Glasgow Math. J. **59** (2017) 533–547.
doi:10.1017/S0017089516000331.

SHARP WEIGHTED BOUNDS FOR GEOMETRIC MAXIMAL OPERATORS

ADAM OSĘKOWSKI

Department of Mathematics, Informatics and Mechanics,
University of Warsaw,
Banacha 2, 02-097 Warsaw, Poland
e-mail: ados@mimuw.edu.pl

(Received 4 August 2015; revised 12 November 2015; accepted 2 January 2016; first published online 10 June 2016)

Abstract. Let $\mathcal{M}$ and G denote, respectively, the maximal operator and the geometric maximal operator associated with the dyadic lattice on $\mathbb{R}^d$.

(i) We prove that for any $0 < p < \infty$, any weight w on $\mathbb{R}^d$ and any measurable f on $\mathbb{R}^d$, we have Fefferman–Stein-type estimate

$$||G(f)||_{L^p(w)} \leq e^{1/p}||f||_{L^p(\mathcal{M}w)}.$$

For each p, the constant $e^{1/p}$ is the best possible.

(ii) We show that for any weight w on $\mathbb{R}^d$ and any measurable f on $\mathbb{R}^d$,

$$\int_{\mathbb{R}^d} G(f)^{1/\mathcal{M}w} w \mathrm{d}x \leq e \int_{\mathbb{R}^d} |f|^{1/w} w \mathrm{d}x$$

and prove that the constant e is optimal.

Actually, we establish the above estimates in a more general setting of maximal operators on probability spaces equipped with a tree-like structure.

2010 *Mathematics Subject Classification.* Primary: 42B25, Secondary: 46E30, 60G42.

1. Introduction. The maximal operator $\mathcal{M}$ on $\mathbb{R}^d$ is an operator acting on measurable functions $f : \mathbb{R}^d \to \mathbb{R}$ by the formula

$$\mathcal{M}f(x) = \sup \left\{ \frac{1}{|Q|} \int_Q |f(u)| \mathrm{d}u \right\}, \qquad x \in \mathbb{R}^d,$$

where the supremum is taken over all cubes Q in $\mathbb{R}^d$ containing x, and the symbol $|Q|$ denotes the Lebesgue measure of Q. There is a modification of this object, called the geometric maximal operator, given by

$$G(f)(x) = \sup \exp \left\{ \frac{1}{|Q|} \int_Q \log(|f(u)|) \mathrm{d}u \right\}, \qquad x \in \mathbb{R}^d,$$

the supremum being taken over the same set as previously. One of basic questions about $\mathcal{M}$ and G, which have interested many mathematicians for more than 40 years now, concern the validity of various weighted inequalities for these operators. Here and below, the word "weight" will refer to a non-negative and integrable function on $\mathbb{R}^d$. In [**7**], Shi introduced the geometric maximal operator and showed the following one-weight estimate.

THEOREM 1.1. *Given a weight w, the following conditions are equivalent:*
(*i*) $w \in A_\infty$*: there exists a finite constant C such that for all cubes Q,*

$$\frac{1}{|Q|}\int_Q w dx \le C \exp\left(\frac{1}{|Q|}\int_Q \log w dx\right).$$

(*ii*) *For $0 < p < \infty$, there is a finite C_p such that the inequality*

$$||G(f)||_{L^p(w)} \le C_p ||f||_{L^p(w)}$$

holds for all $f \in L^p(w)$.

Here, $||f||_{L^p(w)} = \left(\int_{\mathbb{R}^d} |f(x)|^p w(x) \mathrm{d}x\right)^{1/p}$ is the usual weighted p-th norm of f. There is a two-weight version of the above result, established by Yin and Muckenhoupt [**14**] in the one-dimensional setting. Here is the precise statement.

THEOREM 1.2. *Given a pair of weights (u, v) on $\mathbb{R}$, the following are equivalent:*
(*i*) $(u, v) \in W_\infty$*: there exists a constant C such that for all dyadic intervals I,*

$$\int_I G(v^{-1}\chi_I) u dx \le C|I|.$$

(*ii*) *For $0 < p < \infty$, there is a constant $C_p < \infty$ for which the estimate*

$$||G(f)||_{L^p(u)} \le C_p ||f||_{L^p(v)}$$

holds for all $f \in L^p(v)$.

For further results in this direction, see e.g. Cruz-Uribe [**1**], Cruz-Uribe and Neugebauer [**2**], Ortega Salvador and Ramírez Torreblanca [**6**].

The purpose of this paper is to study a certain special class of two-weight estimates, motivated by the works of Fefferman and Stein. It follows from the results in [**3**] that for any dimension d and any constant $p \in (1, \infty)$, there is a finite number $C_{p,d}$ such that

$$||\mathcal{M}f||_{L^p(w)} \le C_{p,d} ||f||_{L^p(\mathcal{M}w)}$$

for any measurable $f : \mathbb{R}^d \to \mathbb{R}$ and any weight w. We will study the above inequality for geometric maximal operators, but in a slightly different context of probability spaces equipped with a tree-like structure. To define the necessary notions, let (X, μ) be a non-atomic probability space. Two measurable subsets A, B of X are said to be almost disjoint if $\mu(A \cap B) = 0$.

DEFINITION 1.3. A set $\mathcal{T}$ of measurable subsets of X will be called a tree, if the following conditions are satisfied:
(i) $X \in \mathcal{T}$ and for every $I \in \mathcal{T}$, we have $\mu(I) > 0$.
(ii) For every $I \in \mathcal{T}$, there is a finite subset $C(I) \subset \mathcal{T}$ containing at least two elements such that:
(a) the elements of $C(I)$ are pairwise almost disjoint subsets of I,

https://doi.org/10.1017/S0017089516000331 Published online by Cambridge University Press

(b) $I = \bigcup C(I)$.

(iii) $\mathcal{T} = \bigcup_{m \geq 0} \mathcal{T}^m$, where $\mathcal{T}^0 = \{X\}$ and $\mathcal{T}^{m+1} = \bigcup_{I \in \mathcal{T}^m} C(I)$.

