Fefferman–Stein inequalities for the dyadic-like maximal operators

MATEUSZ RAPICKI

Abstract. The paper contains a transference theorem which allows to extend a large class of unweighted inequalities for the dyadic maximal operator to their weighted Fefferman–Stein counterparts on general probability spaces.

Mathematics Subject Classification. Primary 42B25, Secondary 60G42.

Keywords. Maximal operator, Weighted inequality, Fefferman–Stein inequality, Martingale, Bellman function method.

1. Introduction. The results of this paper are motivated by certain celebrated bounds for the dyadic maximal operator on $\mathbb{R}^d$. Recall that this operator is given by

$$M^d f(x) = \sup \left\{ \frac{1}{|Q|} \int_Q |f(y)| \, dy : Q \subset \mathbb{R}^d \text{ is a dyadic cube containing } x \right\},$$

where $f: \mathbb{R}^d \to \mathbb{R}$ is locally integrable and the lattice of dyadic cubes in $\mathbb{R}^d$ is formed by the grids $(2^{-N}\mathbb{Z}^d)_{N=0,1,2,...}$. This operator plays a fundamental role in PDEs and harmonic analysis, and the problem of obtaining sharp, or at least good bounds for its norms is of considerable interest. For example, $M^d$ satisfies the weak-type bound

$$\lambda |\{ x \in \mathbb{R}^d : M^d f(x) > \lambda \}| \leq \int_{\{M^d f > \lambda\}} |f(x)| \, dx$$

(1)

for any $f \in L^1(\mathbb{R}^d)$ and $\lambda > 0$, and it also enjoys the Hardy–Littlewood–Doob inequality

$$\|M^d f\|_{L^p(\mathbb{R}^d)} \leq \frac{p}{p-1} \|f\|_{L^p(\mathbb{R}^d)}, \quad 1 < p \leq \infty,$$

(2)
for any \( f \in L^p(\mathbb{R}^d) \). Both of the above estimates are sharp, regardless of the dimension \( d \). These two classical results are of fundamental significance and constitute a starting point for further research. The literature on this topic is very large and we will only quote some statements which are closely connected to the subject of the present paper.

First, note that the inequalities (1) and (2) can be studied in the localized setting. If we restrict ourselves to functions supported on \([0; 1]^d\) and in the definition of \( M^d \) we will consider only the dyadic cubes \( Q \) contained within \([0; 1]^d\), then the obtained local maximal operator \( M^d_{\text{loc}} \) satisfies

\[
\lambda |\{ x \in [0; 1]^d : M^d_{\text{loc}} f(x) > \lambda \}| \leq \int_{\{ M^d_{\text{loc}} f > \lambda \}} |f(x)| \, dx, \tag{3}
\]
and

\[
||M^d_{\text{loc}} f||_{L^p([0;1]^d)} \leq \frac{p}{p-1} ||f||_{L^p([0;1]^d)}, \quad 1 < p \leq \infty. \tag{4}
\]

Actually, these two inequalities are equivalent to (1) and (2), which can be easily proven with the use of standard dilation and translation arguments.

We will be interested in extensions of the inequalities (1) and (2) to the weighted setting. Here and below, the word ‘weight’ refers to a locally integrable, nonnegative function on the base space (e.g., \( \mathbb{R}^d \) or \([0; 1]^d\), this will be clear from the context). It follows from the works of Fefferman and Stein [3] that for any \( f \in L^1(\mathbb{R}^d) \), any \( \lambda > 0 \), and any weight \( w \), we have

\[
\lambda w(\{ x \in \mathbb{R}^d : M^d f(x) > \lambda \}) \leq \int_{\{ M^d f > \lambda \}} |f(x)| M^d w(x) \, dx,
\]
where we have used the notation \( w(A) = \int_A w \, dx \). This, by a straightforward interpolation argument, yields the weighted \( L^p \)-bound

\[
||M^d f||_{L^p(w)} \leq \frac{p}{p-1} ||f||_{L^p(M^d w)}, \quad 1 < p \leq \infty, \tag{5}
\]
where \( ||f||_{L^p(w)} := (\int_{\mathbb{R}^d} |f|^p w \, dx)^{1/p} \). The same weighted inequalities hold in the localized setting. Consult [13] for some dual statements.

