepsilon t'$. So we have, $\limsup_{t\to\infty} \frac{H(x,y,t)}{t} < -\varepsilon$. But,

$$\lim_{t\to\infty} \frac{H(x,y,t)}{t} = \lim_{t\to\infty} \left[B_\tau \left(\frac{x}{t^{\frac{1}{p}}}, \frac{y}{t^{\frac{1}{p}}}, 1 \right) - B_\tau(0,0,1) \right] = 0,$$

by continuity of B_τ at $(0,0,1)$. This gives us a contradiction. Therefore, $H(x,y,t) \geq -\varepsilon t$, for all t and all $\varepsilon > 0$, i.e. H is non-negative concave function on $[|x|^p, \infty)$. So $H(x,y,\cdot)$ is increasing and $H(x,y,|x|^p) = v(x,y)$ by the Dirichlet boundary conditions of B_τ in Proposition 6. □

Proposition 22. *For* $1 < p < \infty, U_{p,\tau}(x,y) = \widetilde{U}_{p,\tau}(x,y)$.

Proof. Suppose $2 \leq p < \infty$ and $|y| \geq (p-1)|x|$. Then

$$\begin{aligned} \widetilde{U}_0(x,y) &= \lim_{t\to\infty} (B_0(x,y,t) - B_0(0,0,1)t) \\ &= \lim_{t\to\infty} \frac{B\left(\frac{x}{x^{\frac{1}{p}}}, \frac{y}{x^{\frac{1}{p}}}, 1 \right)}{1/t} \\ &= \frac{d}{du} B(u^{\frac{1}{p}}x, u^{\frac{1}{p}}y, 1) \Big|_{u=0}. \end{aligned}$$

Now we repeat the same steps and obtain

$$\begin{aligned} \widetilde{U}_\tau(x,y) &= \lim_{t\to\infty} (B_\tau(x,y,t) - B_\tau(0,0,1)t) \\ &= \frac{d}{du} \left[\left(B_0^{\frac{2}{p}}(u^{\frac{1}{p}}x, u^{\frac{1}{p}}y, 1) + \tau^2 \right)^{\frac{p}{2}} \right] \Big|_{u=0} \\ &= \left[\left(B_0^{\frac{2}{p}}(u^{\frac{1}{p}}x, u^{\frac{1}{p}}y, 1) + \tau^2 \right)^{\frac{p-2}{2}} B_0^{\frac{2-p}{p}}(u^{\frac{1}{p}}x, u^{\frac{1}{p}}y, 1) \frac{d}{du} B_0(u^{\frac{1}{p}}x, u^{\frac{1}{p}}y, 1) \right] \Big|_{u=0} \\ &= \left(1 + \frac{\tau^2}{(p-1)^2} \right)^{\frac{p-2}{2}} \widetilde{U}_0(x,y) \\ &= \left(1 + \frac{\tau^2}{(p-1)^2} \right)^{\frac{p-2}{2}} U_0(x,y), \end{aligned}$$

where the last equality is by [7]. Therefore, $\widetilde{U}_\tau(x,y) = U_\tau(x,y)$.

Now suppose $|y| \leq (p-1)|x|$. Looking at the explicit form of B_τ in the region, note that $B_\tau(x,y,\cdot)$ is linear. So $\widetilde{U}_\tau(x,y) = \sup_{t\geq|x|^p}\{B_\tau(x,y,t) - B_\tau(0,0,1)t\} = \sup_{t\geq|x|^p}\{B_\tau(x,y,0)\} = v_\tau(x,y) = U_\tau(x,y)$.

We can apply the same proof to show that $\widetilde{U}_\tau(x,y) = U_\tau(x,y)$ for $1 < p < 2$. □

Proposition 23. *For $\tau \in \left[-\frac{1}{2}, \frac{1}{2}\right]$ and $1 < p < 2$ or $\tau \in [-1,1]$ and $2 \le p < \infty$, U is the least zigzag-biconcave majorant of $v(x,y) = (y^2+\tau^2x^2)^{\frac{p}{2}} - ((p^*-1)^2+\tau^2)^{\frac{p}{2}}|x|^p$.*

Proof. Recall the following facts just proved in Lemma 21 and Proposition 22:

$$U(x,y) = \sup_{t:(x,y,t)\in\Omega} \{\mathcal{B}(x,y,t) - \mathcal{B}(0,0,1)t\} \ \ge v(x,y) = (y^2+\tau^2x^2)^{\frac{p}{2}} - C_pt,$$

where $C_p = ((p^*-1)^2+\tau^2)^{\frac{p}{2}}$.