(iv) we have $\lim_{m \to \infty} \sup_{I \in \mathcal{T}^m} \mu(I) = 0$.

An important example of the above setting is the cube $(0, 1]^d$ endowed with Lebesgue measure and a tree of its dyadic subcubes. Any probability space equipped with a tree gives rise to the corresponding maximal operator $\mathcal{M}_\mathcal{T}$ and the geometric maximal operator $G_\mathcal{T}$, given by

$$\mathcal{M}_\mathcal{T} f(x) = \sup \left\{ \frac{1}{\mu(I)} \int_I |f(u)| \mathrm{d}\mu(u) : x \in I, I \in \mathcal{T} \right\}$$

and

$$G_\mathcal{T}(f)(x) = \sup \left\{ \exp \left(\frac{1}{\mu(I)} \int_I \log(|f(u)|) \mathrm{d}\mu(u) \right) : x \in I, I \in \mathcal{T} \right\}.$$

One of the main goals of this paper is to establish the following result.

THEOREM 1.4. *For any $0 < p < \infty$, any measurable function $f : X \to \mathbb{R}$ and any weight w on X, we have*

$$||G_\mathcal{T}(f)||_{L^p(w)} \leq e^{1/p} ||f||_{L^p(\mathcal{M}_\mathcal{T} w)}. \tag{1.1}$$

For each p, the constant $e^{1/p}$ is the best possible.

The inequality (1.1) can be regarded as a two-weight inequality in the sense that the weights appearing on the left and on the right are different. We will show that we can put the same weight on both sides for the price of appropriate modification of the exponents. Here is the precise statement of our second result.

THEOREM 1.5. *Let w be a positive weight on X. Then, for any measurable function $f : X \to \mathbb{R}$ satisfying*

$$\int_X \log |f| d\mu \geq 0, \tag{1.2}$$

we have

$$\int_X G_\mathcal{T}(f)^{1/\mathcal{M}_\mathcal{T} w} w d\mu \leq e \int_X |f|^{1/w} w d\mu. \tag{1.3}$$

The constant e is the best possible.

Several remarks are in order. First, the estimate (1.3) can be thought of as a certain "two-variable-exponent" bound, and a nice feature of the estimate is that we do not require any growth or regularity assumptions on w. Second, the assumption (1.2) may seem a little unexpected, but we will provide examples showing that (1.3) does not hold in general with any finite constant (in the place of e) if this condition is not imposed. Our final comment is that the above statements are true in the context of the dyadic cube $(0, 1]^d$, and by the standard dilation argument, they extend to the setting of dyadic maximal operators on the whole $\mathbb{R}^d$. In particular, this yields the statements formulated in the abstract.

The proofs of (1.1) and (1.3) will be based on the existence of a certain special functions, enjoying appropriate majorization and concavity properties. This approach, called the Bellman function technique, has gathered a lot of interest in the recent literature: see e.g. [**5**, **8**, **9**, **10**, **11**, **12**, **13**] and the references therein. A nice feature of this paper is that here the Bellman function method yields optimal constants; to the best of our knowledge, this is one of the very few cases in which this approach leads to sharp results for *weighted* estimates.

The paper is organized as follows. In the next section, we establish Theorem 1.4. The final part is devoted to the proof of Theorem 1.5.

2. Proof of theorem 1.4.

2.1. A special function. As mentioned in the preceding section, a central role in the proof of the inequality (1.1) will be played by a certain special function. Introduce $B:\mathbb{R}^2\times[0,\infty)^2\to\mathbb{R}$ by the formula

$$B(x,y,w,v)=\begin{cases} e^y v(y-x-2)+e^y w & \text{if } x+1\geq y,\\ e^y w-e^{x+1}v & \text{if } x+1<y.\end{cases}$$

Some informal steps which lead to the discovery of this object are described in Subsection 2.4. It is straightforward to check that the function B is of class C^1. Further important properties are studied in the two lemmas below.

LEMMA 2.1.

(*i*) *For any* $x\in\mathbb{R}$ *and any* $w\geq 0$*, we have*

$$B(x,x,w,w)\leq 0. \tag{2.1}$$

(*ii*) *For any* $(x,y,w,v)\in\mathbb{R}^2\times[0,\infty)^2$*, we have*

$$B(x,y,w,v)\geq e^y w-e^{x+1}v. \tag{2.2}$$

Proof. The inequality (2.1) is trivial: we have $B(x,x,w,w)=-e^x w\leq 0$. The proof of the estimate (2.2) is also very simple: if $x+1<y$, then both sides of (2.2) are equal; on the other hand, if $x+1\geq y$, then the majorization is equivalent to

$$e^y(y-x-2)+e^{x+1}\geq 0.$$

If we fix x and denote the left-hand side by $F(y)$, we easily compute that $F'(y)=e^y(y-x-1)$. Hence, F is non-increasing on $(-\infty,x+1]$ and consequently, $F(y)\geq F(x+1)=0$. □

The main property of B, which can be regarded as a concavity-type condition, is studied in the following statement.

LEMMA 2.2. *Fix* $(x,y,w,v)\in\mathbb{R}^2\times[0,\infty)^2$ *satisfying* $y\geq x$ *and* $v\geq w$*. Then, for any* $h\in\mathbb{R}$ *and any* $k\geq -w$*, we have the estimate*

$$\begin{aligned} &B\big(x+h,y\vee(x+h),w+k,v\vee(w+k)\big)\\ &\qquad\leq B(x,y,w,v)+B_x(x,y,w,v)h+B_w(x,y,w,v)k. \end{aligned}\tag{2.3}$$

Proof. For the sake of convenience, we have decided to split the reasoning into three intermediate steps.

