WEIGHTED INEQUALITIES FOR THE TWO-DIMENSIONAL ONE-SIDED HARDY-LITTLEWOOD MAXIMAL FUNCTION

L. FORZANI, F.J. MARTÍN-REYES AND S. OMBROSI

Abstract. In this work we characterize the pair of weights (w, v) such that the one-sided Hardy-Littlewood maximal function in dimension two is of weak-type (p, p), $1 \leq p < \infty$, with respect to the pair (w, v). As an application of this result we obtain a generalization of the classic Dunford-Schwartz Ergodic Maximal Theorem for bi-parameter flows of null-preserving transformations.

1. Introduction and main results

In the 30's, it began the study of the one-sided Hardy-Littlewood maximal function

$$M^+f(x) = \sup_{h>0} \frac{1}{h} \int_x^{x+h} |f|,$$

which is defined for measurable functions $f : \mathbb{R} \to \mathbb{R}$. In the same years the basic results about the ergodic maximal operator were obtained. If $(X, \mathcal{F}, \mu)$ is a measure space and $\{\tau^t : t \in \mathbb{R}\}$ is a flow of measure preserving transformations on X, the ergodic maximal function is defined by

$$M_\tau f(x) = \sup_{h>0} \frac{1}{h} \int_0^h |f(\tau^t x)|\, dt$$

for all measurable functions $f : X \to \mathbb{R}$. We notice that M^+ is a particular case of the ergodic maximal operator since $M^+ = M_\tau$ when (X, μ) is $\mathbb{R}$ with the Lebesgue measure and $\tau^t(x) = x + t$. Nowadays it is well known that, by transference arguments, the results of the boundedness for the general operator M_τ can be obtained by the corresponding results for the particular case M^+ (see [18] for a recent exposition in the discrete case).

Although the search started with the one-sided Hardy-Littlewood maximal operator, we notice that in Harmonic Analysis the usual Hardy-Littlewood maximal operator is the two-sided operator

$$Mf(x) = \sup_{h>0} \frac{1}{2h} \int_{x-h}^{x+h} |f|.$$

Key words and phrases. **weights, one-sided maximal function, ergodic maximal theorem**

2000 *Mathematics Subject Classification* Primary: 42B25; Secondary: 28B99.

The research of the second author has been partially supported by Junta de Andalucía grant FQM 354 and spanish goverment grant MTM2005-08350-C03-02. The research of the third author has been partially supported by Universidad Nacional del Sur, grant SGCyT-UNS PGI 24/L058 and by a grant of spanish goverment MEC Res. 26/05/2006.

In $\mathbb{R}^n$, for $n \geq 1$, the Hardy-Littlewood maximal operator is defined by

$$Mf(x) = \sup_{h>0} \frac{1}{|Q(x,h)|} \int_{Q(x,h)} |f|,$$

where $Q(x,h)$ denotes the cube of center x, sides parallel to the axis, and side of length $2h$. This operator has been extensively studied. In particular Muckenhoupt [12] (see also [4]) established necessary and sufficient conditions on a positive function (weight) w for the inequality

$$\int_{\mathbb{R}^n} (Mf)^p w \leq C \int_{\mathbb{R}^n} |f|^p w, \quad 1 < p < \infty,$$

to hold for all measurable functions with a constant independent of f. Muckenhoupt solved also the same problem for the weak type inequality

$$\int_{\{x \in \mathbb{R}^n : Mf(x) > \lambda\}} w \leq \frac{C}{\lambda^p} \int_{\mathbb{R}^n} |f|^p v$$

Since then, a lot of work has been done establishing the same kind of inequalities for other operators. Surprisingly, it took fourteen years until E. Sawyer [17] characterized the good weights for the one-sided Hardy-Littlewood maximal operator in the real line. These results were applied to the ergodic setting (for instance see [11], [14], [2], [8]). We remark that weighted inequalities for other one-sided operators have been studied before and after the seminal work of Muckenhoupt. Examples of these operators are the averaging Hardy operator [1]

$$Tf(x) = \frac{1}{x} \int_0^x f$$

and the Liouville fractional integral operator

$$Tf(x) = \int_x^\infty (t-x)^\alpha f(t)\, dt.$$

We remark that these one-sided operators are defined in the real line.

E. Sawyer [16] studied the weighted inequalities for a one-sided operator in $\mathbb{R}^2$. More precisely, he considered the two-dimensional Hardy operator defined as $Hf(x,y) = \int_0^x \int_0^y f(s,t)\, ds\, dt$ for $x, y > 0$. As Sawyer mentioned in his paper, the proofs do not generalize (at least in an obvious way) to higher dimensions. Until now, the result has not been extended to $\mathbb{R}^n$ with $n \geq 3$.

But, what can be said about the one-sided Hardy-Littlewood maximal operator in $\mathbb{R}^n$ for $n > 1$? It is quite clear that a possible and natural definition is

$$M^+ f(x) = \sup_{h>0} \frac{1}{h^n} \int_{x_1}^{x_1+h} \cdots \int_{x_n}^{x_n+h} |f|,$$

where $x = (x_1, \ldots, x_n)$. Once we have fixed the operator, we settle the question of finding necessary and sufficient conditions for the weighted inequalities

$$\int_{\mathbb{R}^n} (M^+ f)^p w \leq C \int_{\mathbb{R}^n} |f|^p w$$

$$\int_{\{x \in \mathbb{R}^n : M^+ f(x) > \lambda\}} w \leq \frac{C}{\lambda^p} \int_{\mathbb{R}^n} |f|^p v$$

to hold. As far as we know, this problem has not been solved; an early discussion can be found in [10] and the characterization of the weak type of a one-sided dyadic maximal operator in $\mathbb{R}^n$ appears in [13]. However, the usual non-dyadic case seems to be more complicated and therefore it requires a deeper analysis.

In this paper, we answer the question associated with the weighted weak type inequality of M^+ in dimension two. That is, we show a characterization of the pairs of weights (w, v) such that the operator M^+ in $\mathbb{R}^2$ is of weak type (p, p) with respect to the pair (w, v). The conditions in the weights are the expectable geometric conditions similar to the conditions of the classes of Muckenhoupt $A_p\left(\mathbb{R}^2\right)$.

In order to state the main result of this paper we need to introduce some notation.

If $Q = [a-h, a] \times [b-h, b]$ is a square with sides parallel to the axis we set $Q^+ = [a, a+h] \times [b, b+h]$. Now, we define the one-sided Muckenhoupt conditions in $\mathbb{R}^2$.

Definition 1.1. *Let (w, v) be a pair of nonnegative measurable functions on $\mathbb{R}^2$. Let $1 < p < \infty$ and let p' its conjugate exponent, that is, $p + p' = 1$. It is said that (w, v) satisfies $A_p^+\left(\mathbb{R}^2\right)$, or $(w, v) \in A_p^+\left(\mathbb{R}^2\right)$, if there exists a positive constant C such that for all squares Q*

$$\left(\frac{1}{|Q|}\int_Q w\right)^{1/p}\left(\frac{1}{|Q^+|}\int_{Q^+} v^{1-p'}\right)^{1/p'} \leq C.$$

It is said that (w, v) satisfies $A_1^+\left(\mathbb{R}^2\right)$ if there exists a positive constant C such that for all $h > 0$

$$\frac{1}{h^2}\int_{x_1-h}^{x_1}\int_{x_2-h}^{x_2} w \leq Cv(x) \quad \textit{for almost every } x = (x_1, x_2).$$

$A_p^+\left(\mathbb{R}^2\right)$ is similar to the Muckenhoupt $A_p\left(\mathbb{R}^2\right)$ condition. We remind that (w, v) satisfies $A_p\left(\mathbb{R}^2\right)$, $1 < p < \infty$, if there exists a positive constant C such that for all squares Q

$$\left(\frac{1}{|Q|}\int_Q w\right)^{1/p}\left(\frac{1}{|Q|}\int_Q v^{1-p'}\right)^{1/p'} \leq C.$$

It is said that (w, v) satisfies $A_1\left(\mathbb{R}^2\right)$ if there exists a positive constant C such that for all squares Q

$$\frac{1}{|Q|}\int_Q w \leq Cv(x) \quad \text{for almost every } x \in Q.$$

It is easy to see that if (w, v) belongs to the classic Muckenhoupt condition $A_p\left(\mathbb{R}^2\right)$ and if g is a nonnegative function on $\mathbb{R}^2$ which is non decreasing on each variable separately, then $(gw, gv) \in A_p^+\left(\mathbb{R}^2\right)$. In particular, $(g, g) \in A_1^+\left(\mathbb{R}^2\right)$.

Now we are ready to state the main theorem in the paper.

Theorem 1.2. *Let (w, v) be a pair of nonnegative measurable functions on $\mathbb{R}^2$. Let $1 \leq p < \infty$. Then, the following conditions are equivalent:*

(*a*) $(w, v) \in A_p^+\left(\mathbb{R}^2\right)$

(*b*) *There is a constant C such that for every measurable function f and every $\lambda > 0$ the inequality*

$$w\left(\{x : M^+f(x) > \lambda\}\right) \leq C\lambda^{-p}\int_{\mathbb{R}^2} |f|^p\, v$$

holds, where $w(E)$ stands for $\int_E w$.

