# ON WEAK WEIGHTED ESTIMATES OF MARTINGALE TRANSFORM

FEDOR NAZAROV, ALEXANDER REZNIKOV, VASILY VASYUNIN, AND ALEXANDER VOLBERG

Abstract. We consider several weak type estimates for singular operators using the Bellman function approach. We disprove the $A_1$ conjecture, which stayed open after Muckenhoupt–Wheeden's conjecture was disproved by Reguera–Thiele.

## 1. End-point estimates. Notation and facts.

The end-point estimates play an important part in the theory of singular integrals (weighted or unweighted). They are usually the most difficult estimates in the theory, and the most interesting of course. It is a general principle that one can extrapolate the estimate from the end-point situation to all other situations. We refer the reader to the book [1] that treats this subject of extrapolation in depth.

On the other hand, it happens quite often that the singular integral estimates exhibit a certain "blow-up" near the end point. To catch this blow-up can be a difficult task. We demonstrate this hunt for blow-ups by the examples of weighted dyadic singular integrals and their behavior in $L^p(w)$. The end-point $p$ will be naturally 1 (and sometimes slightly unnaturally 2) depending on the martingale singular operator. The singular integrals in this article are the easiest possible. They are dyadic martingale operators on $\sigma$-algebra generated by usual homogeneous dyadic lattice on the real line. We do not consider any non-homogeneous situation, and this standard $\sigma$-algebra generated by a dyadic lattice $\mathcal{D}$ will be provided with Lebesgue measure.

Our goal will be to show how the technique of Bellman function gives the proof of the blow-up of the weighted estimates of the corresponding weighted dyadic singular operators. This blow-up will be demonstrated by certain estimates from below of the Bellman function of a dyadic problem. Interestingly, one can bootstrap then the correct estimates from below of a dyadic operators to the estimate from below of such classical operators as e. g. the Hilbert transform. The same rate of blow-up then persists for the classical operators. But this bootstrapping argument will be carried out in

2010 *Mathematics Subject Classification.* 42B20, 42B35, 47A30.

FN is partially supported by the NSF grant DMS-1265623.

AV and VV are partially supported by the program "Research in Pairs" of the Oberwolfach Institute for Mathematics, and by the NSF grant DMS-1600065; AV is also supported by the NSF grant DMS-1265549, and VV is also supported by the RFBR grant 14-01-00748.

a separate note, here, for simplicity, we work only with dyadic martingale operators.

As to the Bellman function part of our consideration below, this part will be reduced to the task to find the lower estimate for the solutions of the homogeneous Monge–Ampère differential equation.

1.1. **End-point estimates for martingale transform.** We will work with a standard dyadic filtration $\mathcal{D} = \cup_k \mathcal{D}_k$ on $\mathbb{R}$. We consider the martingale transform related to this homogeneous dyadic filtration.

The symbol $\langle f \rangle_I$ denotes average value of $f$ over the set $I$ i. e.,

$$\langle \varphi \rangle_I = \frac{1}{|I|} \int_I \varphi(t)\, dt.$$

We consider martingale differences (recall that the symbol $\mathrm{ch}(J)$ denotes the dyadic children of $J$)

$$\Delta_J \varphi \stackrel{\text{def}}{=} \sum_{I \in \mathrm{ch}(J)} \chi_I (\, \langle \varphi \rangle_I - \langle \varphi \rangle_J).$$

For our case of dyadic lattice on the line we have that $|\Delta_J \varphi|$ is constant on $J$, and

$$\Delta_J \varphi = \frac{1}{2} (\, \langle \varphi \rangle_{J_+} - \langle \varphi \rangle_{J_-}) [\, \chi_{J_+} - \chi_{J_-}] \,.$$

We consider the dyadic $A_1$ class of weights, but we skip the word dyadic in what follows, because we consider here only dyadic operators. A positive function $w$ is called an $A_1$ weight if

$$[w]_{A_1} \stackrel{\text{def}}{=} \sup_{J \in \mathcal{D}} \frac{\langle w \rangle_J}{\inf_J w} < \infty \,.$$

By $Mw$ we will denote the martingale maximal function of $w$, that is $Mw(x) = \sup\{\langle w \rangle_J : J \in \mathcal{D},\, J \ni x\}$. Then $w \in A_1$ with "norm" $Q$ means that

$$Mw \le Q \cdot w \quad a.\, e.\,,$$

and $Q = [w]_{A_1}$ is the best constant in this inequality.

Recall that the martingale transform is the operator given by $T\varphi = \sum_{J \in \mathcal{D}} \varepsilon_J \Delta_J \varphi$. It is convenient to use Haar function $h_J$ associated with dyadic interval $J$,

$$h_J(x) := \begin{cases} \frac{1}{|J|^{1/2}}, & x \in J_+ \,; \\ -\frac{1}{|J|^{1/2}}, & x \in J_- \,. \end{cases}$$

Sometimes it is more convenient to use the Haar functions $H_J$ normalized in $L^\infty$: $H_J = |J|^{1/2} h_J$. In this notations, the martingale transform $\psi$ of a function $\varphi$ is

$$\psi = T\varphi = \sum_{J \in \mathcal{D}} \frac{\varepsilon_J}{|J|} (\varphi, H_J) H_J = \sum_{J \in \mathcal{D}} \varepsilon_J (\varphi, h_J) h_J \,.$$

In all our calculations we always think the sum has only unspecified but finite number of terms, so we may not to worry about the converges of this series. Nevertheless approximation arguments give us the final estimates for arbitrary $L^1$ function $\varphi$. As to the values of the multiplicator coefficients we consider the class $|\varepsilon_J| \le 1$ or its important subclass $\varepsilon_J = \pm 1$.

We are interested in the weak estimate for the martingale transform $T$ in the weighted space $L^1(\mathbb{R}, w\,dx)$, where $w \in A_1$. The end-point exponent is naturally $p = 1$, and we wish to understand the order of magnitude of the constant $C([w]_{A_1})$ in the weak type inequality for the dyadic martingale transform:

$$\frac{1}{|I|} w\{t \in I\colon \sum_{J\in\mathcal{D}(I)} \varepsilon_J (\varphi, h_J) h_J(t) \geq \lambda\} \leq C([w]_{A_1}) \frac{\langle|\varphi|w\rangle_I}{\lambda}\,. \tag{1.1}$$

Here $\varphi$ runs over all functions such that $\operatorname{supp}\varphi \subset I$ and $\varphi \in L^1(I, w\,dt)$, $w \in A_1$. This paper is devoted to the study of the "sharp" order of magnitude of constants $C([w]_{A_1})$ in terms of $[w]_{A_1}$ if $[w]_{A_1}$ is large. We are primarily interested in the estimate of $C([w]_{A_1})$ from below, that is in finding the worst possible $A_1$ weight in terms of weak type estimate (of course this involved also finding the worst test function $\varphi$ as well).

We will prove the following result.

**Theorem 1.1.** *There is a weight $w \in A_1$ such that constant $C([w]_{A_1})$ from* (1.1) *satisfies*

$$C(\,[w]_{A_1}) \geq \frac{1}{515}[w]_{A_1}(\log[w]_{A_1})^{1/3} \quad \text{if } [w]_{A_1} \geq 4.$$

In [3] the following estimate from above has been proved:

**Theorem 1.2.** *There is a positive absolute constant $c$ such that for any weight $w \in A_1$ estimate* (1.1) *holds with*

$$C(\,[w]_{A_1}) = c\,[w]_{A_1} \log[w]_{A_1}\,.$$

**Remark 1.3.** The sharp power remains enigmatic.

## 2. Unweighted estimate of the martingale transform

In this Section we prove the following unweighted analog of inequality (1.1)

$$\frac{1}{|I|}\big|\{t \in I\colon \sum_{J\in\mathcal{D}(I)} \varepsilon_J (\varphi, h_J) h_J(t) \geq \lambda\}\big| \leq 2\frac{\langle|\varphi|\rangle_I}{\lambda}\,. \tag{2.1}$$

We will work not on the whole $\mathbb{R}$ but on a finite interval. The result for the whole axis can be obtain by enlarging the underlying interval and the fact that the estimates will not depend on the interval. So, we are working on $I = [0, 1]$. The symbol $\mathcal{D} = \mathcal{D}(I)$ means the dyadic lattice of subintervals. Let $\varphi$ be a dyadic martingale starting at $x_1$ and $\psi$ is its martingale transform, starting at $x_2$, i. e.,

$$\varphi = x_1 + \sum_{J\subseteq I,\, J\in\mathcal{D}} (\varphi, h_J) h_J\,, \qquad \psi = x_2 + \sum_{J\subseteq I,\, J\in\mathcal{D}} \varepsilon_J (\varphi, h_J) h_J\,.$$

We consider two classes of martingale transforms: 1) the case of $\pm$-transforms, i. e. the case when we assume that $\varepsilon_J = \pm 1$; and 2) the case when the martingale $\psi$ is differentially subordinate to $\varphi$, i. e. the case when we assume that $|\varepsilon_J| \le 1$. The first class of admissible pairs $\{\varphi,\, \psi\}$ we denote by $\mathfrak{A}_\pm$, the second one by $\mathfrak{A}_\varepsilon$.

The desired estimate we deduce to estimating a certain function of three variables related to our inequality, which is called the Bellman function of the problem. In fact the Bellman function related to some inequality is simply the extremal value of the quantity we need to estimate under several fixed parameters related to the problem. Describe now the Bellman function of our problem.

With every pair of functions $\{\varphi,\, \psi\}$ on $I$ we associate the so called Bellman point $\mathfrak{b}_{\varphi,\psi} = x = (x_1, x_2, x_3)$ with coordinates

$$x_1 = \langle\varphi\rangle_I, \qquad x_2 = \langle\psi\rangle_I, \qquad x_3 = \langle|\varphi|\rangle_I.$$

The set of all admissible pairs corresponding to a point $x$ will be denoted by $\mathfrak{A}_\pm(x)$ in the case of $\pm$transform and by $\mathfrak{A}_\varepsilon(x)$ in the case of differential subordination. Our Bellman function is the following one:

$$\mathbf{B}(x) = \mathbf{B}(x_1, x_2, x_3) := \sup_{\mathfrak{A}(x)} \frac{1}{|I|}\Big|\{t \in I\colon \sum_{J\subseteq I, J\in\mathcal{D}} \psi(t) \ge 0\}\Big|\,. \tag{2.2}$$

If we would like to specify that we speak about $\pm$-transform, i. e. supremum is taken over the set $\mathfrak{A} = \mathfrak{A}_\pm$, then the corresponding Bellman function will be written as $\mathbf{B}_\pm$, and we shall write $\mathbf{B}_\varepsilon$ if $\mathfrak{A} = \mathfrak{A}_\varepsilon$. This index will be omitted in any assertion valid in both cases. It is clear, that $\mathbf{B}_\pm \le \mathbf{B}_\varepsilon$, but as we will see at the end these two functions coincide. Note that the function $\mathbf{B}$ should not be indexed by $I$ because it is easy to check that this function does not depend on $I$.