Step 1. A reduction. Fix $x,\ y,\ w,\ v,\ h,\ k$ as in the statement and consider the function $F = F_{x,y,w,v,h,k} : \mathbb{R} \to \mathbb{R}$ given by

$$F(t) = B\big(x + th, y \vee (x + th), w + tk, v \vee (w + tk)\big).$$

The assertion is equivalent to the inequality $F(1) \leq F(0) + F'(0)$. To prove this, observe first that F is concave on the interval

$$I = \{t : x + th \leq y,\ w + tk \leq v\}.$$

Indeed, this follows at once from the fact that for fixed y and v, the function $(x, w) \mapsto B(x, y, w, v)$ is concave (actually, it is a sum of a concave function in x and a linear function in w). So, (2.3) holds true if $x + h \leq y$ and $w + k \leq v$ and it suffices to study the case in which $x + h > y$ or $w + k > v$ (or $1 \notin I$). Let $t_0 = \sup I$. By the aforementioned concavity of F, we have $F(0) + F'(0) \geq F(t_0) + F'(t_0)(1 - t_0)$ and hence it is enough to show that $F(1) \leq F(t_0) + F'(t_0)(1 - t_0)$. This is further equivalent to

$$F_{x',y,w',v,h',k'}(1) \leq F_{x',y,w',v,h',k'}(0) + F'_{x',y,w',v,h',k'}(0),$$

where $x' = x + t_0h$, $w' = w + t_0k$, $h' = (1 - t_0)h$ and $k' = (1 - t_0)k$. But $x + t_0h = y$ or $w + t_0k = v$, by the very definition of t_0 and I. In other words, this shows that it is enough to establish (2.3) under the assumption that $x = y$ and $h \geq 0$, or $w = v$ and $k \geq 0$. We consider these two cases separately.

Step 2. The case $x = y$ and $h \geq 0$. Under these assumptions, we have $x + 1 > y$ and $x + h + 1 > x + h = y \vee (x + h)$, so the inequality (2.3) becomes

$$-2e^{x+h}\big((w + k) \vee v\big) + e^{x+h}(w + k) \leq -2e^x v - e^x zh + e^x(w + k),$$

or

$$-2e^h\big((w + k) \vee v\big) + (e^h - 1)(w + k) \leq -2v - vh.$$

Consider the left-hand side as a function of k (and keep all the other variables fixed). Clearly, this function is increasing on $(-\infty, v - w]$ and decreasing on $[v - w, \infty)$. Hence, it suffices to prove the latter bound for $w + k = v$, which then reads $-2e^h + e^h - 1 \leq -2 - h$. This is further equivalent to the evident estimate $e^h \geq 1 + h$, and hence (2.3) holds true.

Step 3. The case $w = v$ and $k \geq 0$. Here, the reasoning will be a little longer, as there are six subcases to consider.

(a) If $x + 1 \leq y$ and $x + h + 1 \leq y$, then (2.3) reads

$$e^y(w + k) - e^{x+h+1}(w + k) \leq e^y(w + k) - e^{x+1}(v + vh),$$

or $e^h(w + k) \geq w(1 + h)$. This is clearly true.

(b) If $x + 1 \leq y$ and $x + h \leq y \leq x + h + 1$, then (2.3) becomes

$$e^y(w + k)(y - x - h - 2) + e^y(w + k) \leq e^y(w + k) - e^{x+1}(v + vh),$$

https://doi.org/10.1017/S0017089516000331 Published online by Cambridge University Press

or $e^{y-x-1}(w+k)(y-x-h-2)+w(1+h)\leq 0$. Consider the left-hand side as a function of y and denote it by $H(y)$ (all the remaining variables are fixed). We derive that

$$H'(y)=e^{y-x-1}(w+k)(y-x-h-1)\leq 0,$$

so H is non-increasing. From the assumptions of the subcase, we know that $y\geq(x+h)\vee(x+1)$. If $h\in[0,1]$, then $H(y)\leq H(x+1)=-k(1+h)\leq 0$; if $h>1$, then

$$H(y)\leq H(x+h)=-2e^{h-1}(w+k)+w(1+h)\leq -2h(w+k)+w(1+h)\leq 0,$$

since $w+h\geq w$ and $2h\geq 1+h$. This shows that (2.3) holds also in this subcase.

(c) If $x+1\leq y<x+h$, then (2.3) takes the form

$$-2e^{x+h}(w+k)+e^{x+h}(w+k)\leq e^{y}(w+k)-e^{x+1}(v+vh).$$

One easily shows that this is true:

$$\begin{aligned}e^{y}(w+k)-e^{x+1}(v+vh)&\geq e^{x+1}(w+k)-e^{x+1}w(1+h)\\&=e^{x+1}k-e^{x+1}wh\geq -e^{x+1}wh\geq -e^{x+h}(w+k),\end{aligned}$$

since $w+k\geq w$ and $h\leq e^{h-1}$.

(d) Now assume that $y<x+1$ and $y<x+h$. Then, (2.3) becomes

$$-e^{x+h}(w+k)\leq e^{y}w(y-x-2)+e^{y}(w+k)-e^{y}wh.$$

Since $w\leq w+k$, it is enough to show that $-e^{x+h}\leq e^{y}(y-x-1)-e^{y}h$. But this is equivalent to the obvious estimate $e^{x+h-y}\geq 1+x+h-y$.

(e) Next, suppose that $y<x+1$ and $x+h<y<x+h+1$. Observe that under these conditions, the inequality (2.3) reads

$$e^{y}(w+k)(y-x-h-1)\leq e^{y}w(y-x-2)+e^{y}(w+k)-e^{y}wh,$$

or $(w+k)(y-x-h-2)\leq w(y-x-h-2)$. Since $y-x-h-2<0$ and $w+k\geq w$, the bound holds true.

(f) The final subcase is described by the conditions $y<x+1$ and $y>x+h+1$. Then, the inequality (2.3) takes the form

$$e^{y}(w+k)-e^{x+h+1}(w+k)\leq e^{y}w(y-x-2)+e^{y}(w+k)-e^{y}wh,$$

or $0\leq e^{x+h+1-y}(w+k)+w(y-x-h-2)$. But this follows at once from the trivial estimates $w+k\geq w$ and $e^{x+h+1-y}\geq 2+x+h-y$.