The proof is geometric and it is based on Lemma 3.1 which is a covering lemma. The search of this lemma has been inspired by the covering arguments in [13]. It is not clear for us if Lemma 3.1 can be extended to higher dimensions.

We already mentioned that if w is a nonnegative function on $\mathbb{R}^2$ which is non decreasing on each variable separately then $(w, w) \in A_1^+\left(\mathbb{R}^2\right)$. Therefore, we have the following corollary.

Corollary 1.3. *If w is a nonnegative function on $\mathbb{R}^2$ which is increasing (non decreasing) on each variable separately then*

$$\int_{\{x:M^+f(x)>\lambda\}} w \leq \frac{C}{\lambda}\int_{\mathbb{R}^n}|f|w$$

for all $\lambda > 0$ and all measurable function f.

Actually, we want to point out that this corollary follows easily from the Dunford-Schwartz ergodic maximal theorem. In the next theorem we recall the result by Dunford and Schwartz.

Theorem 1.4. [6, Lemma VII.7.11] *Let $(X, \mathcal{F}, \mu)$ be a measure space and let $\mathcal{T} = \{T^t : t = (t_1, \ldots, t_n), t_1, \ldots, t_n > 0\}$ be a strongly measurable semi-group of operators in $L^1(X, \mathcal{F}, \mu)$ with $\|T^t\|_1 \leq 1$ and $\|T^t\|_\infty \leq 1$. Let*

$$M_{\mathcal{T}}f(x) = \sup_{h>0}\frac{1}{h^n}\left|\int_0^h \ldots \int_0^h T^t f(x)\, dt_1 \ldots dt_n\right|.$$

Then there is an absolute constant C_n, which is independent of the semi-group and independent of f, such that

$$\mu(\{x \in X : M_{\mathcal{T}}f(x) > \lambda\}) \leq \frac{1}{C_n\lambda}\int_{\{x\in X:M_{\mathcal{T}}f(x)>C_n\lambda\}}|f|\, d\mu$$

for all $\lambda > 0$.

Observe that if w is a nonnegative function in $\mathbb{R}^n$ which is increasing on each variable separately then the semigroup of operators $T^t f(x) = f(x + t)$, $t \in \mathbb{R}^n$, $t = (t_1, \ldots, t_n)$, $t_i > 0$, is a contraction in $L^1(w)$ and in $L^\infty(w)$. Therefore, we can apply Dunford-Schwartz ergodic theorem and Corollary 1.3 follows not only in $\mathbb{R}^2$ but in $\mathbb{R}^n$ for all n. This result seems to be not well known and the authors have not found it in the literature.

We point out that we do not know any geometric proof of Dunford-Schwartz ergodic theorem. However, Theorem 1.2 gives a geometric proof of Corollary 1.3 which is Dunford-Schwartz ergodic theorem for the semigroup $T^t f(x) = f(x + t)$ in $\mathbb{R}^2$.

As an application in Ergodic Theory of our main result we obtain a theorem which is in some sense an extension of Dunford-Schwartz ergodic Theorem. In order to state it, consider a σ-finite measure space $(X, \mathcal{F}, \mu)$ and let $\{\tau^t : t \in \mathbb{R}^2\}$ be a bi-parameter flow of null-preserving transformations on X, that is,

(a) For all $t \in \mathbb{R}^2$, $\tau^t : X \to X$ is a measurable map such that if $\mu(E) = 0$ then $\mu(\tau^t E) = 0$.

(b) The map $(t,x) \to \tau^t(x)$ from $\mathbb{R}^2 \times X \to X$ is measurable with respect to the completion of the σ-algebra product in $\mathbb{R}^2 \times X$.

(c) $\tau^0(x) = x$ for all $x \in X$ and $\tau^t \circ \tau^s = \tau^{t+s}$ for all $t, s \in \mathbb{R}^2$.

The flow induces a group $\mathcal{T} = \{T^t : t \in \mathbb{R}^2\}$ of operators acting on measurable functions and defined by

$$T^t f(x) = f(\tau^t x).$$

For each $h > 0$ we consider the averages over squares

$$A_h f(x) = \frac{1}{h^2} \int_0^h \int_0^h T^t f(x) dt. \tag{1}$$

To study the convergence of A_h as $h \to +\infty$ the usual thing is to consider the ergodic maximal operator

$$M_{\mathcal{T}} f(x) = \sup_{h>0} |A_h f(x)| .$$

By using Dunford-Schwartz Theorem quoted in the introduction, we have that if each T^t is a contraction in $L^1(\mu)$ then the maximal operator $M_{\mathcal{T}}$ is of weak type $(1,1)$, that is, there exists C such that

$$\mu(\{x \in X : M_{\mathcal{T}} f(x) > \lambda\}) \leq \frac{C}{\lambda} \int_X |f| \, d\mu \tag{2}$$

for all $\lambda > 0$ and all $f \in L^1(\mu)$. Our result in Ergodic Theory, Theorem 1.5, states that (2) holds under the assumption that the group $\mathcal{T}$ is Cesàro bounded in $L^1(\mu)$, which means that there exists $C > 0$ such that

$$\sup_{h>0} \int_X |A_h f| \leq C \int_X |f| \, d\mu. \tag{3}$$

for all measurable function $f \geq 0$. Now we are ready to state the theorem.

Theorem 1.5. *Let $(X, \mathcal{F}, \mu)$ be a σ-finite measure space and let $\{\tau^t : t \in \mathbb{R}^2\}$ be a flow of null-preserving transformations on X. Assume that the group $\mathcal{T}$ is Cesàro bounded in $L^1(\mu)$. Then there is a constant $C > 0$ such that*

$$\mu(\{x \in X : M_{\mathcal{T}} f(x) > \lambda\}) \leq \frac{C}{\lambda} \int_X |f| \, d\mu$$

for all $\lambda > 0$ and all $f \in L^1(\mu)$.

We remark that the assumption about the group together with the properties of the flow assures that $A_h f$ is defined and (3) holds for all $f \in L^1(\mu)$. Notice also that if each T^t is a contraction in $L^1(\mu)$ then the group $\mathcal{T}$ is Cesàro bounded in $L^1(\mu)$. Therefore, our assumption is weaker than the one we need to apply Dunford-Schwartz Theorem (in Final Remarks we observe that there are groups which are Cesàro-bounded but the operators are not contraction in $L^1(\mu)$).

The paper is organized as follows: §2 is dedicated to introduce notations, the definitions of some maximal operators and some results about them; in §3 and §4 we prove the main result and the covering lemma, respectively; while in the last section we prove Theorem 1.5 and we make some remarks.

As usual, if $E \subset \mathbb{R}^n$ is measurable, $|E|$ denotes the Lebesgue measure of E and if w is a measurable function then $w(E) = \int_E w$. Throughout the paper, the letter C will denote a positive constant whose value may change from line to line.

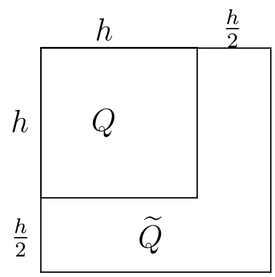

FIGURE 1. Q and $\widetilde{Q}$.

2. Notation and Definitions

If $I = [a, b]$ is a bounded interval we denote $I^+ = [b, 2b - a]$ and $I^- = [2a - b, a]$. By a square we mean a square with sides parallel to the axis. If $Q = I_1 \times I_2$ is a square then $l(Q)$ stands for the length of the side of Q, that is, the length of I_1 or I_2, and we denote the squares $I_1^+ \times I_2^+$, and $I_1^- \times I_2^-$ by Q^+ and Q^-, respectively. We shall say that a square Q is of dyadic size if $l(Q) = 2^k$ for some $k \in \mathbb{Z}$. If Q is a square and α is a positive number, αQ is the square with the same center as Q and $l(\alpha Q) = \alpha l(Q)$. If $Q = [a, a + h] \times [b, b + h]$ then $\widetilde{Q}$ is the dilation of Q to the right and to the bottom in half the length of the side of Q, that is, $\widetilde{Q} = [a, a + \frac{3}{2}h] \times [b - \frac{h}{2}, b + h]$. See Figure 1.