The inequalities (1) and (2) (or rather their local versions (3) and (4)) can also be extended to a more general setting of maximal operators associated with a given tree structure (see, e.g., [6,7,10]). Let us provide the necessary definition.

**Definition 1.** (Tree) Assume that \((X, \mathcal{F}, \mu)\) is a nonatomic probability space. A set \( \mathcal{T} \subseteq \mathcal{F} \) is called a tree if the following conditions are satisfied:

(i) For every \( I \in \mathcal{T} \) we have \( \mu(I) > 0 \).

(ii) For every \( I \in \mathcal{T} \) we have a finite subset \( C(I) \subseteq \mathcal{T} \) such that \( I = \bigcup C(I) \) and for \( A, B \in C(I) \) we have \( \mu(A \cap B) = 0 \).

(iii) \( \mathcal{T} = \bigcup_{n \in \mathbb{N}} \mathcal{T}^n \), where \( \mathcal{T}^0 = \{ X \} \) and \( \mathcal{T}^{n+1} = \bigcup_{I \in \mathcal{T}^n} C(I) \).

(iv) We have \( \lim_{n \to \infty} \sup_{I \in \mathcal{T}^n} \mu(I) = 0 \).
The corresponding maximal operator $M^T_T$ is given by
\[ M^T_T f(\omega) := \sup_{I \in \mathcal{T}, \omega \in I} \frac{1}{\mu(I)} \int_I |f| \, d\mu. \]

Such operator will be called the dyadic-like maximal operator associated with $T$. Observe that the cube $[0; 1]^d$ with its Borel subsets and the Lebesgue measure is a probability space, and the family of dyadic cubes contained in $[0; 1]^d$ forms a tree; furthermore, the associated operator coincides with the local dyadic maximal operator $M^d_{loc}$. Thus the above probabilistic context does generalize the dyadic setting considered previously.

The principal purpose of this paper is to establish a general transference theorem which enables us to pass from unweighted estimates for the local dyadic operator $M^1_{loc}$ on the interval $[0; 1]$ to their Fefferman–Stein counterparts in the context of trees. The precise formulation is given below.

**Theorem 1.** Let $\Phi, \Psi: [0; \infty) \to [0; \infty)$ be nondecreasing functions such that $\Phi$ is convex and $\Psi$ is left-continuous. Suppose in addition that the inequality
\[ ||\Psi(M^1_{loc} f)||_{L^1(0;1)} \leq ||\Phi(|f|)||_{L^1(0;1)} \] (6)

holds for all integrable functions $f: [0; 1] \to \mathbb{R}$. Then for any probability space $(X, \mathcal{F}, \mu)$ equipped with the tree structure $\mathcal{T}$, any weight $w$, and any integrable function $f$, we have
\[ ||\Psi(M^T_T f)||_{L^1(w)} \leq ||\Phi(|f|)||_{L^1(M^T_T w)}. \] (7)

Thus, any ‘integral’ inequality which is valid in the very special one-dimensional dyadic setting automatically extends to the general weighted context. A very nice feature is that the ‘shape’ of the inequality is preserved, i.e., no additional multiplicative constants appear. This in particular implies that if the starting inequality (6) is sharp, then so is the weighted version (7).

Let us mention several applications of the above statement; the theorem covers some well known bounds for the maximal operator.

(i) If we take $\Psi(x) = \lambda 1_{(\lambda; \infty)}(x)$ and $\Phi(x) = x$ for a given $\lambda > 0$, then (6) becomes the unweighted weak-type bound (3), and the assertion (7) is precisely the aforementioned result of Fefferman and Stein, in the context of trees.

(ii) The choice $\Psi(x) = x^p$ and $\Phi(x) = (p/(p - 1))^{p}x^p$ for a given $p > 1$ leads to the tree version of the weighted $L^p$-bound (5).