This completes the proof. □

2.2. Proof of (1.1)**.** Equipped with the properties of B, we turn our attention to the proof of the main inequality. First, note that it is enough to show the assertion for $p=1$: indeed, this follows from the identities

$$||G_{\mathcal{T}}(f)||^p_{L^p(X,w)}=||G_{\mathcal{T}}(|f|^p)||_{L^1(X,w)}\quad\text{and}\quad ||f||^p_{L^p(X,\mathcal{M}w)}=||\,|f|^p\,||_{L^1(X,\mathcal{M}w)}.$$

https://doi.org/10.1017/S0017089516000331 Published online by Cambridge University Press

So, till the end of the proof, we assume that $p = 1$. Let $f : X \to \mathbb{R}$ be a measurable function and let w be an arbitrary weight. By continuity argument, it suffices to show (1.1) under the additional assumption that $\mu(\{x : f(x) = 0\}) = 0$. Define four sequences $(x_n)_{n\geq 0}$, $(y_n)_{n\geq 0}$, $(w_n)_{n\geq 0}$, $(v_n)_{n\geq 0}$ of measurable functions on X as follows. Given a non-negative integer n, an element E of $\mathcal{T}^n$ and a point $x \in E$, set

$$x_n(x) = \frac{1}{\mu(E)} \int_E \log(|f(t)|)\mathrm{d}\mu(t), \qquad w_n(x) = \frac{1}{\mu(E)} \int_E w(t)\mathrm{d}\mu(t)$$

and $y_n(x) = \max_{0\leq k\leq n} x_k(x)$, $v_n(x) = \max_{0\leq k\leq n} w_k(x)$. These objects enjoy the following structural property which will be of crucial importance to the proof. Let n, E be as above and let $E_1, E_2, \ldots, E_m$ be the elements of $\mathcal{T}^{n+1}$ whose union is E. Then, we easily check that

$$\frac{1}{\mu(E)} \int_E x_n(t)\mathrm{d}\mu(t) = \sum_{i=1}^{m} \frac{\mu(E_i)}{\mu(E)} \cdot \frac{1}{\mu(E_i)} \int_{E_i} x_{n+1}(t)\mathrm{d}\mu(t), \tag{2.4}$$

$$\frac{1}{\mu(E)} \int_E w_n(t)\mathrm{d}\mu(t) = \sum_{i=1}^{m} \frac{\mu(E_i)}{\mu(E)} \cdot \frac{1}{\mu(E_i)} \int_{E_i} w_{n+1}(t)\mathrm{d}\mu(t) \tag{2.5}$$

and

$$y_{n+1}(x) = \max\{g_n(x), x_{n+1}(x)\}, \qquad v_{n+1}(x) = \max\{v_n(x), w_{n+1}(x)\}.$$

The main step of the proof is to show that the sequence $\int_X B(x_n, y_n, w_n, v_n)\mathrm{d}\mu$, $n = 0, 1, 2, \ldots$, is non-increasing. To accomplish this, fix an arbitrary integer $n \geq 0$, pick $E \in \mathcal{T}$ and let $E_1, E_2, \ldots, E_m$ be the elements of $\mathcal{T}^{n+1}$ whose union is E. By the very definition, we see that the functions x_n, y_n, w_n and v_n are constant on E; denote the corresponding values by x, y, w and v, and note that $y \geq x$, $v \geq w$. Similarly, x_{n+1} and w_{n+1} are constant on each of the sets $E_1, E_2, \ldots, E_m$; denote the corresponding values by $x + h_1, x + h_2, \ldots, x + h_m$ and $w + k_1, w + k_2, \ldots, w + k_m$. Then, (2.4) implies that $x = \sum_{i=1}^{m} \frac{\mu(E_i)}{\mu(E)}(x + h_i)$, or

$$\sum_{i=1}^{m} \frac{\mu(E_i)}{\mu(E)} h_i = 0. \tag{2.6}$$

Analogously, (2.5) implies

$$\sum_{i=1}^{m} \frac{\mu(E_i)}{\mu(E)} k_i = 0. \tag{2.7}$$

Now, by (2.3) applied to x, y, w, v and h_j, k_j, we get

$$\begin{aligned} &B\big(x + h_j, y \vee (x + h_j), w + k_j, v \vee (w + k_j)\big) \\ &\qquad \leq B(x, y, w, v) + B_x(x, y, w, v)h_j + B_w(x, y, w, v)k_j. \end{aligned}$$

Multiply both sides by $\mu(E_j)/\mu(E)$, sum the obtained inequalities over $j = 1, 2, \ldots, m$ and apply (2.6), (2.7) to get

$$\sum_{j=1}^{m} \frac{\mu(E_j)}{\mu(E)} B\big(x + h_j, y \vee (x + h_j), w + k_j, v \vee (w + k_j)\big) \leq B(x, y, w, v).$$

It is easy to see that this is equivalent to

$$\int_E B(x_{n+1}, y_{n+1}, w_{n+1}, v_{n+1}) \mathrm{d}\mu \leq \int_E B(x_n, y_n, w_n, v_n) \mathrm{d}\mu,$$

and summing over all $E \in \mathcal{T}^n$ yields the aforementioned monotonicity property of the sequence $\big(\int_X B(x_n, y_n, w_n, v_n)\mathrm{d}\mu\big)_{n \geq 0}$. Combining this property with (2.2) gives, for any non-negative integer n,

$$\int_X e^{y_n} w_n - e^{x_n+1} v_n \mathrm{d}\mu \leq \int_X B(x_n, y_n, w_n, v_n) \mathrm{d}\mu \leq \int_X B(x_0, y_0, w_0, v_0) \mathrm{d}\mu \leq 0.$$