Let $x \in \mathbb{R}^2$, $x = (x_1, x_2)$, and let h be a positive real number. We denote $Q_{x,h} = [x_1, x_1 + h] \times [x_2, x_2 + h]$, $Q_{x,h^-} = [x_1 - h, x_1] \times [x_2 - h, x_2]$. With this notation, we define the maximal functions

$$M^+ f(x) = \sup_{h>0} \frac{1}{|Q_{x,h}|} \int_{Q_{x,h}} |f| \quad \text{and} \quad M^- f(x) = \sup_{h>0} \frac{1}{|Q_{x,h^-}|} \int_{Q_{x,h^-}} |f| \,.$$

Now we divide the square $Q_{x,h}$ into four squares (see Figure 2):

$$\begin{aligned} Q_{x,h} &= Q_{x,\frac{h}{2}} \cup Q^1_{x,h} \cup Q^2_{x,h} \cup Q^3_{x,h} \\ Q^2_{x,h} &= [x_1 + \frac{h}{2}, x_1 + h] \times [x_2, x_2 + \frac{h}{2}], \\ Q^3_{x,h} &= [x_1, x_1 + \frac{h}{2}] \times [x_2 + \frac{h}{2}, x_2 + h] \\ &\text{and} \\ Q^1_{x,h} &= [x_1 + \frac{h}{2}, x_1 + h] \times [x_2 + \frac{h}{2}, x_2 + h], \end{aligned}$$

and we define

$$\begin{aligned} M^{+1} f(x) &= \sup_{h>0} \frac{1}{\left|Q^1_{x,h}\right|} \int_{Q^1_{x,h}} |f| \,, \\ M^{+2} f(x) &= \sup_{h>0} \frac{1}{\left|Q^2_{x,h}\right|} \int_{Q^2_{x,h}} |f| \\ M^{+3} f(x) &= \sup_{h>0} \frac{1}{\left|Q^3_{x,h}\right|} \int_{Q^3_{x,h}} |f| \end{aligned}$$

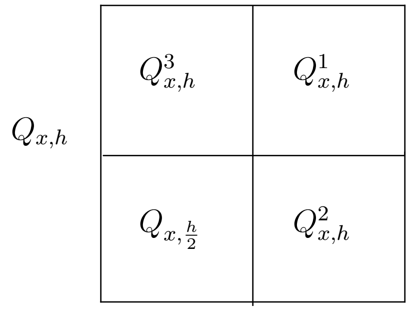

FIGURE 2. Subsquares.

We have that M^+ is essentially equivalent to the sum of the maximal operators M^{+i}, $i = 1, 2, 3$. We state this result in the next proposition.

Proposition 2.1. *The following inequality holds for every measurable function:*

$$\frac{1}{12}(M^{+1}f(x)+M^{+2}f(x)+M^{+3}f(x)) \leq M^+f(x) \leq \frac{1}{3}(M^{+1}f(x)+M^{+2}f(x)+M^{+3}f(x)).$$

Proof. By density arguments, it is enough to prove it for functions $f \in L^1(dx)$. It is clear that for every $h > 0$

$$\frac{1}{\left|Q^i_{x,h}\right|}\int_{Q^i_{x,h}} |f| \leq 4\frac{1}{|Q_{x,h}|}\int_{Q_{x,h}} |f| \leq 4M^+f(x).$$

Therefore $M^{+1}f(x) + M^{+2}f(x) + M^{+3}f(x) \leq 12M^+f(x)$. On the other hand, if $h > 0$ we have

$$\begin{aligned}
\frac{1}{|Q_{x,h}|}\int_{Q_{x,h}} |f| &= \frac{1}{|Q_{x,h}|}\left(\int_{Q_{x,\frac{h}{2}}} |f| + \int_{Q_{x,h}\setminus Q_{x,\frac{h}{2}}} |f|\right)\\
&\leq \frac{1}{4\left|Q_{x,\frac{h}{2}}\right|}\int_{Q_{x,\frac{h}{2}}} |f| + \frac{1}{4}\left(M^{+1}f(x) + M^{+2}f(x) + M^{+3}f(x)\right)\\
&\leq \frac{1}{4}\left(M^+f(x) + M^{+1}f(x) + M^{+2}f(x) + M^{+3}f(x)\right).
\end{aligned}$$

Taking supremum on $h > 0$

$$M^+f(x) \leq \frac{1}{4}\left(M^+f(x) + M^{+1}f(x) + M^{+2}f(x) + M^{+3}f(x)\right).$$

Since $f \in L^1$ we have $M^+f(x) < \infty$ a.e., and therefore

$$M^+f(x) \leq \frac{1}{3}(M^{+1}f(x) + M^{+2}f(x) + M^{+3}f(x)).$$

□

For technical reasons, in the proof of the main result we shall use the maximal operator $\mathcal{M}^+$ defined by

$$\mathcal{M}^+f(x) = \sup_{k\in\mathbb{Z}} \frac{1}{\left|Q_{x,2^k}\right|}\int_{Q_{x,2^k}} |f|,$$

that is, we only take cubes $Q_{x,h}$ of dyadic size. This operator is essentially equivalent to M^+. In fact,

$$\frac{1}{4}M^+ \leq \mathcal{M}^+ \leq M^+.$$

We consider also $\mathcal{M}^{+i}$, $i = 1, 2, 3, 4$, defined by

$$\mathcal{M}^{+i} f(x) \quad = \quad \sup_{k\in\mathbb{Z}} \frac{1}{\left|Q^i_{x,2^k}\right|} \int_{Q^i_{x,2^k}} |f| \,.$$

The operators $\mathcal{M}^+$ and $\mathcal{M}^{+i}$ are related in the same way as M^+ and M^{+i}. We establish this relation in the next proposition.

Proposition 2.2. *The following inequality holds for every measurable function:*

$$\frac{1}{12}(\mathcal{M}^{+1}f(x)+\mathcal{M}^{+2}f(x)+\mathcal{M}^{+3}f(x)) \leq \mathcal{M}^+ f(x) \leq \frac{1}{3}(\mathcal{M}^{+1}f(x)+\mathcal{M}^{+2}f(x)+\mathcal{M}^{+3}f(x)).$$

The proof is the same as the proof of Proposition 2.1.

3. Proof of Theorem 1.2

The relevant implication is $(a) \Longrightarrow (b)$ since $(b) \Longrightarrow (a)$ follows as in the classic case of Muckenhoupt weights and we omit it.

$(a) \Longrightarrow (b)$. Since M^+ and $\mathcal{M}^+$ are essentially equivalent, it is enough to prove (b) for the operator $\mathcal{M}^+$, that is, we are going to prove the following inequality:

$$w\left(\{x : \mathcal{M}^+ f(x) > \lambda\}\right) \leq C\lambda^{-p} \int_{\mathbb{R}^2} |f|^p \, v \tag{4}$$

Observe that (4) follows from the inequality

$$w\left(\{x : \lambda < \mathcal{M}^+ f(x) \leq 2\lambda\}\right) \leq \frac{C}{\lambda^p} \int_{\mathbb{R}^2} |f|^p v, \tag{5}$$

In fact, if (5) holds then

$$\begin{aligned} w\left(\{x : \mathcal{M}^+ f(x) > \lambda\}\right) &= \sum_{k=0}^{\infty} w\left(\{x : 2^k\lambda < \mathcal{M}^+ f(x) \leq 2^{k+1}\lambda\}\right) \\ &\leq \sum_{k=0}^{\infty} \frac{C}{2^{pk}\lambda^p} \int_{\mathbb{R}^2} |f|^p v = \frac{2^p C}{(2^p-1)\lambda^p} \int_{\mathbb{R}^2} |f|^p v. \end{aligned}$$

Proof of (5). By Proposition 2.2, we only have to prove that

$$w\left(\{x : \lambda < \mathcal{M}^{+i} f(x), \mathcal{M}^+ f(x) \leq 2\lambda\}\right) \leq \frac{C}{\lambda^p} \int_{\mathbb{R}^2} |f|^p v \quad \text{for } i = 1, 2, 3, \tag{6}$$

with a constant independent of f and λ. We shall prove it for $i = 2$, being similar for $i = 1, 3$.

Proof of (6) *for* $i = 2$. Let us consider for each $\xi > 0$ the truncated maximal operator

$$\mathcal{M}^{+2}_{\xi} f(x) = \sup_{h=2^k>\xi, k\in\mathbb{Z}} \frac{4}{h^2} \int_{Q^2_{x,h}} |f| \,.$$

Since $\mathcal{M}_\xi^{+2} f \uparrow \mathcal{M}^{+2} f$ as $\xi \downarrow 0^+$ it follows from the monotone convergence theorem that it suffices to prove that

$$w\left(\{x : \lambda < \mathcal{M}_\xi^{+2} f(x), \mathcal{M}^+ f(x) \le 2\lambda\}\right) \le \frac{C}{\lambda^p} \int_{\mathbb{R}^2} |f|^p v, \tag{7}$$

for all $\lambda > 0$ and all measurable f with a constant independent of ξ, λ and f.

To prove (7) we shall need the following covering lemma which is the key result of this paper. We notice that we need similar but different lemmas if we are dealing with $\mathcal{M}^{+i}$, $i = 1, 3$, instead of $\mathcal{M}^{+2}$.