Suppose that w is a zigzag-biconcave function such that $v \le w \le U$. Then $W(x,y,t) := w(x,y) + C_pt$ has restrictive concavity and $W(x,y,t) \ge v(x,y) + C_pt = (y^2+\tau^2x^2)^{\frac{p}{2}}$. Therefore, by Proposition 16, we have $W \ge \mathcal{B}$. So we have

$$\begin{aligned} w(x,y) &= \sup_{t:(x,y,t)\in\Omega} \{W(x,y,t) - C_pt\} \\ &\ge \sup_{t:(x,y,t)\in\Omega} \{\mathcal{B}(x,y,t) - C_pt\} = U(x,y) \quad \square \end{aligned}$$

We now have enough results to easily prove the main result, in terms of the Haar expansion of a $\mathbb{R}-$valued L^p function.

Theorem 24. *Let $1 < p < 2, |\tau| \le \frac{1}{2}$ or $2 \le p < \infty, |\tau| \le 1$. Denote $I = [0,1]$ and let $f, g : [0,1] \to \mathbb{R}$. If $|\langle g\rangle_I| \le (p^*-1)|\langle f\rangle_I|$ and $|(f,h_J)| = |(g,h_J)|$ for all $J \in \mathcal{D}$, then $\langle(\tau^2|f|^2+|g|^2)^{\frac{p}{2}}\rangle_I \le ((p^*-1)^2+\tau^2)^{\frac{p}{2}}\langle|f|^p\rangle_I$, where $((p^*-1)^2+\tau^2)$ is the sharp constant and $p^*-1 = \max\left\{p-1, \frac{1}{p-1}\right\}$.*

Proof. Suppose $2 \le p < \infty$ and $|\tau| \le 1$. The proof relies on the fact that by Proposition 23

$$U(x,y) = \sup_{t\ge|x|^p} \{B(x,y,t) - B(0,0,1)t\}.$$

and by Proposition 19, $B = \mathcal{B}$. Since $|y| \le (p^*-1)|x|$ on Ω, then

$$U(x,y) = v(x,y) = (|y|^2+\tau^2|x|^2)^{\frac{p}{2}} - ((p^*-1)^2+\tau^2)^{\frac{p}{2}}|x|^p \le 0.$$

Then,

$$\sup_{\substack{t>|x|^p \\ |y|\le(p^*-1)|x|}} \{B(x,y,t) - B(0,0,1)t\} \le 0.$$

But, $U(0,0) = 0$, therefore

$$\sup_{\substack{t>|x|^p \\ |y|\le(p^*-1)|x|}} \frac{B(x,y,t)}{t} = B(0,0,1) = ((p^*-1)^2+\tau^2)^{\frac{p}{2}}. \tag{4.1}$$

Observing the relationship, $B = \mathcal{B}$, is enough to get the desired result.

For $1 < p < 2, |\tau| \le \frac{1}{2}$ and $|y| \le (p^* - 1)|x|$,

$$U(x,y) = p\left(1 - \frac{1}{p^*}\right)\left(1 + \frac{\tau^2}{(p^*-1)^2}\right)^{\frac{p-2}{2}} (|x| + |y|)^{p-1}[|y| - (p^*-1)|x|] \le 0,$$

so we have (4.1) by the same reasoning as for $2 \le p < \infty$. □

Remark 25. Note that Minkowski's inequality together with Burkholder's original result gives the estimate from above, $\langle (|g|^2 + \tau^2|f|^2)^{\frac{p}{2}}\rangle_I^{\frac{p}{2}} \le ((p^*-1)^2 + \tau^2)^{\frac{p}{2}}\langle |f|^p\rangle_I$, for $2 \le p < \infty$ and any $\tau \in \mathbb{R}$.

Indeed, if $f \in L^p[0,1]$ and g is the corresponding martingale transform then Minkowski's inequality gives,

$$\begin{aligned} \|g^2 + \tau^2 f^2\|_{L^{\frac{p}{2}}}^{\frac{p}{2}} &\le (\|g^2\|_{L^{\frac{p}{2}}} + \|\tau^2 f^2\|_{L^{\frac{p}{2}}})^{\frac{p}{2}} = (\|g\|_{L^p}^2 + \|\tau f\|_{L^p}^2)^{\frac{p}{2}} \\ &\le \|f\|_{L^p}^p((p^*-1)^2 + \tau^2)^{\frac{p}{2}}. \end{aligned}$$

This is very surprising in the sense that the "trivial" constant $((p^*-1)^2 + \tau^2)^{\frac{p}{2}}$ is actually the sharp constant if τ is small, but it only gives the estimate from above when $2 \le p < \infty$.