(iii) Let $K > 1$ be a fixed parameter. Setting $\Psi(x) = x$ and $\Phi(x) = Kx \log x + K^2/((K - 1)e)$ corresponds to the sharp $L \log L$ bound
\[ ||M^T_T f||_{L^1(w)} \leq K||f \log f||_{L^1(M^T_T w)} + \frac{K^2}{(K - 1)e} \int_X M^T_T w \, d\mu, \]

proved by Gilat [4] in the unweighted case and by Osekowski [11] in the general case.
The geometric maximal operator associated with the tree $T$ is given by
\[ G_T f(\omega) = \sup_{I \in T, \omega \in I} \exp \left( \frac{1}{\mu(I)} \int_I \log |f| \, d\mu \right). \]

It is (see, e.g., [5]) that for $p > 0$ we have $||G_T f||_{L^p([0;1])} \leq e^{1/p} ||f||_{L^p([0;1])}$; we will prove the weighted analogue of this result (cf. [12]). If we apply the bound to $e^{|f|}$, we get (6) with $\Psi(x) = \exp(px)$ and $\Phi(x) = \exp(1 + px)$, so Theorem 1 yields
\[ \int_X \exp(pM_T f)w \, d\mu \leq \int_X \exp(1 + p|f|)M_T w \, d\mu. \]

In particular, applying this bound to $\log |f|$, we obtain the estimate
\[ \int_X (G_T f)^p w \, d\mu \leq \int_X \exp(pM_T (\log |f|))w \, d\mu \leq \int_X \exp(1 + p|\log |f||)M_T w \, d\mu, \]

which is a bit worse than the desired weighted $L^p$ bound for $G_T$. To overcome this problem, assume first that $\log |f|$ is bounded from below by some constant $M$ and replace $f$ by $f e^{-M}$ in the latter estimate. As the result, we get an inequality equivalent to
\[ \int_X (G_T f)^p w \, d\mu \leq e \int_X |f|^p M_T w \, d\mu, \]

as needed. The general case follows by standard limiting arguments: we use the latter bound for $|f| \vee e^{-M_1}$ and the weight $w \wedge M_2$, and finally let $M_1 \to \infty$ and then $M_2 \to \infty$.

Let us say a few words about the proof of our main result. Our argumentation will depend heavily on the so-called Bellman function method, a powerful tool which is used widely to establish various estimates in analysis and probability (cf. [1,2,8,9,14]). Generally speaking, this approach enables to transfer the problem of proving a given inequality to that of constructing a certain special function which enjoys appropriate size and convexity constraints. Roughly speaking, we will proceed as follows. First we will show that the validity of (6) implies the existence of a certain abstract special function of two variables. Then, we will modify this object appropriately to ensure certain additional properties, and thus we will obtain a function of four variables. This special object will allow us to deduce (7).

2. Proof of Theorem 1. For the sake of convenience and clarity, we split this section into a few separate parts.
2.1. Basic definitions, probabilistic interpretation. Let \((X, \mathcal{F}, \mu)\) be a given probability space equipped with some fixed tree \(T\). Given a nonnegative, integrable function \(f\) (random variable) on \(X\), we can define a functional sequence \((f_n)_{n \in \mathbb{N}}\) by the formula
\[
f_n(\omega) := \frac{1}{\mu(I)} \int_I f \, d\mu, \quad \text{where } I \text{ is the element of } T^n, \text{ which contains } \omega.
\]
Note that for almost all \(\omega \in X\) the set \(I\) in the above definition is uniquely determined and hence \(f_n\) is well defined almost everywhere. In the probabilistic language the sequence \((f_n)_{n \in \mathbb{N}}\) is the martingale generated from \(f\) by the filtration \((\mathcal{F}_n)_{n \in \mathbb{N}}\) induced by the tree \(T: \mathcal{F}_n = \sigma(T^n)\). It is easy to see that \(M^T f\) coincides with \(\sup_{n \in \mathbb{N}} f_n\), the maximal function of the martingale induced by \(f\). We will also need the truncated operators \((M^T_n)_{n \in \mathbb{N}}\) given by the formula
\[
M^T_n f(\omega) := \max \left\{ \frac{1}{\mu(I)} \int_I f \, d\mu: \omega \in I \in T^k, \ 0 \leq k \leq n \right\},
\]
which in the probabilistic context become the truncated maximal functions of the martingale generated by \(f\): \(M^T_n f = \max_{0 \leq k \leq n} f_k\). It is easy to see that for any function \(f\) and any nonnegative integers \(m\) and \(n\) we have the identities \(M_m f_n = M f_{m \wedge n}\) and \(M_n f = M f_n\); we will use them in some considerations below. Finally, let \((df_n)_{n \in \mathbb{N}}\) denote the difference sequence of the martingale \((f_n)_{n \in \mathbb{N}}\), given by \(df_0 = f_0\) and \(df_n = f_n - f_{n-1}, n \geq 1\).