### 2.1. **Properties of B.**

2.1.1. *Domain and Range.* Formally the definition of $\mathbf{B}$ is correct for arbitrary $x \in \mathbb{R}^3$, but there is no sense to consider $\mathbf{B}$ at the points where the set of admissible functions is empty, and therefore the corresponding supremum is $-\infty$. We would like to consider the function $\mathbf{B}$ on the domain $\Omega \subset \mathbb{R}^3$:

$$\Omega \stackrel{\text{def}}{=} \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 : |x_1| \le x_3\}\,.$$

For any $x \in \Omega$, the set of test functions $\mathfrak{A}(x)$ is not empty and it is immediately clear from the definition that

$$0 \le \mathbf{B}(x) \le 1\,.$$

2.1.2. *Symmetry.* The function $\mathbf{B}$ is invariant under reflection with respect $x_1$:

$$\mathbf{B}(-x_1, x_2, x_3) = \mathbf{B}(x_1, x_2, x_3)\,,$$

because if $\mathfrak{b}_{\varphi,\psi} = x = (x_1, x_2, x_3)$, then $\widetilde{\varphi} = -\varphi$, $\widetilde{\psi} = \psi$ is an admissible pair corresponding to $\widetilde{x} = (-x_1, x_2, x_3)$.

2.1.3. *Homogeneity.*

$$\mathbf{B}(\tau x_1, \tau x_2, \tau x_3) = \mathbf{B}(x_1, x_2, x_3)\,, \qquad \tau > 0\,,$$

because if $\mathfrak{b}_{\varphi,\psi} = x = (x_1, x_2, x_3)$, then $\widetilde{\varphi} = \tau\varphi$, $\widetilde{\psi} = \tau\psi$ is an admissible pair corresponding to $\widetilde{x} = (\tau x_1, \tau x_2, \tau x_3)$ and the functions $\psi$ and $\widetilde{\psi}$ are positive simultaneously.

2.1.4. *Boundary condition.*

$$\mathbf{B}(0,x_2,0)=\begin{cases}1\,, & \text{if } \ x_2\ge 0\,,\\ 0\,, & \text{if } \ x_2<0\,,\end{cases} \tag{2.3}$$

because the only admissible pair for the point $x=(0,x_2,0)$ is $\varphi=0$, $\psi=x_2$.

2.1.5. *Obstacle condition.*

$$\mathbf{B}(x_1,x_2,|x_1|)\ge\begin{cases}1\,, & \text{if } \ x_2\ge 0\,,\\ 0\,, & \text{if } \ x_2<0\,,\end{cases} \tag{2.4}$$

because the pair of constant functions $\varphi=x_1$, $\psi=x_2$ is an admissible pair for the point $x=(x_1,x_2,|x_1|)$.

By the way, since $\mathbf{B}\le 1$ by the definition, the obstacle condition supplies us with the function on the half of the boundary, namely, $\mathbf{B}(x)=1$ if $x$ is on the boundary and $x_2\ge 0$. We shall see soon that this is not the whole part of the boundary where $\mathbf{B}(x)=1$. However first we derive the main inequality.

### 2.2. **Main inequality.**

**Lemma 2.1.** *Let $x^\pm$ be two points in $\Omega$ such that*

- $|x_2^+-x_2^-|=|x_1^+-x_1^-|$ *in the case* $\mathbf{B}=\mathbf{B}_\pm$;
- $|x_2^+-x_2^-|\le|x_1^+-x_1^-|$ *in the case* $\mathbf{B}=\mathbf{B}_\varepsilon$,

*and $x=\frac12(x^++x^-)$. Then*

$$\mathbf{B}(x)-\frac{\mathbf{B}(x^+)+\mathbf{B}(x^-)}{2}\ge 0\,. \tag{2.5}$$

*Proof.* Fix $x^\pm\in\Omega$, and let $\varphi^\pm$, $\psi^\pm$ be two pairs of test functions giving the supremum in $\mathbf{B}(x^+)$, $\mathbf{B}(x^-)$ respectively up to a small number $\eta>0$. Using the fact that the function $\mathbf{B}$ does not depend on the interval where the test functions are defined, we assume that $\varphi^+$, $\psi^+$ lives on $I^+$ and $\varphi^-$, $\psi^-$ lives on $I^-$, where $I^\pm$ are two halves of the interval $I$:

$$\varphi^\pm=x_1^\pm+\sum_{J\subseteq I^\pm,\,J\in\mathcal{D}}a_{\scriptscriptstyle J}h_{\scriptscriptstyle J}\,,\qquad \psi^\pm=x_2^\pm+\sum_{J\subseteq I^\pm,\,J\in\mathcal{D}}\varepsilon_{\scriptscriptstyle J}a_{\scriptscriptstyle J}h_{\scriptscriptstyle J}\,.$$

And we assume that for these functions the estimates

$$\frac{1}{|I^\pm|}\big|\{t\in I^\pm\colon\psi^\pm(t)\ge 0\}\big|\ge\mathbf{B}(x^\pm)-\eta$$

hold. Consider the functions

$$\varphi(t):=\begin{cases}\varphi^+(t)\,, & \text{if } \ t\in I^+\\ \varphi^-(t)\,, & \text{if } \ t\in I^-\end{cases}=\frac{x_1^++x_1^-}{2}+\frac{x_1^+-x_1^-}{2}h_{\scriptscriptstyle I}+\sum_{J\subseteq I,\,J\in\mathcal{D}}a_{\scriptscriptstyle J}h_{\scriptscriptstyle J}$$

and

$$\psi(t):=\begin{cases}\psi^+(t)\,, & \text{if } \ t\in I^+\\ \psi^-(t)\,, & \text{if } \ t\in I^-\end{cases}=\frac{x_2^++x_2^-}{2}+\frac{x_2^+-x_2^-}{2}h_{\scriptscriptstyle I}+\sum_{J\subseteq I,\,J\in\mathcal{D}}\varepsilon_{\scriptscriptstyle J}a_{\scriptscriptstyle J}h_{\scriptscriptstyle J}\,.$$

Under our assumption about relation between $|x_1^+-x_1^-|$ and $|x_2^+-x_2^-|$ we have $\{\varphi,\psi\}\in\mathfrak{A}_1$ in the first case and $\{\varphi,\psi\}\in\mathfrak{A}_\varepsilon$ in the second one, i. e.

in each case this is an admissible pair of the test functions corresponding to the point $x$. Therefore,

$$\begin{aligned}\mathbf{B}(x) &\geq \frac{1}{|I|}\big|\{t\in I\colon \psi(t)\geq 0\}\big| \\ &= \frac{1}{2|I^+|}\big|\{t\in I^+\colon \psi(t)\geq 0\}\big| + \frac{1}{2|I^-|}\big|\{t\in I^-\colon \psi(t)\geq 0\}\big| \\ &\geq \frac{1}{2}\mathbf{B}(x^+)+\frac{1}{2}\mathbf{B}(x^-)-\eta.\end{aligned}$$

Since this inequality holds for an arbitrarily small $\eta$, we can pass to the limit $\eta\to 0$, what gives us the required assertion. □

It will be convenient to change variables $x_1 = y_1 - y_2$, $x_2 = y_1 + y_2$, $x_3 = y_3$ and introduce a function $\mathbf{M}(y) \stackrel{\text{def}}{=} \mathbf{B}(x)$ defined in the domain $G \stackrel{\text{def}}{=} \{y\in\mathbb{R}^3\colon |y_1-y_2|\leq y_3\}$. Then the main inequality for the function $\mathbf{M}_\pm$ means that it is concave if either $y_1$ is fixed, or $y_2$ is fixed. For the function $\mathbf{M}_\varepsilon$ the condition is more restrictive: it is concave in any direction from the cone $(y_1^+-y_1^-)(y_2^+-y_2^-)\leq 0$, since

$$(x_2^+-x_2^-)^2-(x_1^+-x_1^-)^2 = 4(y_1^+-y_1^-)(y_2^+-y_2^-)\,.$$

### 2.3. **Supersolution.**

**Lemma 2.2.** *Let $B$ a continuous function on $\Omega$ satisfying the main inequality* (2.5) *and the obstacle condition* (2.4)*. Then $\mathbf{B}(x)\leq B(x)$.*

*Proof.* Let us fix a point $x\in\Omega$ and a pair of admissible functions $\varphi$, $\psi$ on $I$ corresponding to $x$, i. e., $\mathfrak{b}_{\varphi,\psi}=x$. Using consequently main inequality for the function $B$ we can write down the following chain of inequalities

$$\begin{aligned}B(\mathfrak{b}_{\varphi,\psi}) &\geq \frac{1}{2}\big(B(\mathfrak{b}_{\varphi|I^+,\psi|I^+})+B(\mathfrak{b}_{\varphi|I^-,\psi|I^-})\big) \\ &\geq \sum_{J\in\mathcal{D},\,|J|=2^{-n}}\frac{1}{|J|}B(\mathfrak{b}_{\varphi|J,\psi|J}) = \int_0^1 B(x^{(n)}(t))dt\,,\end{aligned}$$

where $x^{(n)}(t)=\mathfrak{b}_{\varphi|J,\psi|J}$, if $t\in J$, $|J|=2^{-n}$.

Note that $x^{(n)}(t)\to(\varphi(t),\psi(t),|\varphi(t)|)$ almost everywhere (at any Lebesgue point $t$), and therefore, since $B$ is continuous and bounded, we can pass to the limit in the integral. So, we come to the inequality

$$B(x)\geq\int_0^1 B(\varphi(t),\psi(t),|\varphi(t)|)dt \geq \int\limits_{\{t\colon\,\psi(t)\geq 0\}} dt \;=\; \big|\{t\in I_0\colon \psi(t)\geq 0\}\big|\,, \tag{2.6}$$

where we have used the property $B(x_1,x_2,|x_1|)=1$ for $x_2\geq 0$. Now, taking supremum in (2.6) over all admissible pairs $\varphi$, $\psi$, we get the required estimate $B(x)\geq\mathbf{B}(x)$. □

Now we explain how we will apply this lemma. For a given sequence $\varepsilon=\{\varepsilon_J\}$, we denote

$$T\varphi \stackrel{\text{def}}{=} \sum_{J\subseteq I,\,J\in\mathcal{D}}\varepsilon_J(\varphi,h_J)h_J(x)\,.$$

It is a dyadic singular operator (actually, it is a family of operators enumerated by sequences $\varepsilon$). To prove that it is of weak type $(1,1)$ is the same as to prove

$$\mathbf{B}(x) \le \frac{C\,x_3}{|x_2|}\,.$$

Indeed, if $\varphi$, $\psi$ is an admissible pair corresponding to the point $x$, then $T\varphi = \psi - x_2$. Therefore, for a given $\varphi$ with $\langle\varphi\rangle = x_1$ and $\langle|\varphi|\rangle = x_3$ the best estimate of the value $|\{t\colon T\varphi \ge \lambda\}|$ gives us the function $\mathbf{B}(x)$ with $x_2 = -\lambda$. Thus, would we find any function $B$ with the required estimate and satisfying conditions of Lemma 2.2 we immediately get the needed weak type $(1,1)$, and in fact, more precise information on the level set of $T\varphi$.