Now we will prove the main result for Hilbert-valued martingales. The same ideas can be used to extend the previous result to Hilbert-valued L^p-functions as well. Let $\mathbb{H}$ be a separable Hilbert space with $\|\cdot\|_{\mathbb{H}}$ as the induced norm.

Theorem 26. *Let $1 < p < \infty, (W, \mathcal{F}, \mathbb{P})$ be a probability space and $\{f_k\}_{k\in\mathbb{Z}}, \{g_k\}_{k\in\mathbb{Z}} : W \to \mathbb{H}$ be two $\mathbb{H}-$valued martingales with the same filtration $\{\mathcal{F}_k\}_{k\in\mathbb{Z}}$. Denote $d_k = f_k - f_{k-1}, d_0 = f_0, e_k = g_k - g_{k-1}, e_0 = g_0$ as the associated martingale differences. If $\|e_k(\omega)\|_{\mathbb{H}} \le \|d_k(\omega)\|_{\mathbb{H}}$, for all $\omega \in W$ and all $k \ge 0$ and $\tau \in [-\frac{1}{2}, \frac{1}{2}]$ then*

$$\left\| \left(\sum_{k=0}^n e_k, \tau \sum_{k=0}^n d_k \right) \right\|_{L^p([0,1],\mathbb{H}^2)} \le ((p^*-1)^2 + \tau^2)^{\frac{p}{2}} \left\| \sum_{k=0}^n d_k \right\|_{L^p([0,1],\mathbb{H})},$$

where $((p^*-1)^2 + \tau^2)^{\frac{p}{2}}$ is the best possible constant and $p^* - 1 = \max\{p-1, \frac{1}{p-1}\}$. For $2 < p < \infty$, the result is also true, with the best possible constant, if $|\tau| \le 1$.

In the theorem, "best possible" constant means that if $C_{p,\tau} < ((p^*-1)^2 + \tau^2)^{\frac{1}{2}}$, then for some probability space $(W, \mathcal{G}, P)$ and a filtration $\mathcal{F}$, there exists $\mathbb{H}-$valued martingales $\{f\}_k$ and $\{g\}_k$, such that

$$\|(g_k, \tau f_k)\|_{L^p([0,1],\mathbb{H}^2)} > C_{p,\tau}\, \|f_k\|_{L^p([0,1],\mathbb{H})}.$$

Proof. We will prove the result for $2 \le p < \infty$, since the result for $1 < p < 2$ is similar. Replace $|\cdot|$ with $\|\cdot\|_{\mathbb{H}}$, in $U_{p,\tau}$. Let $f_n = \sum_{k=0}^n d_k$ and $g_n = \sum_{k=0}^n e_k$. Recall that $U := U_{p,\tau}$ is the least zigzag-biconcave majorant of v. As in [8] (pages 77-79),

$$U_{p,\tau}(x+h, y+k) \le U_{p,\tau}(x,y) + \Re(\partial_x U_{p,\tau}, h) + \Re(\partial_y U_{p,\tau}, k), \tag{4.2}$$

for all $x, y, h, k \in \mathbb{H}$, such that $|k| \le |h|$ and $\|x + ht\|_{\mathbb{H}} \|x + kt\|_{\mathbb{H}} > 0$. The result in (4.2) follows from the zigzag-biconcavity and implies that $\mathbb{E}[U(f_k, g_k)]$ is a supermartingale. Lemma 21 gives us that $v(f_n, g_n) \le U(f_n, g_n)$. Therefore,

$$\mathbb{E}[v(f_n, g_n)] \le \mathbb{E}[U(f_n, g_n)] \le \mathbb{E}[U(f_{n-1}, g_{n-1})] \le \cdots \le \mathbb{E}[U(d_0, e_0)].$$

But, $\mathbb{E}[U(d_0, e_0)] \le 0$ in both pieces of U since $2 - p^* \le 0$ and $\|e_0\|_{\mathbb{H}} \le \|d_0\|_{\mathbb{H}}$. Thus, $\mathbb{E}[v(f_n, g_n)] \le 0$. The constant, in the estimate, is best possible, since it was attained in Theorem 24. □

Remark 27. For $1 < p < 2$ and $|\tau| > \frac{1}{2}$, the "trivial" constant, $((p^*-1)^2 + \tau^2)^{\frac{p}{2}}$, in the main result is no longer sharp because of a "phase transition". To give a sense of why this is true one can show that for $1 < p < 2$ fixed, the constant is no longer sharp for τ sufficiently large.