2.2. A special function for the unweighted inequality. Our starting point is to associate a certain special abstract function with the assumed estimate (6). Throughout this subsection, we assume that the underlying probability space is the interval \([0; 1]\) with its Borel subsets and Lebesgue’s measure, equipped with the dyadic tree. Furthermore, for notational simplicity, the associated (truncated and non-truncated) maximal operators will be denoted by \(M_n\) and \(M\), respectively.

Consider the angular domain \(D = \{(x, y) \in \mathbb{R}^2: 0 \leq x \leq y\}\) and introduce the abstract function \(U: D \to \mathbb{R} \cup \{\infty\}\) by the formula
\[
U(x, y) = \sup \left\{ \int_0^1 (\Psi(M_n f \vee y) - \Phi(f_n)) \, du \right\}
\]
\[
= \sup \left\{ \int_0^1 \left( \Psi \left( \max_{0 \leq k \leq n} f_k \vee y \right) - \Phi(f_n) \right) \, du \right\}.
\]
Here the supremum is taken over all nonnegative integers \(n\) and all Lebesgue integrable functions \(f: [0; 1] \to [0; \infty)\) satisfying \(\int_0^1 f = x\). Note that for any \(k\) the random variable \(f_k\) takes only a finite number of values (this is due to the finiteness of \(T^k\), see the second condition in the definition of a tree); this guarantees that there is no problem with the existence of the integrals under the above suprema.
The function $U$ enjoys certain important structural properties, as we will prove now.

**Lemma 2.**

(i) We have $U(x, x) \leq 0$ for any $x \geq 0$.

(ii) For any $(x, y) \in D$ we have $U(x, y) \geq \Psi(y) - \Phi(x)$.

(iii) The function $U$ satisfies the following mid-concavity property: for any $(x, y) \in D$ and any $t \in [0; x]$ we have

$$U(x, y) \geq \frac{1}{2}U(x - t, y) + \frac{1}{2}U(x + t, (x + t) \vee y).$$

**Proof.** To show (i), note that for any $f$ as in the definition of $U(x, x)$ we have

$$M_n f \geq \int_0^1 f \leq x$$

and hence

$$\int_0^1 (\Psi(M_n f \vee x) - \Phi(f_n))du = \int_0^1 (\Psi(M_n f) - \Phi(f_n))du \leq 0.$$ 

Here the last inequality is due to (6) and the identity $M_n f = Mf_n$. The majorization (ii) follows from considering the constant martingale $f \equiv x$ in the definition of $U(x, y)$. Let us turn our attention to (iii). Take two nonnegative integers $k$ and $m$ and two nonnegative integrable functions $f^{(1)}$ and $f^{(2)}$ on $[0; 1]$ satisfying

$$\int_0^1 f^{(1)} = x - t$$

and

$$\int_0^1 f^{(2)} = x + t.$$ 

Let us splice these two functions into one, by the formula

$$f(\omega) = \begin{cases} 
  f_k^{(1)}(2\omega) & \text{if } 0 \leq \omega < 1/2, \\
  f_m^{(2)}(2\omega - 1) & \text{if } 1/2 \leq \omega \leq 1.
\end{cases}$$

Directly from this construction, we see that

$$\int_0^1 f = \frac{1}{2} \int_0^1 f^{(1)} + \frac{1}{2} \int_0^1 f^{(2)} = \frac{1}{2}(x - t) + \frac{1}{2}(x + t) = x.$$ 

Furthermore, by the self-similarity of the dyadic lattice, for any $n \geq 1$ we have

$$f_n(\omega) = \begin{cases} 
  f^{(1)}_{(n-1) \wedge k}(2\omega) & \text{if } 0 \leq \omega < 1/2, \\
  f^{(2)}_{(n-1) \wedge m}(2\omega - 1) & \text{if } 1/2 \leq \omega \leq 1
\end{cases}$$

and therefore,

$$M_n f(\omega) = \begin{cases} 
  M_{(n-1) \wedge k} f^{(1)}(2\omega) \vee x & \text{if } 0 \leq \omega < 1/2, \\
  M_{(n-1) \wedge m} f^{(2)}(2\omega - 1) \vee x & \text{if } 1/2 \leq \omega \leq 1.
\end{cases}$$