Here in the last passage, we have used (2.1) and the equalities $x_0 = y_0$, $w_0 = v_0$. By Jensen's inequality, we see that for each $E \in \mathcal{T}^n$,

$$\int_E e^{x_n+1} v_n \mathrm{d}\mu \leq \int_E e^{\log(|f|)+1} v_n \mathrm{d}\mu = e \int_E |f| v_n \mathrm{d}\mu \leq \int_E |f| \mathcal{M} w \mathrm{d}\mu.$$

Furthermore, $\int_E e^{y_n} w_n \mathrm{d}\mu = \int_E e^{y_n} w \mathrm{d}\mu$. Thus, the preceding estimate gives

$$\int_X e^{y_n} w \mathrm{d}\mu \leq e \int_X |f| \mathcal{M} w \mathrm{d}\mu,$$

so letting $n \to \infty$ implies

$$\int_X G_{\mathcal{T}}(f) w \mathrm{d}\mu \leq e \int_X |f| \mathcal{M} w \mathrm{d}\mu,$$

by Lebesgue's monotone convergence theorem (clearly, $e^{y_n} \nearrow G_{\mathcal{T}}(f)$ almost everywhere with respect to μ). This completes the proof of (1.1).

2.3. Sharpness. Let $(X, \mathcal{T})$ be a given probability space and take the constant weight $w \equiv 1$. Fix an arbitrary constant $\kappa < 1$. We will construct a non-negative, μ-integrable function f for which $\int_X G_{\mathcal{T}}(f) \mathrm{d}\mu \geq e^{\kappa} \int_X f \mathrm{d}\mu$. By a similar argument as that used in the previous section, this will automatically show that $||G_{\mathcal{T}}(f^{1/p})||_{L^p(X,w)} = e^{1/p} ||f^{1/p}||_{L^p(X, \mathcal{M}w)}$ and the sharpness will be established.

We start with the following lemma, which can be found in [**4**].

LEMMA 2.3. *For every $I \in \mathcal{T}$ and every $\alpha \in (0, 1)$, there is a subfamily $F(I) \subset \mathcal{T}$ consisting of pairwise almost disjoint subsets of I such that*

$$\mu\left(\bigcup_{J \in F(I)} J\right) = \sum_{J \in F(I)} \mu(J) = \alpha \mu(I).$$

https://doi.org/10.1017/S0017089516000331 Published online by Cambridge University Press

We will define inductively an appropriate sequence $(A_n)_{n\geq 0}$ of subsets of X, such that each A_n will be a union of pairwise almost disjoint elements of $\mathcal{T}$: $A_n = \bigcup_{I\in F_n} I$. Let $\delta > 0$ be a fixed parameter. We start with setting $A_0 = X$; since $X \in \mathcal{T}^0$, we see that $F_0 = \{X\}$. Suppose that we have successfully constructed $A_n = \bigcup_{I\in F_n} I$. Pick $I \in F_n$ and apply Lemma 2.3 with $\alpha = \kappa/(\kappa+\delta)$. Let F_{n+1} be the union of all the families $F(I)$ corresponding to all the elements $I \in F_n$, and put $A_{n+1} = \bigcup_{I\in F_{n+1}} I$.

Directly from the definition, we see that $\mu(A_n) = (\kappa/(\kappa+\delta))^n$, $n = 0, 1, 2, \ldots$. Put

$$f = \sum_{n=0}^{\infty} e^{1-\kappa+n\delta}\chi_{A_n\setminus A_{n+1}}.$$

Fix a non-negative integer n and let $I \in F_n$ be the "atom" of A_n. Directly from the above construction, we see that

$$\begin{aligned}\frac{1}{\mu(I)}\int_I \log|f|\mathrm{d}\mu &= \sum_{k=n}^{\infty}(1-\kappa+k\delta)\left[\left(\frac{\kappa}{\kappa+\delta}\right)^{k-n} - \left(\frac{\kappa}{\kappa+\delta}\right)^{k-n+1}\right]\\ &= 1-\kappa+\frac{\delta^2}{\kappa}\left(\frac{\kappa+\delta}{\kappa}\right)^n\sum_{k=n}^{\infty}k\left(\frac{\kappa}{\kappa+\delta}\right)^{k+1} = 1+n\delta,\end{aligned}$$

which implies that $G_{\mathcal{T}}(f) \geq e^{1+n\delta}$ on A_n and hence, in particular, $G_{\mathcal{T}}(f) \geq e^{\kappa}f$ on $A_n \setminus A_{n+1}$. Since n was arbitrary, this shows $G_{\mathcal{T}}(f) \geq e^{\kappa}f$ on X and hence $||G_{\mathcal{T}}(f)||_{L^1(X)} \geq e^{\kappa}||f||_{L^1(X)}$. To finish the proof, it suffices to check that f is integrable: as we will see, this is the case if δ is sufficiently small. The first norm is equal to

$$\int_X f\mathrm{d}\mu = \sum_{n=0}^{\infty} e^{1-\kappa+n\delta}\left[\left(\frac{\kappa}{\kappa+\delta}\right)^n - \left(\frac{\kappa}{\kappa+\delta}\right)^{n+1}\right],$$

which is finite if and only if $e^{\delta} < (\kappa+\delta)/\kappa$. Since $\kappa < 1$, the latter estimate holds for sufficiently small positive δ's.

Since $\kappa < 1$ was arbitrary, this shows that the constant $e^{1/p}$ cannot be improved in (1.1).

2.4. On the guess of the optimal constant and the search for the special function. Let us now describe briefly some steps which lead to the discovery of the optimal constant $e^{1/p}$ and the above function B. We would like to stress here that the argument will be informal; its purpose is to guess the *candidate* for the special function, and the formal verification has been already carried out above. We split the argumentation into two parts.