Let us construct such a function f to do this. First of all, $f_n \in L^p[0,1]$ will be chosen so that $f_n \neq 0$ a.e. Let $C_p = (p^* - 1)$. Note that

$$\begin{aligned}
\int (|g_n|^2 + \tau^2 |f_n|^2)^{\frac{p}{2}} &= |\tau|^p \int |f_n|^p \left(1 + \frac{1}{|\tau|^2} \frac{|g_n|^2}{|f_n|^2}\right)^{\frac{p}{2}} \\
&= |\tau|^p \int |f_n|^p + \frac{p}{2|\tau|^{2-p}} \int \frac{|g_n|^2}{|f_n|^{2-p}} \\
&\qquad + \frac{p}{2}\left(\frac{p}{2} - 1\right) \frac{1}{2|\tau|^{4-p}} \int \frac{|g_n|^4}{|f_n|^{4-p}} \left(\frac{1}{1 + \theta_n(x,\tau)}\right)^{2-\frac{p}{2}} \\
&=: |\tau|^p \int |f_n|^p + A + B
\end{aligned}$$

$$
\begin{aligned}
((p^*-1)^2+\tau^2)^{\frac{p}{2}}\int|f_n|^p &= |\tau|^p\left(1+\frac{C_p^2}{|\tau|^2}\right)^{\frac{p}{2}}\int|f_n|^p \\
&= |\tau|^p\int|f_n|^p+\frac{pC_p^2}{2|\tau|^{2-p}}\int|f_n|^p \\
&\quad +\frac{p}{2}\left(\frac{p}{2}-1\right)\frac{C_p^4}{2|\tau|^{4-p}}\int|f_n|^p\left(\frac{1}{1+\tilde\theta_n}\right)^{2-\frac{p}{2}} \\
&=: |\tau|^p\int|f_n|^p+C+D,
\end{aligned}
$$

where $\theta_n(x,\tau),\tilde\theta_n\geq 0$.

Choose $f_n=\chi_{[\frac{1}{8},\frac{1}{4})\cup[\frac{5}{8},\frac{3}{4})}-\chi_{[\frac{3}{8},\frac{1}{2})\cup[\frac{7}{8},1)}-\varepsilon_n\chi_{[0,\frac{1}{8})\cup[\frac{1}{2},\frac{5}{8})}+\varepsilon_n\chi_{[\frac{1}{4},\frac{3}{8})\cup[\frac{3}{4},\frac{7}{8})}$, where $\varepsilon_n>0$ is small. On the sets where $|f_n|=\varepsilon_n$, we can choose the martingale transform g_n of f_n not small. Indeed, without loss of generality choose $x\in[0,\frac{1}{8})$ and denote $J_1=[0,\frac{1}{4})$. Then

$$f_n(x)=\sum_{I:I\supset J_1}(f,h_I)h_I(x)=-\varepsilon_n$$

Define a martingale transform g_n as

$$g_n(x)=\sum_{I:I\supsetneq J_1}(f,h_I)h_I(x)-(f,h_{J_1})h_{J_1}(x).$$

Then

$$f_n(x)-g_n(x)=2(f_n,h_{J_1})h_{J_1}(x)=(2-\varepsilon_n)\sqrt{|J_1|}\left(-\frac{1}{\sqrt{|J_1|}}\right)$$

Therefore, $f_n(x)=-\varepsilon_n$, yet its martingale transform $g_n(x)=2-2\varepsilon_n$, for $x\in J_1$. The same can be done for other intervals of smallness of f_n. Note that $\int|g_n|^2=\int|f_n|^2\to\int|f|^2=\frac{1}{2}$, if we choose $\varepsilon_n\to 0$. So $\int|g_n|^2\approx\frac{1}{2}$, for n sufficiently large.