Consequently, if $x \leq y$ and $n \geq \max\{k, m\} + 1$, then

$$\int_0^1 (\Psi(M_n f \vee y) - \Phi(f_n))du = \frac{1}{2} \int_0^1 (\Psi(M_k f^{(1)} \vee y) - \Phi(f_k^{(1)}))du$$

$$+ \frac{1}{2} \int_0^1 (\Psi(M_m f^{(2)} \vee y) - \Phi(f_m^{(2)}))du.$$
Consequently, by the very definition of $U$, we obtain

$$U(x, y) \geq \frac{1}{2} \int_0^1 \left( \Psi(M_k f^{(1)}(1) \vee y) - \Phi(f_k^{(1)}) \right) du$$

$$+ \frac{1}{2} \int_0^1 \left( \Psi(M_m f^{(2)}(2) \vee y) - \Phi(f_m^{(2)}) \right) du,$$

so taking the supremum over all functions $f^{(1)}$ and $f^{(2)}$ and all integers $k$ and $m$, we get the desired concavity. □

Properties (i) and (iii) imply that $U$ is finite on $D$. We will also need the following technical fact.

**Lemma 3.** For any $y > 0$ we have $\lim_{x \uparrow y} U(x, y) = U(y, y)$ and $\lim_{z \downarrow y} U(z, z) = U(y, y)$.

**Proof.** Let $f$ be an arbitrary nonnegative integrable function $f$ on $[0; 1]$ such that \( \int_0^1 f = y \). For any $z > y$, the function $f + (z - y)$ has integral $z$, so by the definition of $U(z, z)$,

$$U(z, z) \geq \int_0^1 (\Psi(M_n(f + (z - y))) \vee z) - \Phi(f_n + z - y)) du$$

$$= \int_0^1 (\Psi(M_n(f + (z - y))) - \Phi(f_n + z - y)) du$$

$$\geq \int_0^1 (\Psi(M_nf) - \Phi(f_n + z - y)) du,$$

where in the last line we have exploited the monotonicity of $\Psi$. Now, the function $f_n$ takes a finite number of values, so using the convexity of $\Phi$, the above estimate gives

$$\liminf_{z \downarrow y} U(z, z) \geq \int_0^1 (\Psi(M_nf) - \Phi(f_n)) du.$$  

Taking the supremum over all $f$ and $n$, we get that

$$\liminf_{z \downarrow y} U(z, z) \geq U(y, y).$$  

(8)

Next, observe that the condition (iii) implies in particular that for any $y > 0$, the function $x \mapsto U(x, y)$ is midpoint-concave on $[0, y]$. Since this function is bounded from below (by (ii)), it must be merely concave on $[0, y]$ and consequently,

$$\lim_{x \downarrow y} U(x, y) \geq U(y, y).$$  

(9)
However, applying (iii) with \( x = y \), we get
\[
U(y, y) \geq \frac{1}{2} U(y - t, y) + \frac{1}{2} U(y + t, y + t),
\]
and letting \( t \to 0 \) gives the claim, in the light of (8) and (9). \( \square \)

Now we will modify the function \( U \) slightly, to ensure certain additional properties.

**Lemma 4.** The function \( \tilde{U} : D \to \mathbb{R} \) given by \( \tilde{U}(x, y) = \min\{U(x, y), \Psi(y)\} \) satisfies (i), (ii), and the following two conditions.

(iii) For any \( (x, y) \in D, x \neq 0 \), there exists a number \( A(x, y) \) such that for any \( s \geq -x \) we have
\[
\tilde{U}(x + s, (x + s) \vee y) \leq \tilde{U}(x, y) + A(x, y)s. \tag{10}
\]

(iv) For any \( x, y \geq 0 \) we have \( \tilde{U}(x, x \vee y) \leq \Psi(y) \).

Note that (iii') is an extension of the concavity condition (iii), which can be seen by applying (10) to \( s = -t, s = t \) and summing the obtained estimates.

**Proof.** The conditions (i) and (ii) are obviously preserved. The condition (iv) is also straightforward: if \( x \geq y \), then \( \tilde{U}(x, x \vee y) \leq U(x, x) \leq 0 \leq \Psi(y) \); on the other hand, if \( x < y \), then the inequality follows directly from the definition of \( \tilde{U} \). Thus, it remains to handle the property (iii').