2.4. **The Bellman function on the boundary.** First of all we note that the boundary $\partial\Omega$ consists of two independent parts

$$\partial\Omega_+ \stackrel{\text{def}}{=} \{x=(x_1,x_2,x_1)\colon x_1\ge 0, -\infty < x_2 < +\infty\} \qquad \text{and}$$
$$\partial\Omega_- \stackrel{\text{def}}{=} \{x=(x_1,x_2,-x_1)\colon x_1\le 0, -\infty < x_2 < +\infty\}.$$

They are independent in the following sense. If we have a pair of test functions $\varphi$, $\psi$ whose Bellman point $x = \mathfrak{b}_{\varphi,\psi}$ is on the boundary (whence the sign of $\varphi(t)$ is constant on the whole interval), then after splitting the interval we get a pair of Bellman points $x^{\pm}$ from the same part of the boundary. So, the main inequality (2.5) has to be fulfilled separately on $\partial\Omega_+$ and $\partial\Omega_-$. Due to the symmetry condition it is sufficient to find the function, say, on $\partial\Omega_+$ and further we assume that $x_1 \ge 0$.

So we look for a minimal function on the half-plane $\{x_1 \ge 0\}$ satisfying the main inequality and the boundary condition (2.3). We pass to the variable $y$ ($x_1 = y_1 - y_2$ and $x_2 = y_1 + y_2$) and look for a function $M$ in the half-plane $y_2 < y_1$, which satisfies the main inequality (i. e. is concave in each variable: in $y_1$, when $y_2$ is fixed, and in $y_2$, when $y_1$ is fixed) and with the given values on the boundary $y_2 = y_1$: $\mathbf{M} = 1$ if $y_1 = y_2 \ge 0$ and $\mathbf{M} = 0$ if $y_1 = y_2 < 0$.

First, we use concavity of $\mathbf{M}$ with respect to $y_2$ for some fixed $y_1 \ge 0$. Concave function bounded from below cannot decrease, therefore it has to be identically 1 on any such ray due to fixed boundary condition. It remains to find $\mathbf{M}$ in the domain $y_2 < y_1 < 0$. Here we use concavity along $y_1$. We know that our function is 0 at $y_1 = y_2$ and, by what we just said, it is 1 at $y_1 = 0$, therefore between these two points it is at least the linear function $M = 1 - \frac{y_1}{y_2}$, i. e. $\mathbf{M} \ge M$, where

$$M = \begin{cases} 1, & \text{if } y_1 \ge 0,\\ 1 - \frac{y_1}{y_2}, & \text{if } y_1 < 0. \end{cases}$$

To prove the opposite inequality we note that $M$ is concave in each variable and it satisfies the obstacle condition. Therefore, Lemma 2.2 guarantees the required inequality $\mathbf{M}_{\pm} \le M$. To prove that $\mathbf{M}_{\varepsilon} \le M$ we need to check a bit stronger concavity along any direction from the cone $(y_1^+ - y_1^-)(y_2^+ - y_2^-) \le 0$. This will be made below when considering the function in the whole domain.

Returning to variable $x$, we can write $\mathbf{B} = \frac{2x_1}{x_1 - x_2}$ in the half-plane $x_1 \ge 0$.

As a result we have proved the following

**Proposition 2.3.**

$$\mathbf{B}_{\pm}(x_1,x_2,|x_1|)=B(x_1,x_2,|x_1|)\stackrel{\text{def}}{=}\begin{cases}1, & \textit{if}\ \ x_2\ge -|x_1|,\\ \frac{2|x_1|}{|x_1|-x_2}, & \textit{if}\ \ x_2\le -|x_1|.\end{cases}\tag{2.7}$$

2.5. **Full Bellman function for the weak type estimate.** Now we present the full Bellman function:

**Theorem 2.4.** *For the function* $\mathbf{B}$ *defined by* (2.2) *we have the following analytic expression*

$$\mathbf{B}(x)=B(x)\stackrel{\text{def}}{=}\begin{cases}1, & \textit{if}\ \ x_3+x_2\ge 0,\\ 1-\dfrac{(x_3+x_2)^2}{x_2^2-x_1^2}, & \textit{if}\ \ x_3+x_2<0.\end{cases}\tag{2.8}$$

*Proof.* As above we change variables

$$x_1=y_1-y_2\,,\ x_2=y_1+y_2\,,\ x_3=y_3\,,\quad \text{i.e.}\quad y_1=\frac{x_1+x_2}{2}\,,\ y_2=\frac{x_2-x_1}{2}\,.$$

and will be looking for a function $M$

$$M(y)\stackrel{\text{def}}{=}B(x)\,,$$

which is defined in $\Omega\stackrel{\text{def}}{=}\{y=(y_1,y_2,y_3)\,:\,y_3\ge|y_1-y_2|\}$, concave in variables $(y_1,y_3)$ and $(y_2,y_3)$, satisfies boundary condition (2.7), or in term of $M$

$$M(y_1,y_2,|y_1-y_2|)=\begin{cases}1, & \text{if}\ \ y_1\ge 0 \text{ or } y_2\ge 0,\\ 1-\dfrac{\max\{y_1,y_2\}}{\min\{y_1,y_2\}}, & \text{if}\ \ y_1<0 \text{ and } y_2<0.\end{cases}$$

Since the function $\mathbf{B}$ is even with respect to $x_1$, as before it is sufficient to consider the half-space $\{x_1>0\}$, or the half-space $\{y_2<y_1\}$ in $y$-variable. But in fact we can restrict ourselves to the cone $\{x_2<-x_1<0,\ x_3>x_1\}$ or $\{y_2<y_1<0,\ y_3>y_1-y_2\}$, because for $y_1>0$ our function is identically 1 by the same reason as before: it is concave and bounded by 1 on every ray $\{y_1=\text{const}, y_2=\text{const}, y_3>y_1-y_2\}$.

The boundary function is not smooth because the boundary itself is not smooth at the line $\{x_1=x_3=0\}$ and moreover, the boundary condition on this line has a jump. But inside the domain we can look for a smooth candidate $B$. Then it has to satisfy the boundary condition $\frac{\partial B}{\partial x_1}|_{x_1=0}=0$, or in terms of $M$

$$\frac{\partial M}{\partial y_1}\Big|_{y_1=y_2}=\frac{\partial M}{\partial y_2}\Big|_{y_1=y_2}.\tag{2.9}$$

Our function has to be concave in each plane $\{y_1=\text{const}\}$ and in each plane $\{y_2=\text{const}\}$ and we look for a candidate such that its concavity is degenerate in one of these planes, i. e. in that plane the function $M$ satisfies the Monge–Ampère equation. Looking on the boundary we see that the extremals are segments of the lines $\{y_2=\text{const}\}$ and therefore it is natural to look for a solution of the Monge–Ampère equation

$$M_{y_1y_1}M_{y_3y_3}-M^2_{y_1y_3}=0$$

in this plane. (Section of our domain $\Omega$ by this plane is shown on Figure 1.) Note that the half-lines $\{y_3+y_1=\text{const}\colon y_3>y_1-y_2\}$ are in the domain if

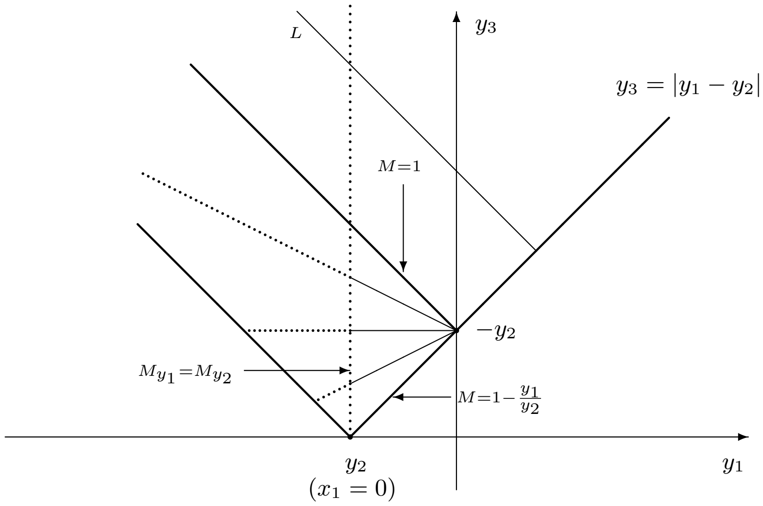


FIGURE 1. Intersection of the domain $\Omega$ with a plane $y_2 = \text{const}$

const $\geq y_2$. Moreover, if const $\geq -y_2$ (recall that $y_2 < 0$), then the boundary value on this ray (the ray $L$ on Fig. 1) is 1, and hence it is identically 1 for $y_3 + y_2 + y_1 \geq 0$, by the same reason as before: concave function bounded from below cannot decrease on an infinite interval.

Therefore we need to solve the Monge–Ampère equation only in the triangle with the vertices $(0, y_2, -y_2)$, $(y_2, y_2, 0)$, and $(y_2, y_2, -2y_2)$:

$$\{y = (y_1, y_2, y_3) \colon y_2 = \text{const}, y_1 > y_2, y_1 - y_2 < y_3 < -y_1 - y_2\}$$

with the boundary conditions

$$M(y_1, y_2, y_1 - y_2) = 1 - \frac{y_1}{y_2}, \qquad M(y_1, y_2, -y_1 - y_2) = 1,$$
$$M_{y_1}(y_2, y_2, y_3) = M_{y_2}(y_2, y_2, y_3)\,.$$

Our function is linear on two sides of the triangle, so the minimal concave function linear on two sides is the linear function it the whole triangle, however this function does not satisfies the boundary condition on the side $y_1 = y_2$. Therefore, the extremal lines cannot intersect inside the triangle and the only way to foliate this triangle without singularities inside the domain is a fan of straight line segments starting from the point $(0, y_2, -y_2)$, which we parametrize by the slope $k$ of each extremal line:

$$y_3 = ky_1 - y_2\,. \tag{2.10}$$

The slope runs over the interval $[-1, 1]$. For $k = -1$ we get the upper side of the triangle $y_1 + y_2 + y_3 = 0$ where $M = 1$, for $k = 1$ we have the lower side $y_3 = y_1 - y_2$ where $M = 1 - \frac{y_1}{y_2}$. On all other extremal lines $M$ is linear in $y_1$ as well

$$M = 1 + m(k, y_2)y_1$$

and our task is to find its slope $m = m(k, y_2)$ with the prescribed values at the points $k = \pm 1$: $m(-1, y_2) = 0$ and $m(1, y_2) = -\frac{1}{y_2}$. We find this function from the boundary condition (2.9) on the third side of the triangle.