Lemma 3.1. *let f be a nonnegative measurable function. Let $A = \{x_j, j = 1, \dots, n\}$ a finite set of points on $\mathbb{R}^2$. Assume that for each $x_j \in A$ we have an associated square Q_j of dyadic size such that its upper right corner is x_j and*

$$\frac{1}{|Q_j|} \int_{Q_j^{+2}} f > \frac{\lambda}{4}.$$

Then, there exists a set $\Gamma \subset \{1, \dots, n\}$ such that, if $\widetilde{Q_j}$ is the dilation of Q_j to the right and to the bottom in half the length of the side of Q_j, we have

$$A \subset \bigcup_{i \in \Gamma} \widetilde{Q_i}, \tag{8}$$

and

$$\frac{1}{|Q_j|} \int_{(\widetilde{Q_j})^+} f > \frac{\lambda}{4} \tag{9}$$

Moreover, $\widetilde{Q_i} \nsubseteq \widetilde{Q_j}$, $i \neq j$, $i, j \in \Gamma$, and the squares $\widetilde{Q_i}$, $i \in \Gamma$, of the same size are almost disjoints, that is, there exists a constant C such that for all l

$$\sum_{\{i \in \Gamma : l(Q_i) = l\}} \chi_{\widetilde{Q_i}} \le C.$$

(Consequently, the squares $(\widetilde{Q_i})^+$ with $i \in \Gamma$ of the same size are almost disjoints too.)
Further, if

$$\frac{1}{|Q_j|} \int_{(\widetilde{Q_j})^+} f \le 8\lambda \tag{10}$$

then there exists a family of sets $\{F_j\}_{j \in \Gamma}$ with $F_j \subset (\widetilde{Q_j})^+$, such that

$$\frac{\lambda}{8} < \frac{1}{|Q_j|} \int_{F_j} f. \tag{11}$$

and they are almost disjoint, i.e, there exists C (independent of everything), such that

$$\sum_{j \in \Gamma} \chi_{F_j}(x) \le C. \tag{12}$$

We postpone the proof of the lemma to the next section.

Proof of (7). Observe first that if $(w_i, v_i) \in A_p^+(\mathbb{R}^2)$, $i = 1, 2$, then

$$(\max\{w_1, w_2\}, \max\{v_1, v_2\}) \in A_p^+(\mathbb{R}^2) \text{ y } (\min\{w_1, w_2\}, \min\{v_1, v_2\}) \in A_p^+(\mathbb{R}^2).$$

In particular, for each $n \in \mathbb{N}$, if $(w, v) \in A_p^+(\mathbb{R}^2)$ then the pairs (w_n, v_n) and $(\widetilde{w}_n, \widetilde{v}_n)$ belong to $A_p^+(\mathbb{R}^2)$ with a uniform constant, where $w_n = \max\{w, \frac{1}{n}\}$, $v_n = \max\{v, \frac{1}{n}\}$, $\widetilde{w}_n = \min\{w, n\}$ and $\widetilde{v}_n = \min\{v, n\}$.

It is enough to prove (7) for bounded functions $f \in L^p(v)$ with compact support. It follows from the above remark that we may assume also that w is locally integrable and there exists $\gamma > 0$ such that

$$0 < \gamma \le w(x) \quad \text{for all } x \in \mathbb{R}^2.$$

Let $E = \{x \in \mathbb{R}^2 : \lambda < \mathcal{M}_\xi^{+2} f(x), \mathcal{M}^+ f(x) \le 2\lambda\}$. We notice that the weighted measure $w(x)\,dx$ is finite on compact sets since w is locally integrable. Therefore it is enough to show that there exists $C > 0$ such that

$$w(K) \le \frac{C}{\lambda^p} \int_{\mathbb{R}^2} |f|^p v. \tag{13}$$

for all compact set $K \subset E$.

Let us fix a compact set $K \subset E$. For each $x = (x_1, x_2) \in K$ there exists a square $Q_x = [x_1 - l, x_1] \times [x_2 - l, x_2]$ with $\xi \le l$ and $l = 2^k$ for some $k \in \mathbb{Z}$ such that

$$\frac{\lambda}{4} < \frac{1}{|Q_x|} \int_{Q_x^{+2}} |f|.$$

Let $Q_{x,2l} = [x_1, x_1 + 2l] \times [x_2, x_2 + 2l]$. It is clear that $(\widetilde{Q_x})^{+2} \subset Q_{x,2l}$ and therefore

$$\begin{aligned} \frac{1}{|Q_x|} \int_{(\widetilde{Q_x})^{+2}} |f| &\le \frac{1}{|Q_x|} \int_{Q_{x,2l}} |f| \\ &= \frac{4}{|Q_{x,2l}|} \int_{Q_{x,2l}} |f| \le 4\mathcal{M}^+ f(x) \le 8\lambda, \end{aligned}$$

where the last inequality follows from $x \in K \subset E$. Consequently, for each $x \in K$ we have a square $Q_x = [x_1 - l, x_1] \times [x_2 - l, x_2]$ such that $\xi \le l$,

$$\frac{\lambda}{4} < \frac{1}{|Q_x|} \int_{Q_x^{+2}} |f| \quad \text{and} \quad \frac{1}{|Q_x|} \int_{(\widetilde{Q_x})^{+2}} |f| \le 8\lambda.$$

Observe that $l \le M$ for certain positive real number M depending on λ and f. This follows from the inequalities

$$|Q_x| \le \frac{4}{\lambda} \int_{Q_x^{+2}} |f| \le \frac{4}{\lambda} \int_{\mathbb{R}^2} |f| < \infty.$$

Since $l \le M$, $x \in K$ and K is compact we have that the union $\cup_{x \in K} \widetilde{Q_x}$ is a bounded set. Thus, there exists a square R such that

$$\cup_{x \in K} \widetilde{Q_x} \subset R.$$

Let us consider the square $2R$. Since w is integrable on $2R$, there exists ε, $0 < \varepsilon < 1$, such that if $Q \subset R$ is a square then

$$w((1 + \varepsilon)Q \setminus Q) \le \gamma \xi^2.$$

If $l(Q) \geq \xi$ and $Q \subset R$ then

$$w((1+\varepsilon)Q \setminus Q) \leq \gamma \xi^2 \leq \gamma |Q| \leq w(Q).$$

Consequently,

$$w((1+\varepsilon)Q) \leq 2w(Q)$$

for all square $Q \subset R$ such that $l(Q) \geq \xi$. In particular

$$w((1+\varepsilon)\widetilde{Q_x}) \leq 2w(\widetilde{Q_x}), \quad \text{for all } x \in K.$$

Let us denote by $B_x(r)$ to the ball of center x and radius r. It is clear that

$$K \subset \cup_{x\in K} B_x(\frac{\xi\varepsilon}{2}).$$

Since K is compact, there exist $x_1, \dots, x_s \in K$ such that $K \subset \cup_{j=1}^{s} B_{x_j}(\frac{\xi\varepsilon}{2})$. Applying the covering lemma to the set $A = \{x_1, \dots, x_s\}$ and the squares $\{Q_{x_j} : j = 1, \dots, s\}$, there exists $\Gamma \subset \{1, \dots, s\}$ such that

$$A = \{x_1, \dots, x_s\} \subset \cup_{i\in\Gamma} \widetilde{Q_{x_i}}$$

Further, there exists $\{F_{x_i} : i \in \Gamma\}$ such that

$$F_{x_i} \subset (\widetilde{Q_{x_i}})^+, \tag{14}$$

$$\frac{\lambda}{8} < \frac{1}{|Q_{x_i}|} \int_{F_{x_i}} |f|. \tag{15}$$

and

$$\sum_{i\in\Gamma} \chi_{F_{x_i}}(x) \leq C. \tag{16}$$

Now, observe that if $x_j \in A$ then there exists $i \in \Gamma$ such that $x_j \in \widetilde{Q_{x_i}}$. This implies

$$B_{x_j}(\frac{\xi\varepsilon}{2}) \subset (1+\varepsilon)\widetilde{Q_{x_i}}.$$

Therefore

$$K \subset \cup_{j=1}^{s} B_{x_j}(\frac{\xi\varepsilon}{2}) \subset \cup_{i\in\Gamma} (1+\varepsilon)\widetilde{Q_{x_i}}$$

and

$$w(K) \leq \sum_{i\in\Gamma} w((1+\varepsilon)\widetilde{Q_{x_i}}) \leq 2\sum_{i\in\Gamma} w(\widetilde{Q_{x_i}}).$$

Assume now that $p > 1$. Then using (15) and Hölder's inequality

$$\begin{aligned} w(K) \leq 2\sum_{i\in\Gamma} w(\widetilde{Q_{x_i}}) &\leq \frac{C}{\lambda^p} \sum_{i\in\Gamma} w(\widetilde{Q_{x_i}}) \left(\frac{1}{|Q_{x_i}|} \int_{F_{x_i}} |f| \right)^p \\ &= \frac{C}{\lambda^p} \sum_{i\in\Gamma} w(\widetilde{Q_{x_i}}) \left(\frac{1}{|Q_{x_i}|} \int_{F_{x_i}} |f| v^{\frac{1}{p}} v^{-\frac{1}{p}} \right)^p \\ &= \frac{C}{\lambda^p} \sum_{i\in\Gamma} \left[\frac{w(\widetilde{Q_{x_i}})}{|Q_{x_i}|^p} \left(\int_{F_{x_i}} v^{-\frac{1}{p-1}} \right)^{p-1} \right] \int_{F_{x_i}} |f|^p v \end{aligned}$$

Since $F_{x_i} \subset (\widetilde{Q_{x_i}})^+$ and from $A_p^+(\mathbb{R}^2)$ condition we obtain

$$\begin{aligned}\frac{w(\widetilde{Q_{x_i}})}{|Q_{x_i}|^p}\left(\int_{F_{x_i}} v^{-\frac{1}{p-1}}\right)^{p-1} &\leq \frac{w(\widetilde{Q_{x_i}})}{|Q_{x_i}|^p}\left(\int_{(\widetilde{Q_{x_i}})^+} v^{-\frac{1}{p-1}}\right)^{p-1}\\ &\leq C\end{aligned}$$

Combining these last estimates and the fact that the sets F_{x_i}, $i \in \Gamma$, are almost disjoint we get (13) for $p > 1$.