As we have seen in the proof of the previous lemma, for any \( y \geq 0 \) the function \( x \mapsto U(x, y) \) is concave on \([0; y]\) and continuous on \((0; y]\); it is clear that the function \( \tilde{U} \) inherits both these properties. Set \( A(x, y) \) to be the corresponding left-sided derivative \( \tilde{U}_x(x-, y) \). Then the inequality (10) is obviously satisfied if \( x + s \leq y \); on the other hand, if \( x + s > y \), then we write
\[
\tilde{U}(x, y) + A(x, y)s = \tilde{U}(x, y) + A(x, y)(y - x) + A(x, y)(x + s - y) \geq \tilde{U}(y, y) + A(y, y)(x + s - y).
\]

Here in the last passage we have used (10) with \( s = y - x \) (which we have just proved) and the inequality \( A(x, y) \geq A(y, y) \) which is due to the concavity of \( x \mapsto \tilde{U}(x, y) \). However, directly from the definition of \( \tilde{U} \), the functions \( U \) and \( \tilde{U} \) coincide on the diagonal \( x = y \) and the left-sided derivative \( A(y, y) = \tilde{U}_x(y-, y) \) is not smaller than the left-sided derivative \( U_x(y-, y) \), so it is enough to prove that
\[
U(x + s, x + s) \leq U(y, y) + U_x(y-, y)(x + s - y).
\]
Suppose that the inequality is not satisfied: for some positive \( \kappa \), we have

\[
U(x + s, x + s) \geq U(y, y) + U_x(y-, y)(x + s - y) + \kappa. \tag{11}
\]

An application of (iii) with \( x = y := (x + s + y)/2 \) and \( t = (x + s - y)/2 \) gives
\[
U\left(\frac{x + s + y}{2}, \frac{x + s + y}{2}\right) \geq \frac{1}{2} U\left(y, \frac{x + s + y}{2}\right) + \frac{1}{2} U(x + s, x + s).
\]
Directly from the definition of $U$ and the monotonicity of $\Psi$, we see that $U(x, y)$ increases as $y$ increases. Consequently, we have $U(y, (x + s + y)/2) \geq U(y, y)$, which combined with the previous estimate and (11) gives

$$U \left( \frac{x + s + y}{2}, \frac{x + s + y}{2} \right) \geq U(y, y) + U_x(y-, y) \left( \frac{x + s + y}{2} - y \right) + \kappa/2.$$ 

Note that this estimate is of the same form as (11) and we may iterate the above reasoning to obtain

$$U \left( y + \frac{x + s - y}{2^n}, y + \frac{x + s - y}{2^n} \right) \geq U(y, y) + U_x(y-, y) \left( \frac{x + s - y}{2^n} \right) + \kappa/2^n$$

for any nonnegative integer $n$. Now, if $n$ is chosen sufficiently large, then $(x + s - y)/2^n \leq y$ and yet another application of (iii) gives

$$U(y, y) \geq \frac{1}{2} U \left( y - \frac{x + s - y}{2^n}, y \right) + \frac{1}{2} U \left( y + \frac{x + s - y}{2^n}, y + \frac{x + s - y}{2^n} \right),$$

which combined with the previous bound implies

$$U(y, y) - U \left( y - \frac{x + s - y}{2^n}, y \right) \geq U_x(y-, y) \left( \frac{x + s - y}{2^n} \right) + \kappa/2^n.$$ 

Dividing by $(x + s - y)/2^n$ and letting $n \to \infty$ yields $U_x(y-, y) \geq U_x(y-, y) + \kappa/(x + s - y)$, a contradiction to (11). \hfill $\Box$

### 2.3. A special function for the weighted inequality

Now we will complicate the function $\tilde{U}$ constructed above to obtain the Bellman function corresponding to (7). Let $D_0 = \{ (x, y, u, v) \in \mathbb{R}^4 : 0 \leq x \leq y, 0 \leq u \leq v \}$ and let $B : D_0 \to \mathbb{R}$ be given by

$$B(x, y, u, v) = (u - v) \Psi(y) + v \tilde{U}(x, y).$$

Observe that the properties (i) and (ii) enjoyed by $\tilde{U}$ imply

(i) $B(x, x, u, u) = u \tilde{U}(x, x) \leq 0$.

(ii) $B(x, y, u, v) \geq u \Psi(y) - v \Phi(x)$.

We will also need a property analogous to (iii').