First we deduce from (2.10) that $k = k(y_1, y_2, y_3) = \frac{y_3+y_2}{y_1}$ and hence

$$\frac{\partial k}{\partial y_1} = -\frac{k}{y_1} \qquad \text{and} \qquad \frac{\partial k}{\partial y_2} = \frac{1}{y_1}\,.$$

Therefore,

$$\frac{\partial M}{\partial y_1} = m + y_1 \frac{\partial m}{\partial k}\,\frac{\partial k}{\partial y_1} = m - k\frac{\partial m}{\partial k}\,,$$

$$\frac{\partial M}{\partial y_2} = y_1\Big(\frac{\partial m}{\partial y_2} + \frac{\partial m}{\partial k}\,\frac{\partial k}{\partial y_2}\Big) = y_1\frac{\partial m}{\partial y_2} + \frac{\partial m}{\partial k}\,.$$

Thus, the boundary condition (2.9) turns into the following equation

$$m - (k+1)\frac{\partial m}{\partial k} = y_2\frac{\partial m}{\partial y_2}\,,$$

which has the general solution of the form

$$m(k, y_2) = (k+1)\Phi\Big(\frac{k+1}{y_2}\Big)\,,$$

where $\Phi$ is an arbitrary function. Since $m(1, y_2) = -\frac{1}{y_2}$, we have $\Phi(t) = -\frac{t}{4}$. And finally

$$M(y) = 1 - \frac{(k+1)^2}{4y_2}y_1 = 1 - \frac{(y_1+y_2+y_3)^2}{4y_1y_2}\,, \qquad \text{if} \quad y_1+y_2+y_3 < 0,$$

or

$$B(x) = 1 - \frac{(x_2+x_3)^2}{x_2^2 - x_1^2}\,, \qquad \text{if} \quad x_2+x_3 < 0.$$

And our function is identically one on the rest of the domain.

Now it is an easy task to check that the found function $M$ satisfies concavity conditions from Lemma 2.1. Since our candidate is $C^1$-smooth function, the desired concavity is sufficient to check only on the subdomain, where our candidate is less than one, i. e., where $y_1 + y_2 + y_3 < 0$. For us there is important that $y_1 < 0$ and $y_2 < 0$ on this part of the domain. We shall check the main inequality (condition 2.5) in the differential form, namely, we check that the quadratic form of the Hessian of $M$ is not positive in the required directions. Direct calculations gives the following expression for the Hessian matrix:

$$\frac{d^2M}{dy^2} = \begin{pmatrix} M_{y_1y_1} & M_{y_1y_2} & M_{y_1y_3} \\ M_{y_2y_1} & M_{y_2y_2} & M_{y_2y_3} \\ M_{y_3y_1} & M_{y_3y_2} & M_{y_3y_3} \end{pmatrix} = \begin{pmatrix} -\dfrac{(y_2+y_3)^2}{2y_1^3y_2} & \dfrac{y_1^2+y_2^2+y_3^2}{4y_1^2y_2^2} & \dfrac{y_2+y_3}{2y_1^2y_2} \\[2ex] \dfrac{y_1^2+y_2^2+y_3^2}{4y_1^2y_2^2} & -\dfrac{(y_1+y_3)^2}{2y_1y_2^3} & \dfrac{y_1+y_3}{2y_1y_2^2} \\[2ex] \dfrac{y_2+y_3}{2y_1^2y_2} & \dfrac{y_1+y_3}{2y_1y_2^2} & -\dfrac{1}{2y_1y_2} \end{pmatrix}$$

and its quadratic form can be written as follows:

$$\Big(\frac{d^2M}{dy^2}\Delta,\Delta\Big)=-\frac{1}{2y_1y_2}\Big(\Delta_3-\frac{y_1+y_3}{y_2}\Delta_2-\frac{y_2+y_3}{y_1}\Delta_1\Big)^2+\frac{(y_1+y_2+y_3)^2}{2y_1^2y_2^2}\Delta_1\Delta_2\,.$$

In our part of the domain we have $y_1<0$ and $y_2<0$, therefore this quadratic form is negative if $\Delta_1\Delta_2\le 0$. So, due to Lemma 2.2 we have inequality $\mathbf{B}_\pm(x)\le\mathbf{B}_\varepsilon(x)\le B(x)$.

To prove the theorem we need to check the converse inequality $\mathbf{B}(x)\ge B(x)$. For this Bellman function it is very easy due to its following special property. Note that the function $M$ is linear on the extremal lines not only in the triangle mentioned above, but also on the continuation of each extremal line as well (see Fig. 1). Indeed, all extremal lines in the triangle under investigation are parametrized by their slope $k$, $-1<k<1$, and have the form

$$y_3=ky_1-y_2,\qquad y_2\le y_1\le 0,$$

and the found function on this line is

$$M(y_1,y_2,ky_1-y_2)=1-\frac{(k+1)^2}{4y_2}y_1\,.$$

Thus, we see that this function is linear not only on the interval $y_1\in(y_2,0)$, but for $y_1<y_2$ as well. So we can continue this extremal line up to its second point of intersection with the boundary $y_3=|y_2-y_1|$, where this $M$ coincides with $\mathbf{M}$. In result we have two points where the concave function $\mathbf{M}$ coincides with the linear function $M$, therefore between these two points we have $\mathbf{M}(y)\ge M(y)$. Since the described continued extremal line foliate the whole domain $y_1+y_2+y_3<0$, we have the desired inequality for arbitrary point $y$ from $\Omega$. $\square$

**Remark 2.5.** We would like to mention that the function (2.8) was published by A. Osękowski in [6]. It was found him absolutely independently, but a bit later than the preliminary version of this paper was accessible in the web (see [4]). However we would like to emphasize that in [6] not only this function is presented supplying us with the estimate of the measure where $\{\psi\ge\lambda\}$, but the more difficult function giving the estimate for the set $\{|\psi|\ge\lambda\}$ is found as well.

2.6. **About coincidence of $\mathbf{B}_\pm$ with $\mathbf{B}_\varepsilon$.** In this subsection we would like to underline that the fact of this coincidence is absolutely not evident. In many cases as in the famous $L^p$ result of Burkholder the estimation for differentially subordinate martingales is the same as for $\pm$-transform. And the natural reason for this is that any differentially subordinate martingale is a convex combination of $\pm$-transforms. Indeed, if we fix a martingale $\psi$ being differentially subordinate to $\varphi$, i. e.

$$T_\varepsilon\varphi\stackrel{\mathrm{def}}{=}\psi=\sum_{J\in\mathcal{D}(I)}\varepsilon_J(\varphi,h_J)h_J,\qquad|\varepsilon_J|\le 1\,,$$

then every number $\varepsilon_J$ can be represented as a convex combination of $\pm1$:

$$\varepsilon_J=\sum_{k=1}^\infty 2^{-k}\varepsilon_{k,J}\,,\qquad \varepsilon_{k,J}=\pm1\,.$$

Therefore,

$$T_\varepsilon = \sum_{k=1}^{\infty} 2^{-k} T_{\varepsilon_k}\,,$$

If we were interested in the estimate of $T_\varepsilon$ in a Banach space $X$ (say, $X = L^p$, $p > 1$), then this representation would show that

$$\sup_{\varepsilon\,:\,\varepsilon_J \in [-1,1]} \|T_\varepsilon\|_X \;=\; \sup_{\varepsilon\,:\,\varepsilon_J \in \{-1,1\}} \|T_\varepsilon\|_X\,.$$

However, we are interested in the case $X = L^{1,\infty}$. Here one can use Lemma of Stein and Weiss:

**Lemma 2.6.** *Let $\{g_j\}$ be a sequence of non-negative measurable functions, such that $\|g_j\|_{L^{1,\infty}} \le 1$ for all $j$. Let $\{c_j\}$ be a sequence of non-negative scalars such that $\sum c_j = 1$ and $\sum c_j \log \frac{1}{c_j} = K < \infty$. Then*

$$\Big\| \sum_j c_j g_j \Big\|_{L^{1,\infty}} \le 2(K+2)\,.$$

See [9] for the proof. From this lemma, we would conclude that

$$\sup_{\varepsilon\,:\,\varepsilon_J\in[-1,1]} \|T_\varepsilon\|_{L^{1,\infty}} \le\; 2\big(2 + \log 2 \sum_{k=1}^{\infty} k 2^{-k}\big) \sup_{\varepsilon\,:\,\varepsilon_J\in\{-1,1\}} \|T_\varepsilon\|_{L^{1,\infty}}\,.$$

However, Theorem 2.4 gives a better result:

**Corollary 2.7.**

$$\sup_{\varepsilon\,:\,\varepsilon_J\in[-1,1]} \|T_\varepsilon\|_{L^{1,\infty}} \;=\; \sup_{\varepsilon\,:\,\varepsilon_J\in\{-1,1\}} \|T_\varepsilon\|_{L^{1,\infty}}\,.$$

## 3. The Bellman function of weak weighted estimate of martingale transform and its properties.

Passing to the weighted case we need to investigate a Bellman function of more variables. Now two additional variables $x_4$ and $x_5$ appear describing a test weight $w$. We put

$$x_4 = \langle w\rangle_I \qquad \text{and} \qquad x_5 = \inf_I w\,.$$

The test weight $w$ will run over the set of all $A_1$ weight with $[w]_{A_1} \le Q$ and with the prescribed parameters $x_4$ and $x_5$. This, by the way, means that these parameters must satisfy the following condition: $x_4 < Qx_5$.

The coordinates $x_1$ and $x_2$ will be the same, but the coordinate $x_3$ we need to change slightly:

$$x_3 = \langle |\varphi| w\rangle_I,$$

because now we fix a wighted norm of the test function $\varphi \in L^1(I, w\,dx)$. A Bellman point $x = (x_1, x_2, x_3, x_4, x_5) = \mathfrak{b}_{\varphi,\psi,w}$ is defined by a dyadic martingale $\varphi$ started at $x_1$, by a subordinated to $\varphi$ martingale $\psi$ started at $x_2$, and by a $A_1$ weight $w$. The Bellman function at this point is defined as follows:

$$\mathbf{B}(x) \overset{\text{def}}{=} \mathbf{B}_Q(x) \overset{\text{def}}{=} \sup \frac{1}{|I|} w\{t \in I : \psi(t) \ge 0\}\,, \tag{3.1}$$

where the supremum is taken over all admissible triples $\varphi$, $\psi$, $w$. We mark the Bellman function by the index $Q$ to emphasize that it depend on a fixed parameter $Q$. And in fact we are interested just in the dependence of $\mathbf{B}_Q$ on this parameter. However during our calculations we will omit this index.

This Bellman function is defined in the following subdomain of $\mathbb{R}^5$:

$$\Omega := \{(x \in \mathbb{R}^5\colon x_3 \ge |x_1|x_5,\ 0 < x_5 \le x_4 \le Qx_5\}\,. \tag{3.2}$$

Note that formally the Bellman function is defined on the whole $\mathbb{R}^5$, but in the domain $\Omega$ we include only the points, for which the set of test functions is not empty and therefore $\mathbf{B}(x) \ne -\infty$ (we would like to assume that $\mathbf{B} \ge 0$).