If $p = 1$ we use that

$$\frac{w(\widetilde{Q_{x_i}})}{|Q_{x_i}|} \leq Cv(x)$$

for almost every $x \in (\widetilde{Q_{x_i}})^+$ (by condition $A_1^+(\mathbb{R}^2)$). Since $F_{x_i} \subset (\widetilde{Q_{x_i}})^+$ we obtain

$$\begin{aligned} w(K) \leq 2\sum_{i\in\Gamma} w(\widetilde{Q_{x_i}}) &\leq \frac{C}{\lambda}\sum_{i\in\Gamma} w(\widetilde{Q_{x_i}})\frac{1}{|Q_{x_i}|}\int_{F_{x_i}} |f|\\ &\leq \frac{C}{\lambda}\sum_{i\in\Gamma}\int_{F_{x_i}} |f|v\\ &\leq \frac{C}{\lambda}\int_{\mathbb{R}^2} |f|v,\end{aligned}$$

where in the last inequality we have used that the sets F_{x_i} are almost disjoint.

4. Proof of Lemma 3.1

Before starting with the proof of the covering lemma, we need to introduce a notion of maximality and to state some previous lemmas.

Definition 4.1. *Assume that $\mathcal{F}$ is a finite family of squares. Let $l_0 = \max\{l(Q) : Q \in \mathcal{F}\}$ and consider $\Sigma_0 = \{Q \in \mathcal{F} : l(Q) = l_0\}$. Now consider $\mathcal{F}_1 = \{Q \in \mathcal{F} : Q \cap R = \emptyset \quad \text{for all } R \in \Sigma_0\}$. If $\mathcal{F}_1 = \emptyset$ then the process stops. If $\mathcal{F}_1 \neq \emptyset$ then we take $l_1 = \max\{l(Q) : Q \in \mathcal{F}_1\}$ and we set $\Sigma_1 = \{Q \in \mathcal{F}_1 : l(Q) = l_1\}$. Now we continue the process considering $\mathcal{F}_2 = \{Q \in \mathcal{F}_1 : Q \cap R = \emptyset \quad \text{for all } R \in \Sigma_1\}$. Since the family F is finite the process stops in a finite number of steps. Assume that $\Sigma_0, \ldots, \Sigma_{k_0}$ have been chosen and $\mathcal{F}_{k_0+1} = \{Q \in \mathcal{F}_{k_0} : Q \cap R = \emptyset \quad \text{for all } R \in \Sigma_{k_0}\} = \emptyset$. We shall say that the squares belonging to $\cup_{i=0}^{k_0}\Sigma_i$ are maximal squares in $\mathcal{F}$.*

Remark 4.2. *Observe that if $\mathcal{F}$ is a finite family of squares and $Q \in \mathcal{F}$ then either Q is maximal in $\mathcal{F}$ or there exists a maximal square Q_m such that $l(Q) < l(Q_m)$ and $Q_m \cap Q \neq \emptyset$. We have also that if Q_i and Q_j are maximal in $\mathcal{F}$ then $l(Q_i) = l(Q_j)$ or $Q_i \cap Q_j = \emptyset$.*

Before stating the next lemma, remind the notation introduced in §2. In particular, $\widetilde{Q}$ is the dilation of Q to the right and to the bottom in half the length of the side of Q.

Lemma 4.3. *Fix a square Q. Let $\mathcal{F} = \{Q_j : j \in \Gamma\}$ a finite family of squares of dyadic size and such that the squares $\widetilde{Q}_j$, $j \in \Gamma$, of the same size are almost disjoints, i.e., there exists a constant A such that for all l*

$$\sum_{\{i\in\Gamma : l(Q_i)=l\}} \chi_{\widetilde{Q}_i} \leq A.$$

Let $\Gamma_0 = \{j \in \Gamma : \left(\widetilde{Q_j} \cup (\widetilde{Q_j})^+\right) \cap \partial Q \neq \emptyset, |Q_j| < |Q|\}$, *where* ∂Q *denotes the border of* Q. *Then there exists a constant* C *depending only on* A *such that*

$$\sum_{j\in\Gamma_0} \left|\widetilde{Q_j}\right| \le C\,|Q|\,.$$

Proof. Let $l\,(Q) = l_0$. Then there exists $k_0 \in \mathbb{Z}$ such that $2^{-k_0} < l_0 \le 2^{-k_0+1}$. Therefore

$$\Gamma_0 = \cup_{k=k_0}^{\infty}\{j \in \Gamma : \left(\widetilde{Q_j} \cup (\widetilde{Q_j})^+\right) \cap \partial Q \neq \emptyset, |Q_j| = (1/2^k)^2\}.$$

If L_i, $i = 1, 2, 3, 4$, are the sides of the square Q it is clear that it will suffice to prove

$$\sum_{k=k_0}^{\infty} \sum_{j\in\Gamma_{k,i}} \left|\widetilde{Q_j}\right| \le C\,|Q|\,,$$

where $\Gamma_{k,i} = \{j \in \Gamma : \left(\widetilde{Q_j} \cup (\widetilde{Q_j})^+\right) \cap L_i \neq \emptyset, |Q_j| = (1/2^k)^2\}$. We will prove it for one of the sides, being similar for the others. If $Q = [a,b]\times[c,d]$ let $L_1 = \{a\}\times[c,d]$. Let $k \ge k_0$. Then there exists a rectangle R_k such that $|R_k| = 42 l_0 2^{-k}$ and

$$\bigcup_{j\in\Gamma_{k,1}} \widetilde{Q_j} \cup (\widetilde{Q_j})^+ \subset R_k.$$

(Take $R_k = [a - \frac{3}{2^k}, a + \frac{3}{2^k}] \times [c - 3l_0, d + 3l_0]$). Then

$$\begin{aligned}
\sum_{j\in\Gamma_{k,1}} \left|\widetilde{Q_j}\right| &= C \sum_{k=k_0}^{\infty} \sum_{j\in\Gamma_{k,1}} |Q_j| \\
&\le C \sum_{k=k_0}^{\infty} \left|\bigcup_{j\in\Gamma_{k,1}} Q_j\right| \\
&\le C \sum_{k=k_0}^{\infty} |R_k| \\
&\le C \sum_{k=k_0}^{\infty} l_0 2^{-k} = C l_0 2^{-k_0} \le C\,|Q|
\end{aligned}$$

where we have used in the first inequality the fact that the squares of equal size are almost disjoints. $\square$

Lemma 4.4. *Fix a square* Q. *Let* $\mathcal{F} = \{Q_j : j \in \Gamma\}$ *a finite family of squares of dyadic size and such that the squares* $\widetilde{Q}_j$, $j \in \Gamma$, *of the same size are almost disjoints, i.e., there exists a constant* A *such that for all* l

$$\sum_{\{i\in\Gamma: l(Q_i)=l\}} \chi_{\widetilde{Q_i}} \le A.$$

Assume that $\widetilde{Q_j}$ *is not included in* $\widetilde{Q_i}$ *for all* $j, i \in \Gamma$, $j \neq i$. *Let* $\Gamma_0 = \{j \in \Gamma : (\widetilde{Q_j})^+ \cap (\widetilde{Q})^+ \neq \emptyset, |Q_j| < |Q|\}$. *Then there exists a constant* C *depending only on* A *such that*

$$\sum_{j\in\Gamma_0} \left|\widetilde{Q_j}\right| \le C\,|Q|\,.$$

Proof. Let $\Gamma_1 = \{j \in \Gamma_0 : \left(\widetilde{Q_j} \cup (\widetilde{Q_j})^+\right) \cap \partial(\widetilde{Q})^+ \neq \emptyset\}$ and $\Gamma_2 = \{j \in \Gamma_0 : \widetilde{Q_j} \cup (\widetilde{Q_j})^+ \subset (\widetilde{Q})^+\}$. It is clear that $\Gamma_0 = \Gamma_1 \cup \Gamma_2$. Then, applying Lemma (4.3) with $Q = (\widetilde{Q})^+$, we can see that $\sum_{j\in\Gamma_1} \left|\widetilde{Q_j}\right| \leq C\,|Q|$. Now, we will see that $\sum_{j\in\Gamma_2} \left|\widetilde{Q_j}\right| \leq C\,|Q|$. We consider the maximal squares $\widetilde{Q_i}$ in the sense of Definition 4.1 for the family $\{\widetilde{Q_j} : j \in \Gamma_2\}$. Then

$$\sum_{j\in\Gamma_2} \left|\widetilde{Q_j}\right| \leq \sum_{\substack{i\in\Gamma_2: \\ \widetilde{Q_i}\max}} \left(\sum_{\substack{j\in\Gamma_2 \\ \widetilde{Q_i}\cap\widetilde{Q_j}\neq\emptyset \\ |\widetilde{Q_j}|<|\widetilde{Q_i}|}} \left|\widetilde{Q_j}\right| + |\widetilde{Q_i}| \right)$$