**Lemma 5.** The function $B$ satisfies the following concavity-type property. Pick an arbitrary point $(x, y, u, v) \in D_0$ and two numbers $s \geq -x$, $t \geq -u$. If $x \neq 0$, then

$$B(x + s, y \vee (x + s), u + t, v \vee (u + t)) \leq B(x, y, u, v) + svA(x, y) + t\Psi(y)$$

where $A$ is the function guaranteed by Lemma 4. If $x = 0$, then

$$B(x, y, u + t, v \vee (u + t)) \leq B(x, y, u, v) + t\Psi(y).$$
Proof. The second estimate is straightforward. Indeed, if \( u + t \leq v \), then both sides are equal, while for \( u + t > v \) we use the bound \( \bar{U}(0, y) \leq \Psi(y) \) (the property (iv)) to get

\[
B(x, y, u + t, v \lor (u + t)) = (u + t)\bar{U}(x, y) \leq (u + t - v)\Psi(y) + v\bar{U}(x, y)
= B(x, y, u, v) + t\Psi(y).
\]

To prove the first inequality of the lemma, we consider three cases.

(1) Assume \( x + s \leq y \) and \( u + t \leq v \). From the property (iii') we have

\[
\bar{U}(x + s, y) \leq \bar{U}(x, y) + sA(x, y)
\]

and therefore

\[
B(x + s, y \lor (x + s), u + t, v \lor (u + t)) = B(x + s, y, u + t, v)
= (u + t - v)\Psi(y) + v\bar{U}(x + s, y)
\leq (u + t - v)\Psi(y) + v(\bar{U}(x, y) + sA(x, y))
= B(x, y, u, v) + svA(x, y) + t\Psi(y).
\]

(2) If \( x + s > y \) and \( u + t \leq v \), then the property (iii') yields \( \bar{U}(x + s, x + s) \leq \bar{U}(x, y) + sA(x, y) \). Moreover, notice that \( \Psi(y) \leq \Psi(x + s) \), since \( \Psi \) is nondecreasing. Therefore

\[
B(x + s, y \lor (x + s), u + t, v \lor (u + t)) = B(x + s, x + s, u + t, v)
= (u + t - v)\Psi(x + s) + v\bar{U}(x + s, x + s)
\leq (u + t - v)\Psi(y) + v(\bar{U}(x, y) + sA(x, y))
= B(x, y, u, v) + svA(x, y) + t\Psi(y).
\]

(3) Finally, suppose that \( u + t > v \). Then using the property (iii') and the inequality \( \Psi(y) \geq \bar{U}(x + s, y \lor (x + s)) \) (guaranteed by the property (iv)), we have

\[
B(x, y, u, v) + svA(x, y) + t\Psi(y)
= (u + t - v)\Psi(y) + v(\bar{U}(x, y) + sA(x, y))
\geq (u + t - v)\Psi(y) + v\bar{U}(x + s, y \lor (x + s))
\geq (u + t)\bar{U}(x + s, y \lor (x + s))
= B(x + s, y \lor (x + s), u + t, v \lor (u + t))
\]

This completes the proof. \(\square\)

2.4. Proof of Theorem 1. Fix an arbitrary probability space \( (X, \mathcal{F}, \mu) \) equipped with some tree structure \( T \), and let \( f \) be an integrable function. We may assume that \( f \) is nonnegative, since (7) depends on \( f \) through \( |f| \). Furthermore, with no loss of generality we may assume that \( \int_X \Phi(f)M^2 \, d\mu < \infty \), since otherwise there is nothing to prove. Fix a nonnegative integer \( n \) and an element
The functions $f_n, M_n^T f, w_n,$ and $M_n w$ are constant on $I$; denote the corresponding values by $x, y, u,$ and $v$. If $x \neq 0$, we apply the first inequality of Lemma 5 (with $s = df_{n+1}$ and $t = dw_{n+1}$) to obtain that the above integral does not exceed
\[
\int I \left\{ B(x, y, u, v) + df_{n+1} \cdot u A(x, y) + dw_{n+1} \Psi(y) \right\} d\mu
\]
\[
= \int I B(f_n, M_n^T f, w_n, M_n^T w) d\mu.
\]

On the other hand, if $x = 0$, then we necessarily have $df_{n+1} = 0$ (if a non-negative martingale gets to zero, it stops ultimately) and hence the second inequality of Lemma 5 implies that (12) does not exceed
\[
\int I B(x, y, u, v) + dw_{n+1} \Psi(y) d\mu = \int I B(f_n, M_n^T f, w_n, M_n^T w) d\mu.
\]