Part I. The first step is to try to investigate the simpler, unweighted case, i.e., the estimate

$$||G_{\mathcal{T}}(f)||_{L^1(X)} \leq C||f||_{L^1(X)}, \tag{2.8}$$

for some optimal C to be identified (using the same reasoning as above, if one solves this problem, one immediately obtains the unweighted L^p inequalities with optimal constant $C^{1/p}$). Repeating the arguments from the preceding subsections, we see that it is enough to find a C^1 function $b : \mathbb{R}^2 \to \mathbb{R}$ which satisfies the conditions

$$b(x, x) \le 0 \qquad \text{for all } x \in \mathbb{R}, \tag{2.9}$$

$$b(x, y) \ge e^y - Ce^x \qquad \text{for all } x,\ y \in \mathbb{R}, \tag{2.10}$$

and the following: for any $x,\ y,\ h \in \mathbb{R}$ such that $x \le y$, we have

$$b(x + h, y \vee (x + h)) \le b(x, y) + b_x(x, y)h. \tag{2.11}$$

We will show how to find such a function if $C \ge e$. The key is to take a closer look at (2.11). In particular, this condition implies that for any fixed y the function $b(\cdot, y)$ is concave on $(-\infty, y]$. Actually, since we are interested in the value of the optimal constant, it seems reasonable to assume that for each y, $b(\cdot, y)$ is *linear* in x. Furthermore, (2.10) suggests that for each y there is a point $\gamma = \gamma(y)$ such that

$$b(\gamma(y), y) = e^y - Ce^{\gamma(y)}.$$

Indeed, otherwise we would be able to decrease the function b (keeping (2.9), (2.10) and (2.11) valid), so that this condition is satisfied. The above observations imply that b must be of the form

$$b(x, y) = e^y - Ce^{\gamma(y)} - Ce^{\gamma(y)}(x - \gamma(y)), \tag{2.12}$$

and hence we need to find a suitable γ. To do this, take $x = y$ and $h \ge 0$; then (2.11) implies

$$b(x + h, x + h) - b_x(x, y)h \le b(x, y),$$

and hence the derivative of the left-hand side, considered as a function of h, must be non-positive at zero. This is equivalent to $b_y(x, x) \le 0$ for all $x \in \mathbb{R}$. But, repeating the previous argument, since we are interested in the sharp estimate, it is natural to expect that this will actually be an equality: $b_y(x, x) = 0$ for all x. Plugging this into (2.12) leads to the differential equation

$$e^{x-\gamma(x)} = C(x - \gamma(x))\gamma'(x).$$

It is easy to see that one of the solutions is the function $\gamma(x) = x - a$, where $e^a = Ca$. But the latter equation has solutions (in a) if and only if $C \ge e$; this leads to the guess that $C = e$ is optimal. Assuming this equality, we get $a = 1$, $\gamma(x) = x - 1$ and arrive at b given by

$$b(x, y) = e^y(y - x - 1).$$

This function satisfies (2.9), (2.10) and (2.11), so (2.8) holds with the constant e. Is this the only special function which leads to the unweighted estimate? No. One can easily see that there is some freedom for modifications of b on the set $\{(x, y) : x < \gamma(y)\} = \{(x, y) : x < y - 1\}$ – we can put there any function so that the concavity of $b(\cdot, y)$ and the majorization (2.10) are preserved. For instance, the function

$$\tilde{b}(x, y) = \begin{cases} e^y(y - x - 1) & \text{if } y \le x + 1, \\ e^y - e^{x+1} & \text{if } y > x + 1 \end{cases} \tag{2.13}$$

also works fine.

https://doi.org/10.1017/S0017089516000331 Published online by Cambridge University Press

Part II. Now, we turn our attention to the weighted setting and study the estimate

$$||G_{\mathcal{T}}f||_{L^1(w)} \le C||f||_{L^1(\mathcal{M}_{\mathcal{T}}w)}.$$

A priori, we do not know the optimal constant, and our first guess is to try to show that $C = e$. If this is true, we can hope that the special function B will have some similarities with the function b, or rather $\tilde{b}$, constructed above. A look at (2.13) suggests that we should take $B(x, y, w, v) = e^y w - e^{x+1}v$ on the set $\{(x, y) : y > x + 1\}$. Then, the majorization (2.2) will actually be an equality on this set, a phenomenon which is also true in the unweighted setting (when the function $\tilde{b}$ is used). Then, taking into account the desired C^1 regularity of the function B, a natural choice on the remaining part of the domain is

$$B(x, y, w, v) = e^y w + e^y v(y - x - 2).$$

This is precisely the special function used above. We should also point out here that the alternative choice

$$B(x, y, w, v) = e^y w + e^y v(y - x - 2), \qquad (x, y, w, v) \in \mathbb{R}^2 \times [0, \infty)^2,$$

based on the function b, does not work; the condition (2.3) is not satisfied.

3. Proof of Theorem 1.5.

3.1. A special function. Clearly, it is enough to establish (1.3), since the sharpness of this estimate follows at once from the above considerations (again, we take $w \equiv 1$). Let $B : \mathbb{R} \times [0, \infty) \times (0, \infty)^2 \to \mathbb{R}$ be given by

$$B(x, y, w, v) = e^{y/v}\left(\frac{y}{v} - \frac{x}{w} - 1\right)w.$$

In Subsection 3.4, we explain the reasoning which leads to this function. Let us study some crucial properties of this object.