3.1. **The properties of $\mathbf{B}_Q$.**

3.1.1. *The first property: boundary conditions.* On the boundary $x_4 = x_5$ the weight is a constant function $w = x_4 = x_5$, and therefore

$$\mathbf{B}(x_1, x_2, x_3, x_5, x_5) = \begin{cases} x_5, & \text{if } \ x_3 + x_2x_5 \ge 0, \\ x_5\Big(1 - \dfrac{(\frac{x_3}{x_5} + x_2)^2}{x_2^2 - x_1^2}\Big), & \text{if } \ x_3 + x_2x_5 < 0. \end{cases}$$

3.1.2. *The second property: the homogeneity.* It is clear that if $\{\varphi, \psi, w\}$ is the set of admissible triples for a point $x \in \Omega$, then the set $\{s_1\varphi, s_1\psi, s_2w\}$ is admissible for the point

$$\tilde{x} = (s_1x_1, s_1x_2, s_1s_2x_3, s_2x_4, s_2x_5)$$

for an arbitrary pair of positive numbers $s_1$, $s_2$. Then by the definition of the Bellman function we have

$$\mathbf{B}(\tilde{x}) = s_2\mathbf{B}(x)\,. \tag{3.3}$$

In what follows we deal mainly with the restriction $\mathcal{B}$ of $\mathbf{B}$ to the three-dimensional affine plane

$$G = \{x \in \Omega\colon x_2 = -1,\ x_5 = 1\} = \{(x_1, x_3, x_4)\colon |x_1| \le x_3,\ 1 \le x_4 \le Q\}\,, \tag{3.4}$$

i. e. the function

$$\mathcal{B}(x_1, x_3, x_4) = \mathbf{B}(x_1, -1, x_3, x_4, 1)\,.$$

For $x_2 \ge 0$ we always have $\mathbf{B}(x) = x_4$ because for any such point the constant test function $\psi = x_2$ is admissible, and for $x_2 < 0$ we can reconstruct $\mathbf{B}$ from $\mathcal{B}$ due to homogeneity (3.3): choosing $s_1 = -x_2^{-1}$ and $s_2 = x_5^{-1}$ we get

$$\mathbf{B}(x) = x_5\mathbf{B}(-\frac{x_1}{x_2}, -1, -\frac{x_3}{x_2x_5}, \frac{x_4}{x_5}, 1) = x_5\mathcal{B}(-\frac{x_1}{x_2}, -\frac{x_3}{x_2x_5}, \frac{x_4}{x_5}\,)\,. \tag{3.5}$$

3.1.3. *The third property: special form of concavity.* Here we state our main inequality, the weighted analog of Lemma 2.1.

**Lemma 3.1.** *Let $x^\pm$ be two points in $\Omega$ such that $|x_2^+ - x_2^-| \le |x_1^+ - x_1^-|$ and let the point $x$ with $x_i = \frac{1}{2}(x_i^+ + x_i^-)$ for $1 \le i \le 4$ and $x_5 = \min\{x_5^+, x_5^-\}$ be in $\Omega$ as well. Then*

$$\mathbf{B}(x) - \frac{\mathbf{B}(x^+) + \mathbf{B}(x^-)}{2} \ge 0\,. \tag{3.6}$$

*Proof.* We repeat almost verbatim the proof of Lemma 2.1. Fix $x^\pm \in \Omega$, and take two triples of test functions $\varphi^\pm$, $\psi^\pm$, $w^\pm$ giving the supremum in $\mathbf{B}(x^+)$, $\mathbf{B}(x^-)$ respectively up to a small number $\eta > 0$. Using the fact that the function $\mathbf{B}$ does not depend on the interval where the test functions are defined, we assume that $\varphi^+$, $\psi^+$, $w^+$ live on $I_0^+$ and $\varphi^-$, $\psi^-$, $w^-$ live on $I_0^-$, i.e.,

$$\varphi^\pm = x_1^\pm + \sum_{I\subseteq I_0^\pm,\, I\in\mathcal{D}} a_I h_I\,, \qquad \psi^\pm = x_2^\pm + \sum_{I\subseteq I_0^\pm,\, I\in\mathcal{D}} \varepsilon_I a_I h_I\,, \quad |\varepsilon_I| \le 1\,.$$

Consider

$$\varphi(t) := \begin{cases} \varphi^+(t)\,, & \text{if } t\in I^+ \\ \varphi^-(t)\,, & \text{if } t\in I^- \end{cases} = \frac{x_1^+ + x_1^-}{2} + \frac{x_1^+ - x_1^-}{2} h_{I_0} + \sum_{I\subseteq I_0,\, I\in D} a_I h_I\,,$$

$$\psi(t) := \begin{cases} \psi^+(t)\,, & \text{if } t\in I^+ \\ \psi^-(t)\,, & \text{if } t\in I^- \end{cases} = \frac{x_2^+ + x_2^-}{2} + \frac{x_2^+ - x_2^-}{2} h_{I_0} + \sum_{I\subseteq I_0,\, I\in D} \varepsilon_I a_I h_I\,.$$

and

$$w(t) := \begin{cases} w^+(t)\,, & \text{if } t\in I^+, \\ w^-(t)\,, & \text{if } t\in I^-. \end{cases}$$

Since $|x_2^+ - x_2^-| \le |x_1^+ - x_1^-|$ and all $|\varepsilon_I| \le 1$, $\psi$ is subordinated to $\varphi$. Moreover, according to hypothesis of the Lemma, the point $x$ is in $\Omega$, whence $x_4 \le Qx_5$, i. e. $[w]_{A_1} \le Q$. Therefore the triple $\varphi$, $\psi$, $w$ is an admissible triple of the test functions corresponding to the point $x$, and

$$\begin{aligned} \mathbf{B}(x) &\ge \frac{1}{|I_0|} w\big(\{t\in I_0\colon \psi(t)\ge 0\}\big) \\ &= \frac{1}{2|I_0^+|} w^+\big(\{t\in I_0^+\colon \psi(t)\ge 0\}\big) + \frac{1}{2|I_0^-|} w^-\big(\{t\in I_0^-\colon \psi(t)\ge 0\}\big) \\ &\ge \frac12 \mathbf{B}(x^+) + \frac12 \mathbf{B}(x^-) - 2\eta. \end{aligned}$$

Since this inequality holds for an arbitrary small $\eta$, we can pass to the limit $\eta\to 0$, what gives us the required assertion. □

3.1.4. *The forth property:* $\mathbf{B}$ *decreases in* $x_5$*.* This is a corollary of the preceding property, i. e. it follows from the main inequality. Indeed if we put in the hypotheses of Lemma 3.1 $x_i^+ = x_i^-$ for $1\le i\le 4$ and $x_5^+ > x_5^-$, then $x_5 = x_5^-$ and inequality 3.6 turn into

$$\mathbf{B}(x_1,x_2,x_3,x_4,x_5^-) - \mathbf{B}(x_1,x_2,x_3,x_4,x_5^+) \ge 0\,, \tag{3.7}$$

just what we need.

3.1.5. *The fifth property: function* $t\mapsto \frac1t \mathcal{B}(x_1, tx_3, tx_4)$ *is increasing.* This is in fact the preceding property rewritten in terms of $\mathcal{B}$. Indeed, if we put $x_2 = -1$ and use (3.5), then (3.7) with $t^+ = \frac{1}{x_5^-}$ and $t^- = \frac{1}{x_5^+}$ gives the required monotonicity.

3.1.6. *The sixth property: function* $\mathcal{B}$ *is concave.* Lemma 3.1 applied to the case $x_2^+ = x_2^- = -1$ and $x_5^+ = x_5^- = 1$ guarantees the stated concavity.

3.1.7. *The seventh property: the symmetry and monotonicity in $x_1$.* It is easy to see from the definition that $\mathbf{B}$, and hence $\mathcal{B}$ as well, is even in its variable $x_1$.

Concavity of $\mathcal{B}$ (in $x_1$) and this symmetry together imply that $x_1 \mapsto B(x_1, x_3, x_4)$ is increasing on $[-x_3, 0]$ and decreasing on $[0, x_3]$.

3.2. **The goal and the idea of the proof.** It would be natural now to solve the corresponding boundary value problem for the Monge–Ampère equation, to find the function $\mathbf{B}$, as it was done in the unweighted case, and then to find the constant we are interested in:

$$C(Q) = \sup\left\{\frac{|x_2|\mathbf{B}(x)}{x_3} : x_2 < 0,\ x_3 \geq |x_1|x_5,\ x_5 \leq x_4 \leq Qx_5\right\}.$$

However for now this task is too difficult for us. So, we use the listed properties of $\mathbf{B}$ to prove the following estimate from below on function $\mathcal{B}$.

**Theorem 3.2.** *If $Q \geq 4$ then*

$$\mathcal{B}(x_1, x_3, x_4) \geq \frac{1}{515} Q(\log Q)^{1/3} x_3 . \tag{3.8}$$

*at some point $(x_1, x_3, x_4) \in G$.*

**Remark 3.3.** It is a subtle result and it will take some space below to prove. Recall that Muckenhoupt conjectured that for the Hilbert transform $H$ and any weight $w \in A_1$ the following two estimates hold on a unit interval $I$:

$$w\{x \in I : |Hf(x)| > \lambda\} \leq \frac{C}{\lambda} \int_I |f| Mw dx\,, \tag{3.9}$$

$$w\{x \in I : |Hf(x)| > \lambda\} \leq \frac{C\,[w]_{A_1}}{\lambda} \int_I |f| w dx\,. \tag{3.10}$$

Obviously if (3.9) holds then (3.10) is valid as well. It took many years to *disprove* (3.9). This was done by Maria Reguera and Christoph Thiele [7], [8]. The constructions involve a very irregular (almost a sum of delta measures) weight $w$, so there was a hope that such an effect cannot appear when the weight is regular in the sense that $w \in A_1$. Theorem 3.2 gives a counterexample to this hope for the case when the Hilbert transform is replaced by the martingale transform on a usual homogeneous dyadic filtration. The reader can consult [5] to see that for the Hilbert transform a counterexample also exists, and so (3.10) fails as well. The counterexample for the Hilbert transform is the transference of a counterexample we build here for the martingale transform. Notice that Theorem 3.2 implicitly gives a certain counterexample for the Hilbert transform. We will explain in a separate note how to make this transference.

Now a couple of words about the idea of the proof of Theorem 3.2. Ideally we would like to find the formula for $\mathcal{B}$ (and therefore for $\mathbf{B}$ because of (3.5)). To proceed we rewrite the third property of $\mathbf{B}$ (see subsection 3.1.3) as a PDE on $\mathcal{B}$. Then, using the boundary conditions on $\mathcal{B}$ on $\partial G$ (the domain $G$ is defined in (3.4)), we may hope to solve this PDE. Unfortunately there are many roadblocks on this path, starting with the fact that the third property of $\mathbf{B}$ is not a PDE, it is rather a partial differential inequality in discrete form. It the not weighted case we pay no attention to this important fact. We simply assume the required smoothness of our function to find a smooth candidate. After such a candidate ws found we have proved that it coincides with the required Bellman function. Now we cannot find a candidate and we will work with the abstractly defined Bellman function whose smoothness is unknown. We will write the inequality in discrete form as a pointwise partial differential inequality, but for that we will need a subtle result of Aleksandrov.