Since $\widetilde{Q_j}$ is not included in $\widetilde{Q_i}$, we have that

$$\{j \in \Gamma_2 : \widetilde{Q_j} \cap \widetilde{Q_i} \neq \emptyset, |\widetilde{Q_j}| < |\widetilde{Q_i}|\} = \{j \in \Gamma_2 : \widetilde{Q_j} \cap \partial\widetilde{Q_i} \neq \emptyset, |\widetilde{Q_j}| < |\widetilde{Q_i}|\}.$$

Now, applying Lemma 4.3 with $Q = \widetilde{Q_i}$, we get

$$\sum_{\substack{j\in\Gamma_2 \\ \widetilde{Q_i}\cap\widetilde{Q_j}\neq\emptyset \\ |\widetilde{Q_j}|<|\widetilde{Q_i}|}} \left|\widetilde{Q_j}\right| \leq C|\widetilde{Q_i}|.$$

Since the squares $\widetilde{Q_i}$ are almost disjoint and $\widetilde{Q_i} \subset (\widetilde{Q})^+$ we obtain that

$$\begin{aligned} \sum_{j\in\Gamma_2} \left|\widetilde{Q_j}\right| &\leq C \sum_{\substack{i\in\Gamma_2: \\ \widetilde{Q_i}\max}} \left|\widetilde{Q_i}\right| \\ &\leq C \left|(\widetilde{Q})^+\right| = C\,|Q|\,. \end{aligned}$$

□

Proof of Lemma 3.1. We shall do two selections.

First selection: Let B_1 be the set of points of A such that their second coordinate is the biggest one among the second coordinates of the points in A. Let x_{i_1} be the point of B_1 with the smaller first coordinate. Assume that $x_{i_1}, \ldots, x_{i_k}$ have been chosen. We define $A_{k+1} = A \setminus \bigcup_{j=1}^{k} \widetilde{Q_{i_j}}$. If $A_{k+1} = \emptyset$ then we do not choose more points. If $A_{k+1} \neq \emptyset$ then we consider the set B_{k+1} of points of A_{k+1} such that the second coordinate is the biggest one among the second coordinates of the points in A_{k+1} and we choose $x_{i_{k+1}}$ as the point in B_{k+1} with the smaller first coordinate. Since we have a finite number of squares the process stops. Let $\Delta = \{i : x_i \text{ was chosen at some moment}\}$. Then we have the following properties:

(a) $A \subset \bigcup_{i\in\Delta} \widetilde{Q_i}$

(b) $\dfrac{1}{|Q_i|} \displaystyle\int_{(\widetilde{Q_i})^+} f > \dfrac{\lambda}{4}$ for all $i \in \Delta$.

(c) If $l > 0$ and Q_j and Q_k are two squares with $j, k \in \Delta$ and $l(Q_j) = l(Q_k) = l$ then the norm $\|x_j - x_k\|_\infty$ is greater than $l/2$.

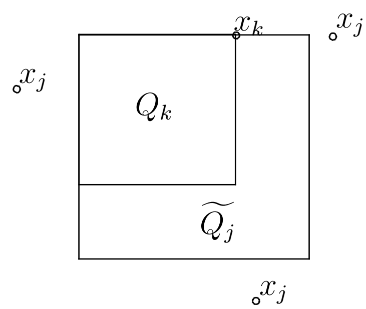

FIGURE 3. Possible places for x_j if x_k was chosen before than x_j

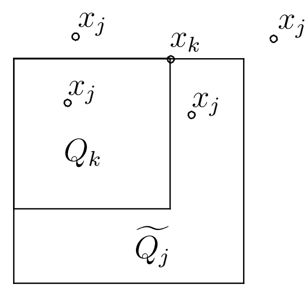

FIGURE 4. Impossible places for x_j if x_k was chosen before than x_j

(d) The squares $\widetilde{Q_i}$, $i \in \Delta$, of the same size are almost disjoints, i.e., there exists a constant C such that for all l

$$\sum_{\{i\in\Delta: l(Q_i)=l\}} \chi_{\widetilde{Q_i}} \leq C$$

($C = 36$ is valid).

Property (a) is clear. In fact, if A is not included in $\bigcup_{i\in\Delta} \widetilde{Q_i}$ then $A \setminus \bigcup_{i\in\Delta} \widetilde{Q_i} \neq \emptyset$ and the process would continue. Property (b) follows from $Q_i^{+2} \subset (\widetilde{Q_i})^+$ and the assumption in the lemma. In order to see that property (c) holds we may assume that the point x_k was chosen before x_j. Then (c) follows since the second coordinate of x_k is bigger or equal than the second coordinate of x_j, $x_j \notin \widetilde{Q_k}$ and $l(Q_k) = l$ (see Figures 3 and 4).

Finally, property (d) follows from (c). Let us fix a square $\widetilde{Q_{j_0}} = \left[z_{j_0} - \frac{3}{2}l, z_{j_0}\right] \times \left[y_{j_0} - \frac{3}{2}l, y_{j_0}\right]$, of the selection, i.e., (z_{j_0}, y_{j_0}) is the upper-right corner. Then if $\widetilde{Q_j}$ is a square of the selection with $l(\widetilde{Q_j}) = \frac{3}{2}l$ and $\widetilde{Q_j} \cap \widetilde{Q_{j_0}} \neq \emptyset$ then the upper-right corner (z_j, y_j) of $\widetilde{Q_j}$ belongs to the square $R_{j_0} = \left[z_{j_0} - \frac{3}{2}l, z_{j_0} + \frac{3}{2}l\right] \times \left[y_{j_0} - \frac{3}{2}l, y_{j_0} + \frac{3}{2}l\right]$. Now, we divide R_{j_0} in 36 disjoint and equal squares of size $\frac{l}{2}$. The point (z_j, y_j) belongs only to one of those squares. Therefore, it follows from (c) that there are not more than 36 squares $\widetilde{Q_j}$ in the collection (one of them is $\widetilde{Q_{j_0}}$) of size $l\left(\widetilde{Q_j}\right) = l\left(\widetilde{Q_{j_0}}\right)$ that intersect $\widetilde{Q_{j_0}}$. This proves (d).

Second selection: Let $\Gamma_1 = \Delta$. We take $\widetilde{Q_{i_1}} \in \{\widetilde{Q_i} : i \in \Gamma_1\}$ such that $l(\widetilde{Q_{i_1}}) = \max\{l(\widetilde{Q_i}) : i \in \Gamma_1\}$. Once we have chosen $\widetilde{Q}_1, \ldots, \widetilde{Q}_k$, we consider

$$\Gamma_{k+1} = \{i \in \Delta : \text{ such that } \widetilde{Q_i} \text{ is not contained in } \widetilde{Q_l} \text{ for } l = 1, \ldots, k\}.$$

If $\Gamma_{k+1} = \emptyset$ we do not choose more squares. If $\Gamma_{k+1} \neq \emptyset$ we choose $\widetilde{Q_{k+1}} \in \{\widetilde{Q_i} : i \in \Gamma_{k+1}\}$ such that $l(\widetilde{Q_{i_{k+1}}}) = \max\{l(\widetilde{Q_i}) : i \in \Gamma_{k+1}\}$. Since we have a finite number of squares, the process stops in a finite number of steps. Let

$$\Gamma = \{i \in \Delta : \text{ such that } \widetilde{Q_i} \text{ was chosen in one of the steps above }\}.$$

The family $\{\widetilde{Q_i} : i \in \Gamma\}$ has the properties(a)-(d) and it satisfies also the following property:

(e) If $i, j \in \Gamma$, $i \neq j$, then $\widetilde{Q_j} \nsubseteq \widetilde{Q_i}$.