LEMMA 3.1.
(*i*) *For any $x \ge 0$ and any $w > 0$, we have*

$$B(x, x, w, w) \le 0. \tag{3.1}$$

(*ii*) *For any $x \in \mathbb{R}$, $y \ge 0$ and any $w, v > 0$, we have*

$$B(x, y, w, v) \ge e^{y/v}w - e^{x/w+1}w. \tag{3.2}$$

Proof. We have $B(x, x, w, w) = -e^{x/w}w < 0$, which yields (3.1). The inequality (3.2) is equivalent to

$$\frac{y}{v} - \frac{x}{w} - 1 \ge 1 - e^{x/w - y/v + 1},$$

or to the obvious bound $e^s \ge s + 1$ after the substitution $s = x/w - y/v + 1$. This completes the proof. □

Next, we study the concavity property of B.

https://doi.org/10.1017/S0017089516000331 Published online by Cambridge University Press

LEMMA 3.2. *Fix $(x, y, w, v) \in \mathbb{R} \times [0, \infty) \times (0, \infty)^2$ satisfying $y \geq x$ and $v \geq w$. Then, for any $h \in \mathbb{R}$ and any $k > -w$, we have the estimate*

$$\begin{aligned} &B\big(x+h, y \vee (x+h), w+k, v \vee (w+k)\big) \\ &\qquad \leq B(x, y, w, v) + B_x(x, y, w, v)h + B_w(x, y, w, v)k. \end{aligned}$$

Proof. The inequality reads

$$\begin{aligned} &\exp\left(\frac{y \vee (x+h)}{v \vee (w+k)}\right)\left(\frac{y \vee (x+h)}{v \vee (w+k)} - \frac{x+h}{w+k} - 1\right)(w+k) \\ &\leq e^{y/v}\left(\frac{y}{v} - 1\right)(w+k) - e^{y/v}(x+h). \end{aligned} \tag{3.3}$$

Now, we consider separately four cases.

Case I: $x + h \leq y$**,** $w + k \leq v$**.** Then, both sides of (3.3) are equal.

Case II: $x + h \geq y$**,** $w + k \geq v$**.** Then, (3.3) is equivalent to

$$\exp\left(\frac{x+h}{y+k}\right)(w+k) + e^{y/v}\left(\frac{y}{v} - 1\right)(w+k) - e^{y/v}(x+h) \geq 0. \tag{3.4}$$

Fix x, y, w, v and k. The left-hand side, considered as a function of $h \in [y-x, \infty)$, is convex; furthermore, a straightforward analysis of the derivative shows that the minimal value of the function is attained for h such that $(x+h)/(w+k) = y/v$. It remains to note that if h satisfies this equation, then (3.4) becomes an equality.

Case III: $x + h \geq y$**,** $w + k \leq v$**.** The estimate can be rewritten in the form

$$\begin{aligned} &\left[\exp\left(\frac{x+h}{v}\right)\left(\frac{x+h}{v} - 1\right) - e^{y/v}\left(\frac{y}{v} - 1\right)\right](w+k) \\ &\leq \left[\exp\left(\frac{x+h}{v}\right) - e^{y/v}\right](x+h). \end{aligned}$$

Since the function $s \mapsto e^s(s-1)$ is increasing on $[0, \infty)$, the expression in the square brackets is non-negative. This implies that it is enough to establish the inequality for the largest value of k, i.e., for $w + k = v$. This brings us back to Case II.

Case IV: $x + h \leq y$**,** $w + k \geq v$**.** Then, (3.3) takes the form

$$\left(e^{y/v} - e^{y/(w+k)}\right)(x+h) + \big(\text{terms not depending on } h\big) \leq 0.$$

But $y \geq 0$ and $w + k \geq v$, so $e^{y/v} - e^{y/(w+k)} \geq 0$ and it is enough to show the estimate for largest h, i.e., $h = y - x$. However, this has been already considered in Case II above. □

3.2. Proof of (1.3)**.** The argumentation goes along the same lines as in the proof of (1.1). With the sequences $(x_n)_{n\geq 0}$, $(y_n)_{n\geq 0}$, $(w_n)_{n\geq 0}$ and $(v_n)_{n\geq 0}$ defined as previously, we show that (3.3) implies

$$\int_X B(x_n, y_n, w_n, v_n)\mathrm{d}\mu \leq \int_X B(x_0, y_0, w_0, v_0)\mathrm{d}\mu.$$

This estimate, combined with (3.1) and (3.2), yields

$$\int_X \exp\left(\frac{y_n}{v_n}\right) w_n \mathrm{d}\mu \le e \int_X \exp\left(\frac{x_n}{w_n}\right) w_n \mathrm{d}\mu.$$

Now, observe that

$$\int_X \exp\left(\frac{y_n}{v_n}\right) w_n \mathrm{d}\mu = \int_X \exp\left(\frac{y_n}{v_n}\right) w \mathrm{d}\mu \ge \int_X \exp\left(\frac{y_n}{\mathcal{M}_{\mathcal{T}} w}\right) w \mathrm{d}\mu.$$

Furthermore, the function $(x, w) \mapsto e^{x/w} w$ is convex: its Hessian matrix

$$\begin{bmatrix} e^{x/w} w^{-1} & -e^{x/w} x w^{-2} \\ -e^{x/w} x w^{-2} & e^{x/w} x^2 w^{-3} \end{bmatrix}$$

is non-negative-definite. Thus, by Jensen's inequality,

$$\int_X \exp\left(\frac{x_n}{w_n}\right) w_n \mathrm{d}\mu \le \int_X \exp\left(\frac{\log |f|}{w}\right) w \mathrm{d}\mu.$$

Putting all the above facts together, we see that

$$\int_X \exp\left(\frac{y_n}{\mathcal{M}_{\mathcal{T}} w}\right) w \mathrm{d}\mu \le e \int_X \exp\left(\frac{\log |f|}{w}\right) w \mathrm{d}\mu,$$

and it remains to let $n \to \infty$ and apply Lebesgue's monotone convergence theorem. This completes the proof of (1.3).