3.3. **From discrete inequality to differential inequality via Aleksandrov's theorem.** As it was mentioned in Subsection 3.1.6 the function $\mathcal{B}$ is concave on its domain of definition $G$. By the result of Aleksandrov, see Theorem 6.9 of [2], $\mathcal{B}$ has all second derivatives almost everywhere in $G$. Second property (homogeneity) of function $\mathbf{B}$ (see (3.5)) implies that the function $\mathbf{B}$ has all second derivatives almost everywhere in $\Omega$.

First, using this fact we rewrite the homogeneity condition (see Subsection 3.1.2) in the following differential form:

$$x_1\mathbf{B}_{x_1} + x_2\mathbf{B}_{x_2} + x_3\mathbf{B}_{x_3} = 0\,; \tag{3.11}$$

$$x_3\mathbf{B}_{x_3} + x_4\mathbf{B}_{x_4} + x_5\mathbf{B}_{x_5} = \mathbf{B}\,. \tag{3.12}$$

These equalities we have got by differentiating (3.3) with respect to $s_1$ and with respect to $s_2$ and taking the result for $s_1 = s_2 = 1$.

Our second step is to replace the main inequality in discrete form by the inequality in the form of a pointwise partial differential inequality. Lemma 3.1 implies that the quadratic form

$$\sum_{i,j=1}^{4} \mathbf{B}_{x_ix_j}\Delta_i\Delta_j \tag{3.13}$$

is non-positive at almost any interior point of $\Omega$ and for all vectors $\Delta \in \mathbb{R}^4$ such that $|\Delta_2| \le |\Delta_1|$.

We consider three partial cases of (3.13) with $\Delta_1 = \Delta_2$, with $\Delta_1 = -\Delta_2$, and with $\Delta_2 = 0$. Moreover, to reduce our investigation to consideration of $2\times 2$ matrices we choose some special relation between $\Delta_3$ and $\Delta_4$. In the first case we consider the quadratic form on the vector $\Delta$ with

$$\Delta_1 = \Delta_2 = -\delta_1, \quad \Delta_3 = x_3(\delta_1+\delta_2), \quad \Delta_4 = x_4\delta_2.$$

In the second case we put

$$\Delta_1 = -\Delta_2 = \delta_1, \quad \Delta_3 = x_3(\delta_1+\delta_2), \quad \Delta_4 = x_4\delta_2.$$

Then we get two quadratic forms

$$\sum_{i,j=1}^{4} \mathbf{B}_{x_ix_j}\Delta_i\Delta_j = \sum_{i,j=1}^{2} K^{\pm}_{ij}\delta_i\delta_j\,,$$

where

$$K^{\pm} = \begin{pmatrix} \begin{matrix} \mathbf{B}_{x_1x_1} \pm \mathbf{B}_{x_1x_2} \mp x_3\mathbf{B}_{x_1x_3} \pm \mathbf{B}_{x_1x_2} \\ + \mathbf{B}_{x_2x_2} - 2x_3\mathbf{B}_{x_2x_3} \mp x_3\mathbf{B}_{x_1x_3} + x_3^2\mathbf{B}_{x_3x_3} \end{matrix} & \begin{matrix} \mp x_3\mathbf{B}_{x_1x_3} - x_3\mathbf{B}_{x_2x_3} + x_3^2\mathbf{B}_{x_3x_3} \\ \mp x_4\mathbf{B}_{x_1x_4} - x_4\mathbf{B}_{x_2x_4} + x_3x_4\mathbf{B}_{x_3x_4} \end{matrix} \\ \\ \begin{matrix} \mp x_3\mathbf{B}_{x_1x_3} - x_3\mathbf{B}_{x_2x_3} + x_3^2\mathbf{B}_{x_3x_3} \\ \mp x_4\mathbf{B}_{x_1x_4} - x_4\mathbf{B}_{x_2x_4} + x_3x_4\mathbf{B}_{x_3x_4} \end{matrix} & x_3^2\mathbf{B}_{x_3x_3} + 2x_3x_4\mathbf{B}_{x_3x_4} + x_4^2\mathbf{B}_{x_4x_4} \end{pmatrix}.$$

These matrices are non-positive and their half sum is the following non-positive matrix

(3.14)
$$\begin{pmatrix} \mathbf{B}_{x_1x_1}+\mathbf{B}_{x_2x_2}-2x_3\mathbf{B}_{x_2x_3}+x_3^2\mathbf{B}_{x_3x_3} & -x_3\mathbf{B}_{x_2x_3}+x_3^2\mathbf{B}_{x_3x_3}-x_4\mathbf{B}_{x_2x_4}+x_3x_4\mathbf{B}_{x_3x_4} \\ -x_3\mathbf{B}_{x_2x_3}+x_3^2\mathbf{B}_{x_3x_3}-x_4\mathbf{B}_{x_2x_4}+x_3x_4\mathbf{B}_{x_3x_4} & x_3^2\mathbf{B}_{x_3x_3}+2x_3x_4\mathbf{B}_{x_3x_4}+x_4^2\mathbf{B}_{x_4x_4} \end{pmatrix}.$$

Before proceed further we rewrite this matrix in terms of the function $\mathcal{B}$. For this aim we have to get rid of the derivatives with respect to $x_2$ in this matrix. We are able to do this by using (3.11):

$$\begin{aligned} -x_2\mathbf{B}_{x_2x_3} &= \mathbf{B}_{x_3} + x_1\mathbf{B}_{x_1x_3} + x_3\mathbf{B}_{x_3x_3}\,; \\ -x_2\mathbf{B}_{x_2x_4} &= x_1\mathbf{B}_{x_1x_4} + x_3\mathbf{B}_{x_3x_4}\,; \\ x_2^2\mathbf{B}_{x_2x_2} &= 2x_1\mathbf{B}_{x_1} + 2x_3\mathbf{B}_{x_3} + x_1^2\mathbf{B}_{x_1x_1} + 2x_1x_3\mathbf{B}_{x_1x_3} + x_3^2\mathbf{B}_{x_3x_3}\,. \end{aligned}$$

Using these expressions at the point $x=(x_1,-1,x_3,x_4,1)$ we can rewrite the matrix (3.14) as follows

(3.15)
$$\begin{pmatrix} (1+x_1^2)\mathcal{B}_{x_1x_1}+2x_1\mathcal{B}_{x_1} & -x_1\mathcal{B}_{x_3}-x_1x_3\mathcal{B}_{x_1x_3}-x_1x_4\mathcal{B}_{x_1x_4} \\ -x_1\mathcal{B}_{x_3}-x_1x_3\mathcal{B}_{x_1x_3}-x_1x_4\mathcal{B}_{x_1x_4} & x_3^2\mathcal{B}_{x_3x_3}+2x_3x_4\mathcal{B}_{x_3x_4}+x_4^2\mathcal{B}_{x_4x_4} \end{pmatrix}.$$

Now we consider the matrix $K^0$, that appears if we take $\Delta_1 = x_1\delta_1$, $\Delta_2=0$, $\Delta_3=x_3\delta_2$, and $\Delta_4=x_4\delta_2$ in our quadratic form

$$\sum_{i,j=1}^4 \mathbf{B}_{x_ix_j}\Delta_i\Delta_j = \sum_{i,j=1}^2 K^0_{ij}\delta_i\delta_j\,.$$

In result we get

$$K^0 = \begin{pmatrix} x_1^2\mathbf{B}_{x_1x_1} & x_1x_3\mathbf{B}_{x_1x_3}+x_1x_4\mathbf{B}_{x_1x_4} \\ x_1x_3\mathbf{B}_{x_1x_3}+x_1x_4\mathbf{B}_{x_1x_4} & x_3^2\mathbf{B}_{x_3x_3}+2x_3x_4\mathbf{B}_{x_3x_4}+x_4^2\mathbf{B}_{x_4x_4} \end{pmatrix}.$$

The same matrix at the point $x=(x_1,-1,x_3,x_4,1)$ is

$$\begin{pmatrix} x_1^2\mathcal{B}_{x_1x_1} & x_1x_3\mathcal{B}_{x_1x_3}+x_1x_4\mathcal{B}_{x_1x_4} \\ x_1x_3\mathcal{B}_{x_1x_3}+x_1x_4\mathcal{B}_{x_1x_4} & x_3^2\mathcal{B}_{x_3x_3}+2x_3x_4\mathcal{B}_{x_3x_4}+x_4^2\mathcal{B}_{x_4x_4} \end{pmatrix}. \tag{3.16}$$

Taking the sum of (3.15) and (3.16) we get the following non-positive matrix

(3.17)
$$\begin{pmatrix} (1+2x_1^2)\mathcal{B}_{x_1x_1}+2x_1\mathcal{B}_{x_1} & -x_3\mathcal{B}_{x_3} \\ -x_3\mathcal{B}_{x_3} & 2(x_3^2\mathcal{B}_{x_3x_3}+2x_3x_4\mathcal{B}_{x_3x_4}+x_4^2\mathcal{B}_{x_4x_4}) \end{pmatrix} \le 0\,.$$

**Definition 3.4.** Consider a subdomain of $G$,

$$G_1 := \{(x_1,x_3,x_4)\in G\colon x_3>2|x_1|,\ 2<x_4<Q\}\,.$$

Fix now $x=(x_1,x_3,x_4)\in G_1$ and a parameter $t\in[1/2,1]$. Consider inequality (3.17) at the point $x^t=(x_1,tx_3,tx_4)$.

Let us introduce a new function $\beta$, which is certain averaging of $\mathcal{B}$, namely, for any $x\in G_1$ we put

$$\beta(x) \stackrel{\text{def}}{=} 2\int_{1/2}^1 \mathcal{B}(x^t)\,dt\,.$$

Notice several simple facts. First of all

$$x_i\beta_{x_i}(x) = 2\int_{1/2}^{1} x_i^t \mathcal{B}_{x_i}(x^t)\,dt, \qquad x_i^2\beta_{x_ix_i} = 2\int_{1/2}^{1} (x_i^t)^2\mathcal{B}_{x_ix_i}(x^t)\,dt\,.$$

For every function $F$ on $G$ we introduce the notation,

$$\gamma_{_F}(x) \stackrel{\text{def}}{=} x_3^2F_{x_3x_3} + 2x_3x_4F_{x_3x_4} + x_4^2F_{x_4}\,,$$

then

$$\gamma_{_\beta}(x) = 2\int_{1/2}^{1}\gamma_{_\mathcal{B}}(x^t)dt\,. \tag{3.18}$$

Now integrate (3.17) on the interval $t \in [1/2, 1]$. The previous simple observations allow us now to rewrite our reduced concavity condition in the form

$$\begin{pmatrix} (1+2x_1^2)\beta_{x_1x_1} + 2x_1\beta_{x_1} & -x_3\beta_{x_3} \\ -x_3\beta_{x_3} & 2\gamma_{_\beta} \end{pmatrix} \le 0\,. \tag{3.19}$$

The reader may wonders why we are so keen to replace (3.17) by a virtually the same (3.19)? The answer is because we can give a very good pointwise estimate on $\gamma_{_\beta}(x)$, $x \in G_1$. Unfortunately we cannot give any pointwise estimate on $\gamma_{_\mathcal{B}}(x)$, $x \in G$.