To show that $\int(g^2+\tau^2f^2)^{\frac{p}{2}}>((p^*-1)^2+\tau^2)^{\frac{p}{2}}\int|f|^p$ it suffices to show $A+B>C+D$. As $D\leq 0$ it is enough to prove $A+B-C>0$. Let $A'=\tau^{2-p}A$, $C'=\tau^{2-p}C$, $B'=\tau^{4-p}B$. Regardless of the choice of τ we have $A'>2C'$ if n is chosen sufficiently large. In fact looking at A' we see that it is bigger than the integral, where integrand has numerator close to 2 and denominator equal ε_n. On the other hand C' involves just an integral with uniformly (in n) bounded integrand. Then we fix n, of course $|B'|$ is very large, but we notice that choosing τ to be very large makes the following inequality true:

$$A+B-C \;=\; \frac{1}{|\tau|^{2-p}}[(A'-C')-\frac{1}{|\tau|^2}|B'|]>0\,.$$

This completes the example.

References

1. N. Boros, P. Janakiraman, A. Volberg, *Perturbation of Burkholder's martingale transform and Monge–Ampère equation*, (2011), preprint.
2. N. Boros, P. Janakiraman, A. Volberg, *Sharp Lp-bounds for a perturbation of Burkholders Martingale Transform*, C. R. Acad. Sci. Paris, Ser. I, doi:10.1016/j.crma.2011.01.001, (2011).
3. R. Banuelos, P. Janakiraman L^p-*bounds for the Beurling–Ahlfors transform* Trans. Amer. Math. Soc. 360 (2008), no. 7, 3603–3612.
4. D. L. Burkholder, *Boundary value problems and sharp inequalities for martingale transforms*, The Annals of Probability, vol. 12 (1984), No. 3, pp. 647–702.
5. D. Burkholder, *Boundary value problems and sharp estimates for the martingale transforms*, Ann. of Prob. **12** (1984), 647–702.
6. D. Burkholder, *An extension of classical martingale inequality*, Probability Theory and Harmonic Analysis, ed. by J.-A. Chao and W. A. Woyczynski, Marcel Dekker, (1986).
7. D. Burkholder, *Sharp inequalities for martingales and stochastic integrals*, Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987), Astérisque No. 157-158 (1988), 75–94.
8. D. Burkholder, *Differential subordination of harmonic functions and martingales*, (El Escorial 1987) Lecture Notes in Math., **1384** (1989), 1–23.
9. D. Burkholder, *Explorations of martingale theory and its applications*, Lecture Notes in Math. **1464** (1991), 1–66.
10. D. Burkholder, *Strong differential subordination and stochastic integration*, Ann. of Prob. **22** (1994), 995–1025.
11. D. Burkholder, *A proof of the Peczynski's conjecture for the Haar system*, Studia Math., **91** (1988), 79–83.
12. K.P. Choi, *Some sharp inequalities for martingale transforms*, Trans. Amer. Math. Soc. 307 (1988) 279300.
13. S. Geiss, S. Montgomery-Smith, E. Saksman *On singular integral and martingale transforms*,Trans. Amer. Math. Soc. **362** (2010), 553-575.
14. T. Iwaniec, *Extremal inequalities in Sobolev spaces and quasiconformal mappings*, Z. Anal. Anwendungen 1 (1982), 1–16.
15. O. Lehto, *Remarks on the integrability of the derivatives of quasiconformal mappings*, Ann. Acad. Sci. Fenn. Ser. A I No. **371** (1965), 3–8.
16. F. Nazarov and A. Volberg *Heating of the Ahlfors–Beurling operator and estimates of its norm*, St. Petersburg Math. J., 14 (2003) no. 3.
17. S. Pichorides,*On the best value of the constants in the theorems of Riesz, Zygmund, and Kolmogorov*, Studia Math. **44** (1972), 165-179
18. A. V. Pogorelov, Extrinsic geometry of convex surfaces, Translations of Mathematical Monographs, Amer. Math. Soc., v. 35, 1973.

19. V. Vasyunin, A. Volberg, *Monge–Ampère equation and Bellman optimization of Carleson Embedding Theorems*, arXiv:0803.2247.
20. V. Vasyunin, A. Volberg, *Bellster and others*, Preprint, 2008.
21. V. Vasyunin, A. Volberg, *Burkholder's function via Monge–Ampère equation*, arXiv:1006.2633v1

Nicholas Boros, Dept. of Math., Michigan State University. borosnic@msu.edu

Prabhu Janakiraman, Dept. of Math., Michigan State University. pjanakir1978@gmail.com

Alexander Volberg, Dept. of Math., Michigan State University. volberg@math.msu.edu