We have already proved the first part of the lemma. Now we assume

$$\frac{1}{|Q_j|} \int_{(\widetilde{Q_j})^+} f \leq 8\lambda, \tag{17}$$

and we proceed to select the family $\{F_j\}$. We can apply Lemma 4.4 to the family $\{Q_j : j \in \Gamma\}$. Fixed $j \in \Gamma$ and $\Gamma_j = \{i \in \Gamma : (\widetilde{Q_i})^+ \cap (\widetilde{Q_j})^+ \neq \emptyset, |Q_i| < |Q_j|\}$ we have by Lemma 4.4 that

$$\sum_{i \in \Gamma_j} |Q_i| \leq \sum_{i \in \Gamma_j} |\widetilde{Q_i}| \leq C|Q_j|.$$

Consequently,

$$\begin{aligned} \sum_{i \in \Gamma_j} \int_{(\widetilde{Q_i})^+} f &\leq 8\lambda \sum_{i \in \Gamma_j} |Q_i| \\ &\leq 8C\lambda |Q_j| \\ &< 32C \int_{(\widetilde{Q_j})^+} f. \end{aligned}$$

Therefore, we have proved that there exists a natural number N such that

$$\sum_{i \in \Gamma_j} \int_{(\widetilde{Q_i})^+} f \leq N \int_{(\widetilde{Q_j})^+} f, \tag{18}$$

where N is independent of f, λ and j. Let s be the number of elements of Γ. If $s \leq 2N$ we choose $F_j = (\widetilde{Q_j})^+$ and there is nothing to prove. Suppose $s > 2N$. We define for each n, $1 \leq n \leq s$, a subset E_n^j of $(\widetilde{Q_j})^+$, in the following way:

$$E_n^j = \{x \in (\widetilde{Q_j})^+ : \text{there exist at least } n \text{ numbers } i \in \Gamma \text{ such that } x \in (\widetilde{Q_i})^+ \text{ and } |Q_i| < |Q_j|\}$$

Clearly $E_{n+1}^j \subset E_n^j$. Moreover, for $x \in \left(\widetilde{Q_j}\right)^+$

$$\sum_{n=1}^{s} \chi_{E_n^j}(x) \leq \sum_{i \in \Gamma_j} \chi_{(\widetilde{Q_i})^+}(x). \tag{19}$$

In fact, if $\sum_{n=1}^{s} \chi_{E_n^j}(x) = k$ then $x \in (\widetilde{Q_i})^+$ for k indexes i with $|Q_i| < |Q_j|$. But these k indexes belong to Γ_j; therefore $\sum_{i \in \Gamma_j} \chi_{(\widetilde{Q_i})^+}(x) \geq k$ and the inequality is true. Now, by (18) and (19) we get

$$\begin{aligned}\sum_{n=1}^{s}\int_{E_n^j} f(x)\,dx &\leq \int_{(\widetilde{Q_j})^+} f(x)\left(\sum_{i\in\Gamma_j}\chi_{(\widetilde{Q_i})^+}(x)\right)dx\\ &\leq \sum_{i\in\Gamma_j}\int_{(\widetilde{Q_i})^+} f \leq N\int_{(\widetilde{Q_j})^+} f.\end{aligned}$$

This inequality, $s>2N$ and $\int_{E_{n+1}^j} f \leq \int_{E_n^j} f$ give

$$2N\int_{E_{2N}^j} f \leq \sum_{n=1}^{s}\int_{E_n^j} f \leq N\int_{(\widetilde{Q_j})^+} f.$$

Therefore,

$$\int_{E_{2N}^j} f \leq \frac{1}{2}\int_{(\widetilde{Q_j})^+} f.$$

From this last inequality, if we define $F_j = (\widetilde{Q_j})^+ - E_{2N}^j$, we get

$$\int_{F_j} f \geq \frac{1}{2}\int_{(\widetilde{Q_j})^+} f.$$

Thus

$$\frac{1}{|Q_j|}\int_{F_j} f \geq \frac{1}{2|Q_j|}\int_{(\widetilde{Q_j})^+} f > \frac{\lambda}{8}.$$

It only remains to prove that the sets F_j are almost disjoint. We will prove

$$\sum_{j\in\Gamma}\chi_{F_j}(x) \leq 72N. \tag{20}$$

Let $x\in\cap_{i=1}^{k}F_{j_i}$, with $j_i\in\Gamma$. We will show that $k\leq 72N$. Since $F_{j_i}\subset(\widetilde{Q_{j_i}})^+$ we have $x\in(\widetilde{Q_{j_i}})^+$. Consider all the squares $(\widetilde{Q_{j_i}})^+$ of maximum radio (there cannot be more than 36). For any of them (suppose $(\widetilde{Q_{j_{i_0}}})^+$) there can be no more than $2N$ of smaller size such that x belongs to those squares. For suppose this were not true, then x belongs to s squares with $s>2N$ of size smaller than $(\widetilde{Q_{j_{i_0}}})^+$ and therefore $x\in E_s^{j_{i_0}}\subset E_{2N}^{j_{i_0}}$, a contradiction since $x\in F_{j_{i_0}} = (\widetilde{Q_{j_{i_0}}})^+ - E_{2N}^{j_{i_0}}$. Therefore (20) follows.

□

5. Proof of Theorem 1.5

To prove Theorem 1.5 we need some facts and to introduce some notation. For each $t\in\mathbb{R}^2$ we consider the measures μ_t defined by

$$\mu_t(E) = \mu(\tau^t(E)).$$

These measures have the same sets of measure zero than μ since the transformations are null-preserving. If H_t is the Radon-Nikodym derivative of μ_t with respect to μ then

$$\mu(E) = \int_X (T^t\chi_E)H_t\,d\mu$$

and

$$H_{t+s} = (T^tH_s)H_t. \tag{21}$$

It follows that the operators $S^t f = H_t T^t f$ are isometries in $L^1(\mu)$. Consequently, by using [6, Lemma III.11.6] we may assume without loss of generality that $H_t(x)$ is measurable with respect to the completion of the σ-algebra product (see also [15]).

Proof of Theorem 1.5. Since $\mathcal{T}$ is Cesàro bounded in $L^1(\mu)$, we have by Tonelli's Theorem that

$$\frac{1}{h^2}\int_0^h\int_0^h\int_X T^{t-s}f(x)\,d\mu\,dt \le C\int_X T^{-s}f(x)\,d\mu$$

for every $h>0$, all $s=(s_1,s_2)\in\mathbb{R}^2$ and each measurable function $f\ge 0$. But

$$\int_X T^t f(x)\,d\mu = \int_X f(x)H_{-t}(x)\,d\mu.$$

Therefore

$$\int_X f(x)\left(\frac{1}{h^2}\int_0^h\int_0^h H_{s-t}(x)\,dt\right)\,d\mu \le C\int_X f(x)H_s(x)\,d\mu$$

for all nonnegative measurable function f, which implies

$$\frac{1}{h^2}\int_{s_1-h}^{s_1}\int_{s_2-h}^{s_2} H_t(x)\,dt\,d\mu \le CH_s(x) \quad \text{a.e. } x.$$

It follows that for almost every $x\in X$

$$\frac{1}{h^2}\int_{s_1-h}^{s_1}\int_{s_2-h}^{s_2} H_t(x)\,dt\,d\mu \le CH_s(x) \quad \text{for a.e. } s=(s_1,s_2),$$

or, in other words, for almost every x the functions $t\to H_t(x)$ satisfy $A_1^+(\mathbb{R}^2)$ with a constant independent of x. Now we obtain the weak type $(1,1)$ inequality by transference arguments.

We can assume that $f\ge 0$. For each $\eta>0$, let us consider $M_\eta^+ f(x) = \sup_{0<h\le\eta} A_h f(x)$. Let $\lambda>0$ and $E_\lambda=\{x\in X : M_\eta^+ f(x)>\lambda\}$. Let us fix $R>0$. Then

$$\begin{aligned}\mu(E_\lambda) &= \frac{1}{R^2}\int_0^R\int_0^R\int_X T^t\chi_{E_\lambda}(x)H_t(x)d\mu(x)dt\\ &= \int_X \frac{1}{R^2}\int_0^R\int_0^R T^t\chi_{E_\lambda}(x)H_t(x)d\mu(x)dt.\end{aligned} \tag{22}$$

If we define $g^x(t)=T^tg(x)$, we have that if $R>0$, $t=(t_1,t_2)$, $t_1\le R$, $t_2\le R$ and $T^t\chi_{E_\lambda}(x)=1$ then $M^+(f^x\chi_{[0,R+\eta]\times[0,R+\eta]})(t)>\lambda$. Therefore

$$\mu(E_\lambda)\le\int_X\frac{1}{R^2}\int_{\{t:M^+(f^x\chi_{[0,R+\eta]\times[0,R+\eta]})(t)>\lambda\}} H_t(x)d\mu(x)dt. \tag{23}$$