Our reduced concavity condition (3.19) is equivalent to the assertion that $\gamma_{_\beta} \le 0$ and the determinant of the matrix in (3.19) is non-negative, i. e.,

$$[-\gamma_{_\beta}]\cdot[-(1+2x_1^2)\beta_{x_1x_1} - 2x_1\beta_{x_1}] \ge x_3^2\beta_{x_3}^2\,. \tag{3.20}$$

Let us denote

$$R \stackrel{\text{def}}{=} \sup\frac{\mathcal{B}(x)}{x_3}\,, \qquad x = (x_1, x_3, x_4) \in G\,.$$

Our goal formulated in (3.8) is to prove $R \ge cQ(\log Q)^\varepsilon$. We are still not too close, but notice that automatically $\mathcal{B}(x) \le Rx_3$, $x \in G$.

3.4. **Logarithmic blow-up.** First we find a pointwise estimate on $\gamma_{_\beta}$.

**Lemma 3.5.** *If $x = (x_1, x_3, x_4)$ is such that $|x_1| \le \frac{1}{4}x_3$ and $x_4 \ge 4$, then*

$$-\gamma_{_\beta}(x) \le 8R(|x_1| + \frac{x_3}{x_4})\,.$$

*Proof.* Consider the following functions

$$\rho(t) \stackrel{\text{def}}{=} \mathcal{B}(x^t), \quad x \in G_1, \qquad \text{and} \qquad r(t) \stackrel{\text{def}}{=} \rho(1)t - \rho(t)$$

on the interval $[t_0, 1]$, where $t_0 = \max(\frac{|x_1|}{x_3}, \frac{1}{x_4})$.

Recall that the function $\rho(t)/t$ is increasing (see property five of $\mathbf{B}$ in Section 3.1). Therefore, $\rho(t)/t \le \rho(1)$, i. e. $r(t) \ge 0$. Since $r$ is convex (because $\rho$ is concave) and $r(1) = 0$, $r$ is a decreasing function on $[t_0, 1]$, in particular $r'(1) \le 0$. Let us estimate the maximal value of $r$ in the following way:

$$r(t_0) < \rho(1)t_0 \le Rx_3t_0 < R\big(|x_1| + \frac{x_3}{x_4}\big)\,. \tag{3.21}$$

Under the hypotheses of the Lemma we have $t_0 \le \frac{1}{4}$, and therefore

$$-\int_{1/2}^{1} \rho''(t)\,dt \le \int_{1/2}^{1} r''(t)\,dt \le 4\int_{1/2}^{1} (t-t_0)r''(t)\,dt$$
$$\le 4\int_{t_0}^{1} (t-t_0)r''(t)\,dt = 4r'(1)(1-t_0) - 4r(1) + 4r(t_0)\,.$$

Using estimate (3.21) and the properties of $r$ ($r'(1) \le 0$ and $r(1) = 0$) we get

$$-\int_{1/2}^{1} \rho''(t)\,dt \le 4R(|x_1| + \frac{x_3}{x_4})\,.$$

The equality $\gamma_{\mathcal{B}}(x^t) = t^2\rho''(t)$ implies

$$-\int_{1/2}^{1} \gamma_{\mathcal{B}}(x^t)\,dt \le 4R(|x_1| + \frac{x_3}{x_4})\,.$$

So, by (3.18) this is the stated in the Lemma estimate. □

Now we would like to get an estimate for $\beta_{x_3}$ from below. For this aim we construct a pair of test functions $\varphi$, $\psi$ and a test weight $w$, which supply us with the following estimate for the function $\mathcal{B}$.

**Lemma 3.6.** *If $x = (x_1, x_3, x_4)$ is such that $2x_3 + x_1 \ge 1$, then*

$$\mathcal{B}(x) \ge \frac{2x_4 - 1}{4}\,.$$

*Proof.* Let us take the following test functions on the interval $[0, 1]$

$$\varphi = x_1 + x_3 H_{(0,1)} + (x_3 - x_1)H_{(0,\frac{1}{2})} + (x_3 + x_1)H_{(\frac{1}{2},1)}\,;$$
$$\psi = -1 + x_3 H_{(0,1)} + (x_3 - x_1)H_{(0,\frac{1}{2})} - (x_3 + x_1)H_{(\frac{1}{2},1)}\,;$$
$$w = 1 + 2(x_4 - 1)\chi_{(\frac{1}{4},\frac{3}{4})}\,.$$

The Bellman point corresponding to this triple is $(x_1, -1, x_3, x_4, 1)$. The function $\psi$ on the interval $(\frac{1}{2}, \frac{3}{4})$ has the value $2x_3 + x_1 - 1$, where the weight $w$ is $2x_4 - 1$. Therefore, if $2x_3 + x_1 \ge 1$, then by the definition $\mathcal{B}(x) \ge (2x_4 - 1)/4$. □

**Corollary 3.7.** *If $x_3 + x_1 \ge 1$, then*

$$\beta(x) \ge \frac{3x_4 - 2}{8}\,.$$

*Proof.* If $x_3 + x_1 \ge 1$, then $2tx_3 + x_1 \ge 1$ for all $t \in [\frac{1}{2}, 1]$. And therefore,

$$\beta(x) = 2\int_{1/2}^{1} \mathcal{B}(x^t)\,dt \ge \frac{1}{2}\int_{1/2}^{1} (2tx_4 - 1)dt = \frac{3x_4 - 2}{8}\,.$$

□

**Corollary 3.8.** *If $x_3 + x_1 \ge 1$ and $x_4 \ge 2$, then*

$$\beta(x) \ge \frac{x_4}{4}\,.$$

**Corollary 3.9.** *If $x_4 \ge 2$, then*

$$\beta(x_1, 1, x_4) \ge \frac{x_4}{4}\,.$$

*Proof.* Since the function $\mathbf{B}$ is even in $x_1$, the functions $\mathcal{B}$ and $\beta$ are even as well. Therefore without loss of generality we cann assume that $x_1 \ge 0$. Hence for $x_3 = 1$ the condition $x_3 + x_1 \ge 1$ holds, and we have the required estimate. $\square$

**Corollary 3.10.** *If $x_4 \ge 2$, then there exists an $a = a(x_4) \in (0, 1]$ such that $\beta(0, a, x_4) = \frac{x_4}{8}$.*

*Proof.* Since the function $\beta$ is continuous, the conditions

$$\beta(0, 0, x_4) = 0 \quad \text{and} \quad \beta(0, 1, x_4) \ge \frac{x_4}{4}$$

guarantee the existence of the required $a$. $\square$

**Remark 3.11.** The function $\beta$ is increasing in $x_3$ because it is positive, concave, and defined on an infinite interval $(0, \infty)$. Therefore the root $a$ is unique.

**Lemma 3.12.** *For any $x \in G$ we have*

$$\beta(x) \ge \big(1 - \frac{2|x_1|}{x_3}\big)\beta(0, x_3, x_4)\,. \tag{3.22}$$

*Proof.* Since $\beta$ is even in $x_1$, we can assume $x_1 > 0$. The stated estimate is immediate consequence of the following two facts, $\beta$ is non-negative and concave in $x_1$:

$$\beta(x) \ge \big(1 - \frac{2x_1}{x_3}\big)\beta(0, x_3, x_4) + \frac{2x_1}{x_3}\beta(\frac{x_3}{2}, x_3, x_4)\,.$$

$\square$

**Lemma 3.13.** *Let $a = a(x_4)$ be the function described in Corollary 3.10. If $x = (x_1, x_3, x_4)$ is such that $4x_1 \le x_3 \le a$, $2 \le x_4 \le Q$, then*

$$\beta_{x_3}(x) \ge \max\Big\{\frac{x_4 - 16Rx_3}{16a}, \frac{x_4}{8}\Big\}\,. \tag{3.23}$$

*Proof.* Since $\beta$ is concave with respect to $x_3$, and $\beta_{x_3} \ge 0$ for $x_3 \in (0, a)$ we can write

$$a\beta_{x_3}(x) \ge (a - x_3)\beta_{x_3}(x) \ge \beta(x_1, a, x_4) - \beta(x_1, x_3, x_4)\,.$$

Assuming that $x_1 \ge 0$ we can use Lemma 3.12:

$$\beta(x_1, a, x_4) \ge \big(1 - \frac{2x_1}{a}\big)\beta(0, a, x_4) = \big(1 - \frac{2x_1}{a}\big)\frac{x_4}{8} \ge \frac{1}{16}x_4\,.$$

Together with the general estimate $\beta(x) \le Rx_3$ we obtain

$$\beta_{x_3}(x) \ge \frac{x_4 - 16Rx_3}{16a}\,.$$

To get the second inequality we estimate $\beta_{x_3}(a)$:

$$\beta_{x_3}(a) \ge \frac{\beta(x_1, 1, x_4) - \beta(x_1, a, x_4)}{1 - a} \ge \beta(x_1, 1, x_4) - \beta(x_1, a, x_4)\,.$$

Now we use Corollary 3.9 together with the property of $\beta$ to decrease with respect to $x_1$ for $x_1 > 0$:

$$\beta(x_1, 1, x_4) \geq \frac{x_4}{4} \qquad \text{and} \qquad \beta(x_1, a, x_4) \leq \beta(0, a, x_4) = \frac{x_4}{8}\,.$$

In result we get the required estimate:

$$\beta_{x_3}(x_1, a, x_4) \geq \frac{x_4}{4} - \frac{x_4}{8} = \frac{x_4}{8}\,.$$

□

Let us denote the function on the right hand side of (3.23) by $m$. We can rewrite it in the following form:

$$m(x_3, x_4) = \begin{cases} \frac{x_4 - 16Rx_3}{16a}, & \text{if } \ x_3 \leq \frac{(1-2a)x_4}{16R}\,; \\ \frac{x_4}{8}, & \text{if } \ x_3 \geq \frac{(1-2a)x_4}{16R}\,. \end{cases} \tag{3.24}$$

All preparations are made and we are ready to prove Theorem 3.2.