Summing over all $I \in T^n$ we obtain that the sequence $(\int X B(f_n, M_n^T f, w_n, M_n^T w) d\mu)_{n \in \mathbb{N}}$ is nonincreasing. Using properties (ii) and (i), conclude that
\[
\int X (\Psi(M_n^T f)w_n - \Phi(f_n)M_n^T w) d\mu \leq \int X B(f_n, M_n^T f, w_n, M_n^T w) d\mu
\]
\[
\leq \int X B(f_0, f_0, w_0, w_0) d\mu \leq 0.
\]

We obtain the assertion by letting $n \to \infty$ and using standard limiting theorems. More precisely, we have $\int X \Psi(M_n^T f)w_n d\mu = \int X \Psi(M_n^T f)wd\mu \to \int X \lim_{n \to \infty} \Psi(M_n^T f)wd\mu$ by the Lebesgue monotone convergence theorem. Then we use left-continuity of $\Psi$ and the fact that $M_n^T f \nearrow M^T f$ to obtain $\int X \Psi(M_n^T f)wd\mu = \int X \Psi(M^T f)wd\mu$. On the right-hand side we use Jensen’s inequality: $\int X \Phi(f_n)M_n^T wd\mu \leq \int X \Phi(f)M_n^T wd\mu \leq \int X \Phi(f)M^T wd\mu$.

**Acknowledgements.** Funding was provided by Narodowe Centrum Nauki (Grant No. DEC-2014/14/E/ST1/00532).

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the
source, provide a link to the Creative Commons license, and indicate if changes
were made.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdic-
tional claims in published maps and institutional affiliations.

**References**

[1] Bellman, R.: Dynamic Programming. Princeton University Press, Princeton
(1957)

[2] Burkholder, D.L.: Explorations in Martingale Theory and Its Applications. École
d’Été de Probabilités de Saint-Flour XIX–1989, pp. 1–66, Lecture Notes in Math-
ematics, vol. 1464. Springer, Berlin (1991)

[3] Fefferman, C., Stein, E.M.: Some maximal inequalities. Am. J. Math. **93**, 107–
115 (1971)

[4] Gilat, D.: The best bound in the LlogL inequality of Hardy and Littlewood and
its martingale counterpart. Proc. Am. Math. Soc. **97**(3), 429–436 (1986)

[5] Hytönen, T., Pérez, C.: Sharp weighted bounds involving $A_{\infty}$. Anal. PDE **6**, 
777–818 (2013)

[6] Melas, A.D.: The Bellman functions of dyadic-like maximal operators and related
inequalities. Adv. Math. **192**, 310–340 (2005)

[7] Melas, A.D., Nikolaidakis, E.: On weak type inequalities for dyadic maximal func-
tions. J. Math. Anal. Appl. **348**, 404–410 (2008)

[8] Nazarov, F., Treil, S.: The hunt for Bellman function: applications to estimates of
singular integral operators and to other classical problems in harmonic analysis.
Algebra Anal. **8**, 32–162 (1997)

[9] Osekowski, A.: Sharp Martingale and Semimartingale Inequalities. Monografie
Matematyczne, vol. 72, Birkhäuser Basel (2012)

[10] Osekowski, A.: Sharp weak-type estimates for the dyadic-like maximal operators.
Taiwan. J. Math. **19**, 1031–1050 (2015)

[11] Osekowski, A.: Sharp weighted logarithmic bound for maximal operators. Arch.
Math. **107**, 635–644 (2016)

[12] Osekowski, A.: Sharp weighted bounds for geometric maximal operators. Glasg.
Math. J. **59**, 533–547 (2017)

[13] Osekowski, A.: Sharp $L_p$-bounds for the martingale maximal function. Tohoku
Math. J. **70**, 121–138 (2018)

[14] Vasyunin, V., Volberg, A.: Monge–Ampère equation and Bellman optimization of
Carleson embedding theorems. In: Linear Operators and Complex Analysis, pp. 195–238. American Mathematical Society, Providence (2009)
MATEUSZ RAPICKI
Wydział Matematyki Informatyki i Mechaniki
Uniwersytet Warszawski
Warsaw
Poland
e-mail: rapicki@mimuw.edu.pl

Received: 18 September 2018