Since, for almost every x, the functions $t \to H_t(x)$ satisfy $A_1^+(\mathbb{R}^2)$ with a constant independent of x we obtain by Theorem 1.2 that the last term is dominated by

$$
\begin{aligned}
\frac{C}{\lambda}\int_X \frac{1}{R^2}\int_0^{R+\eta}\int_0^{R+\eta} & f^x(t)H_t(x)dtd\mu(x) \\
&= \frac{C}{\lambda}\frac{1}{R^2}\int_0^{R+\eta}\int_0^{R+\eta}\int_X T^t f(x)H_t(x)d\mu(x)dt \\
&= \frac{C}{\lambda}\frac{1}{R^2}\int_0^{R+\eta}\int_0^{R+\eta}\int_X f(x)d\mu(x)dt \\
&= \frac{C}{\lambda}\left(\frac{R+\eta}{R}\right)^2\int_X f(x)d\mu(x).
\end{aligned} \tag{24}
$$

Letting R go to infinity we obtain

$$\mu(E_\lambda) \le \frac{C}{\lambda}\int_X f(x)d\mu(x).$$

Letting η tend to infinity we obtain the inequality that we wished to prove. □

Final Remarks 5.1. *It is clear that $M_{\mathcal{T}}$ is bounded in $L^\infty(\mu)$. Therefore, under the assumptions in Theorem 1.5 we have that $M_{\mathcal{T}}$ is of strong type (p,p) for all $p>1$. It can be proved from this fact that $A\oplus B$ is dense in $L^p(\mu)$, where*

$$A = \{f \in L^p(\mu) : T^t f = f \text{ for all } t=(t_1,t_2), t_1,t_2>0\}$$

and B is the linear manifold generated by

$$\{f - T^t f : f \in L^p(\mu), t=(t_1,t_2), t_1,t_2>0\}.$$

It is clear that the averages $A_h f$ converge a.e. as $h \to +\infty$ for all f in this dense set. Therefore, by the weak type $(1,1)$ inequality and the strong type inequality (p,p), $p>1$, we have that $A_h f$ converge a.e. as $h\to+\infty$ for all $f \in \cup_{p\ge 1}L^p(\mu)$ (a detailed proof in the one dimensional case can be seen in [3]*).*

It is worth noting that there exist flows which are Cesàro bounded in $L^1(\mu)$. In order to see this, consider a flow of measure preserving transformations, that is, $\mu(\tau^t E) = \mu(E)$. Let us take the ergodic maximal operator

$$\mathcal{N}_{\mathcal{T}} f(x) = \sup_{h>0}\frac{1}{h^2}\left|\int_0^h\int_0^h f(\tau^{-t}x)\,dt\right|.$$

It is known that $\mathcal{N}$ is of weak type $(1,1)$ and of strong type (p,p) for $p>1$. Fix p, $1<p<\infty$, and a positive function $g \in L^p(\mu)$. Let A be a constant such that

$$||\mathcal{N}_{\mathcal{T}} f||_{L^p(\mu)} \le A||f||_{L^p(\mu)}$$

and define

$$w = \sum_{i=0}^\infty \frac{\mathcal{N}_{\mathcal{T}}^{(i)} g}{(2A)^i},$$

where $\mathcal{N}_{\mathcal{T}}^{(i)}$ is the ith-iteration of $\mathcal{N}_{\mathcal{T}}$. It is clear that $w \in L^p(\mu)$, $||w||_{L^p(\mu)} \le 2||g||_{L^p(\mu)}$, $g \le w$ and $\mathcal{N}_{\mathcal{T}} w \le 2Aw$ a.e.. Consider now the measure $\tilde\mu = w\,d\mu$. The last property of w implies that the flow is Cesàro bounded in $L^1(\tilde\mu)$ and it is clear that the transformations τ^t are null-preserving transformations with respect to $\tilde\mu$. Further, if we have that for some $t\in\mathbb{R}^2$ the transformation τ^t is ergodic and the *function $g \notin L^\infty(\mu)$ then the operators T^t are not contractions, more over they are not power bounded, that is, there is not a positive constant such that for all t*

$$\int_X |T^t f|\, d\tilde{\mu} \le C \int_X f\, d\tilde{\mu}.$$

A more detailed discussion in the one-dimensional case for the two-sided case can be found in [9].

Finally, we point out that Theorem 1.5 remains true for $1 < p < \infty$, that is, if the group $\mathcal{T}$ is Cesàro bounded in $L^p(\mu)$, $1 < p < \infty$, which means that there exists C such that

$$\sup_{h>0} \int_X |A_h f|^p \; d\mu \le C \int_X |f|^p \, d\mu. \tag{25}$$

for all measurable function $f \geq 0$, then there is a positive constant C such that

$$\mu(\{x \in X : M_{\mathcal{T}} f(x) > \lambda\}) \le \frac{C}{\lambda^p} \int_X |f|^p \, d\mu$$

for all $\lambda > 0$ and all $f \in L^p(\mu)$. The weak type inequality follows by transference arguments, as in the case $p = 1$, by using that the functions $t \to H_t(x)$ satisfy $A_p^+(\mathbb{R}^2)$ for almost every x with an uniform constant. All we have to show is that (25) *implies that the functions $t \to H_t(x)$ satisfy $A_p^+(\mathbb{R}^2)$ for almost every x. The proof is similar to the one-dimensional case* [3], *it uses the ideas of the factorization of weights* [5, 7] *but is not so direct as the corresponding one for $p = 1$.*

References

[1] K. Andersen and B. Muckenhoupt *Weighted weak type Hardy inequalities with applications to Hilbert transforms and maximal functions*, Studia Math. 72 (1982), no. 1, 9–26.

[2] E, Berkson and T.A. Gillespie, *Mean-boundedness and Littlewood-Paley for separation-preserving operators*, Trans. Amer. Math. Soc. 349 (1997), no. 3, 1169–1189.

[3] A.L. Bernardis, M. Lorente, F.J. Martín-Reyes, M.T. Martínez and A. de la Torre., *Differences of ergodic averages for Cesro bounded operators*, to appear in The Quarterly Journal of Mathematics.

[4] R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals*, Studia Math, 51 (1974), 241–250

[5] R. Coifman, P.W. Jones and J.L. Rubio de Francia, *Constructive decomposition of BMO functions and factorization of A_p weights*, Proc. Amer. Math. Soc. 87 (1983), no. 4, 675–676.

[6] N. Dunford and J.T. Schwartz, *Linear Operators. I.*, Pure and Applied Mathematics, Vol. 7 Interscience Publishers, Inc., New York; Interscience Publishers, Ltd., London 1958.

[7] J. García-Cuerva and J.L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland Mathematics Studies, 116. Notas de Matemática [Mathematical Notes], 104. North-Holland Publishing Co., Amsterdam, 1985.

[8] T.A. Gillespie and J.L. Torrea, *Weighted ergodic theory and dimension free estimates.*, Q. J. Math. 54 (2003), no. 3, 257–280.

[9] M. Lorente and F.J. Martín-Reyes *The ergodic Hilbert transform for Cesro bounded flows*, Tohoku Math. J. (2) 46 (1994), no. 4, 541–556.

[10] F.J. Martín-Reyes, *On the one-sided Hardy-Littlewood maximal function in the real line and in dimensions greater than one*, Fourier analysis and partial differential equations (Miraflores de la Sierra, 1992), 237–250, Stud. Adv. Math., CRC, Boca Raton, FL, 1995.

[11] F.J. Martín-Reyes and A. de la Torre, *The dominated ergodic estimate for mean bounded, invertible, positive operators*, Proc. Amer. Math. Soc. 104 (1988), no. 1, 69–75.

[12] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. 165 (1972), 207–226.

[13] S. Ombrosi, *Weak weighted inequalities for a dyadic one-sided maximal function in* $\mathbb{R}^n$, Proc. Amer. Math. Soc. 133 (2005), no. 6, 1769–1775.
[14] R. Sato, *A remark on the ergodic Hilbert transform*, Math. J. Okayama Univ. 28 (1986), 159–163 (1987).
[15] R. Sato, *An ergodic maximal equality for nonsingular flows and applications*, Acta Math. Hungar. 61 (1993), no. 1-2, 51–58.
[16] E. Sawyer, *Weighted inequalities for the two-dimensional Hardy operator*, Studia Math. 82 (1985), no. 1, 1–16.
[17] E. Sawyer, *Weighted inequalities for the one-sided Hardy-Littlewood maximal functions*, Trans. Amer. Math. Soc. 297 (1986), no. 1, 53–61.
[18] E. Stein and R. Shakarchi, *Real analysis. Measure theory, integration, and Hilbert spaces*, Princeton Lectures in Analysis, III. Princeton University Press, Princeton, NJ, 2005.

(Liliana Forzani) IMAL-CONICET, Facultad de Ingeniería Química, Universidad Nacional del Litoral, Güemes 3450, 3000 Santa Fe, Argentina
E-mail address: liliana.forzani@gmail.com

(Francisco J. Martín-Reyes) Departamento de Análisis Matemático, Facultad de Ciencias, Universidad de Málaga, 29071-Málaga, Spain
E-mail address: martin_reyes@uma.es

(Sheldy J. Ombrosi) Departamento de Matemática, Universidad NAcional del Sur, Bahía Blanca, 8000, Argentina
E-mail address: sombrosi@uns.edu.ar