*Proof of* **Theorem 3.2.** Now we combine Lemmas 3.5 and 3.13 to deduce from (3.20) the following inequality

$$-(1+2x_1^2)\beta_{x_1x_1} - 2x_1\beta_{x_1} \geq \frac{x_3^2(\beta_{x_3})^2}{-\gamma_\beta} \geq \frac{x_3^2 m^2}{8R(|x_1| + \frac{x_3}{x_4})}\,,$$

that holds under assumptions $4|x_1| \leq x_3 \leq a \leq 1$ and $4 \geq x_4 \leq Q$. Dividing both part of this inequality over $\sqrt{1+2x_1^2}$ we can rewrite it in the form

$$-\frac{\partial}{\partial x_1}\Big(\sqrt{1+2x_1^2}\,\beta_{x_1}\Big) \geq \frac{x_3^2 m^2}{8R(|x_1| + \frac{x_3}{x_4})\sqrt{1+2x_1^2}}\,.$$

Integrating this inequality and taking into account that $\beta$ is even in $x_1$ (i. e. $\beta_{x_1}(0, x_3, x_4) = 0$) we get

$$\begin{aligned} -\sqrt{1+2x_1^2}\,\beta_{x_1} &\geq \frac{x_3^2 m^2}{8R}\int_0^{x_1} \frac{dt}{(t + \frac{x_3}{x_4})\sqrt{1+2t^2}} \\ &\geq \frac{x_3^2 m^2}{8R\sqrt{1+2x_1^2}}\int_0^{x_1} \frac{dt}{t + \frac{x_3}{x_4}} = \frac{x_3^2 m^2}{8R\sqrt{1+2x_1^2}} \log\Big(1 + \frac{x_4}{x_3}x_1\Big)\,. \end{aligned}$$

Once more we divide over the square root and integrate in $x_1$:

$$\begin{aligned} \beta(0, x_3, x_4) - \beta(x_1, x_3, x_4) &\geq \frac{x_3^2 m^2}{8R}\int_0^{x_1} \log\Big(1 + \frac{x_4}{x_3}t\Big)\frac{dt}{1+2t^2} \\ &\geq \frac{x_3^2 m^2}{8R(1+2x_1^2)}\int_0^{x_1} \log\Big(1 + \frac{x_4}{x_3}t\Big)\,dt \\ &= \frac{x_3^3 m^2}{8Rx_4(1+2x_1^2)}\Big[\Big(1 + \frac{x_4}{x_3}x_1\Big)\log\Big(1 + \frac{x_4}{x_3}x_1\Big) - \frac{x_4}{x_3}x_1\Big] \\ &\geq \frac{x_3^3 m^2}{9Rx_4}\Big[\Big(1 + \frac{x_4}{x_3}x_1\Big)\log\Big(1 + \frac{x_4}{x_3}x_1\Big) - \frac{x_4}{x_3}x_1\Big]\,. \end{aligned}$$

In the last estimate we use the restriction $4|x_1| \leq x_3 \leq 1$, whence $1+2x_1^2 \leq \frac{9}{8}$.

Now we use inequality (3.22) from Lemma 3.12 and the general inequality $\beta(x) \le Rx_3$:

$$\beta(0,x_3,x_4) - \beta(x_1,x_3,x_4) \le \frac{2x_1}{x_3}\beta(0,x_3,x_4) \le 2x_1 R\,.$$

Combining with the preceding inequality we come to the following estimate

$$\frac{x_3^3 m^2}{18R^2 x_1 x_4}\Big[\Big(1+\frac{x_4}{x_3}x_1\Big)\log\Big(1+\frac{x_4}{x_3}x_1\Big) - \frac{x_4}{x_3}x_1\Big] \le 1\,.$$

Recall that this estimate we obtained in the following domain of variables:

$$0 \le 4x_1 \le x_3 \le a \quad \text{and} \quad 4 \le x_4 \le Q\,.$$

Let us now choose the values of this variables. Since the function $t \mapsto \frac{1+t}{t}\log(1+t)$ monotonously increases, we get the best possible estimate when take the maximal possible value of $x_1$, i. e. $x_1 = \frac14 x_3$:

$$\frac{x_3^2 m^2}{18R^2 x_4}\Big[\Big(1+\frac{x_4}{4}\Big)\log\Big(1+\frac{x_4}{4}\Big) - \frac{x_4}{4}\Big] \le 1\,.$$

Since the behaviour of the function $a(x_4)$ is unknown, we cannot choose the best possible value of $x_4$, we take the largest value $x_4 = Q$:

$$\frac{x_3^2 m^2}{18R^2 Q}\Big[\Big(1+\frac{Q}{4}\Big)\log\Big(1+\frac{Q}{4}\Big) - \frac{Q}{4}\Big] \le 1\,,$$

where, of course, $a = a(Q)$ and $m = m(x_3, Q)$. To simplify this expression we use the following elementary estimate:

$$\big(1+\frac{t}{4}\big)\log\big(1+\frac{t}{4}\big) - \frac{t}{4} \ge \frac{t}{16}\log t \quad \text{for} \quad t \ge 4\,.$$

To check this inequality we consider the function

$$f(t) \stackrel{\text{def}}{=} 16\big(1+\frac{t}{4}\big)\log\big(1+\frac{t}{4}\big) - 4t - t\log t$$

and check that $f(t) \ge 0$ for $t \ge 4$.

$$\begin{aligned}
f(4) &= 32\log 2 - 16 - 4\log 4 = 8\log\frac{8}{e^2} > 0\,;\\
f'(t) &= 4\log\big(1+\frac{t}{4}\big) - \log t - 1\,;\\
f'(4) &= 4\log 2 - \log 4 - 1 = \log\frac{4}{e} > 0\,;\\
f''(t) &= \frac{4}{t+4} - \frac{1}{t} = \frac{3t-4}{t(t+4)} > 0 \quad \text{for} \quad t \ge 4\,.
\end{aligned}$$

In result we get

$$\frac{x_3^2 m^2}{288R^2}\log Q \le 1 \qquad \text{for any} \quad x_3 \in [0,a]\,. \tag{3.25}$$

Now we need to investigate the function $x_3 \mapsto x_3 m(x_3, Q)$ on the interval $[0,a]$. If $a \ge \frac14$ then this function is increasing and takes its maximal value at the point $x_3 = a$, and (3.25) yields

$$\frac{a^2 Q^2}{288R^2 \cdot 8^2}\log Q \le 1\,.$$

or

(3.26)
$$R \geq \frac{a}{96\sqrt{2}} Q\big(\log Q\big)^{1/2} \geq \frac{\big(\log 4\big)^{1/6}}{4\cdot 96\sqrt{2}} Q\big(\log Q\big)^{1/3} \geq \frac{1}{515} Q\big(\log Q\big)^{1/3}.$$

We specially make the exponent of logarithm worth ($\frac{1}{3}$ instead of $\frac{1}{2}$), because we can get only such exponent for other values of the unknown parameter $a$.

From now on we assume that $a < \frac{1}{4}$. In this case the function has a local maximum at the point $x_3 = \frac{Q}{32R}$. Indeed, since $aR \geq \beta(0, a, Q) = \frac{Q}{8}$, we have $a \geq \frac{Q}{8R} > \frac{Q}{32R}$, therefore the point $x_3 = \frac{Q}{32R}$ is in the domain. The value of the function $x_3 m(x_3, Q)$ at this the point is $\frac{Q}{32R} \cdot \frac{Q}{32a}$. On the other hand at the end of the interval for $x_3 = a$ we have the value $am(a, Q) \geq \frac{aQ}{8}$. If $a^2 < \frac{Q}{128R}$ then we use the first estimate:
$$1 \geq \Big(\frac{Q}{32R} \cdot \frac{Q}{32a}\Big)^2 \frac{1}{288R^2} \log Q \geq \frac{Q^4}{9\cdot 2^{25} R^4} \cdot \frac{2^7 R}{Q} \log Q \geq \Big(\frac{Q}{134R}\Big)^3 \log Q\,,$$
or
$$R \geq \frac{1}{134} Q\big(\log Q\big)^{1/3}.$$
In the case if $a^2 \geq \frac{Q}{128R}$ we use the second estimate:
$$1 \geq \Big(\frac{aQ}{8}\Big)^2 \frac{1}{288R^2} \log Q \geq \frac{Q^3}{9\cdot 2^{18} R^3} \log Q \geq \Big(\frac{Q}{134R}\Big)^3 \log Q\,,$$
and again

(3.27)
$$R \geq \frac{1}{134} Q\big(\log Q\big)^{1/3}.$$

Therefore, if $a < \frac{1}{4}$ estimate 3.27 holds.

Comparing the estimates we got for different possible values of the unknown parameter $a$, namely, (3.26) and (3.27) we see that the estimate
$$R \geq \frac{1}{515} Q\big(\log Q\big)^{1/3}$$
is true in all cases. This completes the proof of Theorem 3.2, and therefore the proof of Theorem 1.1.

**Aknowledgement.** This paper was prepared mainly during the stay of V. Vasyunin and A. Volberg in the Mathematisches Forschungsinstitut Oberwolfachthe (the program “Research in Pairs”). We are very grateful for this opportunity and the remarkable work conditions there.

## References

[1] David V. Cruz-Uribe, Jos Maria Martell, Carlos Prez, *Weights, extrapolation and the theory of Rubio de Francia*. Operator Theory: Advances and Applications, Vol. 215. Birkhuser/Springer Basel AG, Basel, 2011. xiv+280 pp.

[2] L. C. Evans, R. F. Gariepy, *Measure Theory and Fine Properties of Functions.*, 1992.

[3] A. Lerner, S. Ombrosi, C. Pérez, *$A_1$ bounds for Calderón–Zygmund operators related to a problem of Muckenhoupt and Wheeden*, Math. Res. Lett., **16** (2009) no. 1, 149–156.

[4] F. Nazarov, A. Reznikov, V. Vasyunin, A. Volberg, *A1 conjecture: weak norm estimates of weighted singular operators and Bellman function*, http://sashavolberg.files.wordpress.com/2010/11/a11_7loghilb11_21_2010.pdf.

[5] F. Nazarov, A. Reznikov, V. Vasyunin, A. Volberg, *A Bellman function counterexample to the $A_1$ conjecture: the blow-up of weak norm estimates of weighted singular operators*, arXiv:1506.04710v1, pp. 1–23.

[6] A. Osękowski, *Some sharp estimate for the Haar system and other bases in $L^2(0,1)$*, Mathematica Scandinavica **115** (2014), pp. 123–142.

[7] M. C. Reguera, *On Muckenhoupt–Wheeden conjecture*, Adv. Math. **227** (2011), no. 4, 1436–1450.

[8] M. C. Reguera, C. Thiele, *The Hilbert transform does not map $L^1(Mw)$ to $L^{1,\infty}(w)$*, arXiv:1011.1767, Math. Res. Lett. **19** (2012), no. 1, 1–7.

[9] E. M. Stein and J. N. Weiss, *On the convergence of Poisson integrals*, Trans. AMS **140** (1960), 39–54.

Department of Mathematics, Kent Sate University, USA
*E-mail address*: `nazarov@math.kent.edu` (F. Nazarov)

Department of Mathematics, Michigan Sate University, East Lansing, MI. 48823
*E-mail address*: `reznikov@math.msu.edu` (A. Reznikov)

V. A. Steklov Math. Inst., Fontanka 27, St. Petersburg, Russia
*E-mail address*: `vasyunin@pdmi.ras.ru` (V. Vasyunin)

Department of Mathematics, Michigan Sate University, East Lansing, MI. 48823
*E-mail address*: `volberg@math.msu.edu` (A. Volberg)