3.3. On the condition (1.2)**.** Now, we will show that if (1.2) is not assumed, then (1.3) does not hold in general, even if we replace e by a larger constant. The example is the following. Let $E_1, E_2, \ldots, E_m$ be the pairwise almost disjoint elements of $\mathcal{T}^1$ whose union is X. Fix $\varepsilon \in (0, 1)$ and set

$$w = \varepsilon \chi_{E_1} + \frac{1 - \varepsilon\mu(E_1)}{1 - \mu(E_1)} \chi_{X \setminus E_1}, \qquad f = \chi_{E_1} + e^{x/(1-\mu(E_1))} \chi_{X \setminus E_1}.$$

Of course, w and f are measurable with respect to the σ-algebra generated by $\mathcal{T}^1$. Hence, when computing $\mathcal{M}_{\mathcal{T}} w$ and $G_{\mathcal{T}}(f)$, it suffices to consider the averages $\mu(I)^{-1} \int_I w \mathrm{d}\mu$ and $\mu(I)^{-1} \int_I \log |f| \mathrm{d}\mu$ for $I \in \{X, E_1, \ldots, E_m\}$ only. But

$$\int_X w \mathrm{d}\mu = 1, \qquad \int_X \log |f| \mathrm{d}\mu = x$$

(in particular, the condition (1.2) is violated), so we see that

$$\mathcal{M}_{\mathcal{T}} w = \chi_{E_1} + \frac{1 - \varepsilon\mu(E_1)}{1 - \mu(E_1)} \chi_{X \setminus E_1}, \qquad G_{\mathcal{T}}(f) = \chi_{E_1} + e^{x} \chi_{X \setminus E_1}.$$

Therefore,

$$\frac{\int_X G_{\mathcal{T}}(f)^{1/\mathcal{M}_{\mathcal{T}} w} w \mathrm{d}\mu}{\int_X |f|^{1/w} w \mathrm{d}\mu} = \frac{\varepsilon + e^{x(1-\mu(E_1))/(1-\varepsilon\mu(E_1))}(1 - \varepsilon\mu(E_1))}{\varepsilon + e^{x/(1-\varepsilon\mu(E_1))}(1 - \varepsilon\mu(E_1))} \to e^{-x\mu(E_1)},$$

https://doi.org/10.1017/S0017089516000331 Published online by Cambridge University Press

as $\varepsilon \to 0$. Consequently, if $-x$ is sufficiently big and ε is sufficiently close to 0, the above ratio can be made arbitrarily large. This shows that the inequality (1.3) does not hold with any finite constant and proves the necessity of (1.2).

3.4. On the guess of the optimal constant and the search for the special function. First, note that in the unweighted setting (i.e., for $w \equiv 1$), the inequality (1.3) reduces to the estimate (2.8) in which the constant $C = e$ is optimal. This leads us to the guess that e is also optimal in the general case, and we can hope that the special function will be similar to the functions constructed in Subsection 2.4. This time, our starting point is the function

$$b(x, y) = e^y(y - x - 1).$$

The function B we search for must satisfy the majorization

$$B(x, y, w, v) \geq e^{y/v}w - e^{x/w+1}w = w(e^{y/v} - e^{x/w+1}).$$

Comparing this to (2.10), we immediately get the natural choice

$$B(x, y, w, v) = wb(x/w, y/v) = e^{y/v}\left(\frac{y}{v} - \frac{x}{w} - 1\right)w.$$

This is precisely the function used above. Now, it can be shown that Lemma 3.2 does not hold true on the whole domain $\mathbb{R}^2 \times (0, \infty)^2$; using the alternative "base" function $\tilde{b}$ (as in Subsection 2.4) also does not work. This enforces us to restrict the domain to nonnegative y's. This corresponds to the assumption (1.2) which, as we have rigorously proved above, is necessary.

Acknowledgements. The author would like to thank an anonymous referee for the careful reading of the paper and several helpful remarks. This research was supported by the NCN grant DEC-2014/14/E/ST1/00532.

REFERENCES

1. D. Cruz-Uribe, The minimal operator and the geometric maximal operator in $\mathbb{R}^n$, *Studia Math.* **144**(1) (2001), 1–37.

2. D. Cruz-Uribe and C. J. Neugebauer, Weighted norm inequalities for the geometric maximal operator, *Publ. Mat.* **42**(1) (1998), 239–263.

3. C. Fefferman and E. M. Stein, Some maximal inequalities, *Amer. J. Math.* **93**(1) (1971), 107–115.

4. A. D. Melas, The Bellman functions of dyadic-like maximal operators and related inequalities, *Adv. Math.* **192**(2) (2005), 310–340.

5. F. Nazarov and S. Treil, The hunt for Bellman function: Applications to estimates of singular integral operators and to other classical problems in harmonic analysis, *Algebra i Analis* **8**(5) (1997), 32–162.

6. P. Ortega Salvador and C. Ramírez Torreblanca, Weighted inequalities for the one-sided geometric maximal operators, *Math. Nachr.* **284**(11–12) (2011), 1515–1522.

7. X. Shi, Two inequalities related to geometric mean operators, *J. Zhejiang Teachers College* **1** (1980), 21–25.

8. L. Slavin, A. Stokolos and V. Vasyunin, Monge-Ampère equations and Bellman functions: The dyadic maximal operator, *C. R. Acad. Sci. Paris, Ser. I* **346**(9–10) (2008), 585–588.

9. L. Slavin and V. Vasyunin, Sharp results in the integral-form John-Nirenberg inequality, *Trans. Amer. Math. Soc.* **363**(8) (2011), 4135–4169.

10. L. Slavin and A. Volberg, Bellman function and the H^1-BMO duality, *Harmonic analysis, partial differential equations, and related topics*, Contemp. Math., vol. 428 (American Mathematical Society, Providence, RI, 2007), 113–126.

11. V. Vasyunin and A. Volberg, Monge-Ampére equation and Bellman optimization of Carleson embedding theorems, *Linear and complex analysis*, Amer. Math. Soc. Transl. Ser., vol. 2, 226 (American Mathematical Society, Providence, RI, 2009), pp. 195–238.

12. V. Vasyunin and A. Volberg, Burkholder's function via Monge-Ampére equation, *Illinois J. Math.* **54**(4) (2010), 1393–1428.

13. J. Wittwer, Survey article: A user's guide to Bellman functions, *Rocky Mountain J. Math.* **41**(3) (2011), 631–661.

14. X. Yin and B. Muckenhoupt, Weighted inequalities for the maximal geometric mean operator, *Proc. Amer. Math. Soc.* **124**(1) (1996), 75–81.