FEFFERMAN-STEIN INEQUALITIES FOR THE HARDY-LITTLEWOOD MAXIMAL FUNCTION ON THE INFINITE ROOTED k-ARY TREE

SHELDY OMBROSI, ISRAEL P. RIVERA-RÍOS, AND MARTÍN D. SAFE

Abstract. In this paper weighted endpoint estimates for the Hardy-Littlewood maximal function on the infinite rooted k-ary tree are provided. Motivated by Naor and Tao [20] the following Fefferman-Stein estimate

$$w\left(\{x\in T\,:\,Mf(x)>\lambda\}\right)\leq c_s\frac{1}{\lambda}\int_T|f(x)|M(w^s)(x)^{\frac{1}{s}}dx\qquad s>1$$

is settled and moreover it is shown it is sharp, in the sense that it does not hold in general if $s=1$. Some examples of non trivial weights such that the weighted weak type $(1,1)$ estimate holds are provided. A strong Fefferman-Stein type estimate and as a consequence some vector valued extensions are obtained. In the Appendix a weighted counterpart of the abstract theorem of Soria and Tradacete on infinite trees [32] is established.

1. Intoduction and main results

The centered Hardy-Littlewood maximal function on $\mathbb{R}^d$ is defined as

$$Mf(x)=\sup_{r>0}\frac{1}{|B(x,r)|}\int_{B(x,r)}|f(y)|dy.$$

Due to fact that the Lebesgue measure is doubling, namely, since

$$|B(x,2r)|\leq 2^d|B(x,r)|$$

it is not hard to check that $Mf(x)\simeq M^uf(x)$ where

$$M^uf(x)=\sup_{x\in B}\frac{1}{|B|}\int_B|f(y)|dy$$

and B is any ball. Furthermore we may replace balls by cubes with their sides parallel to the axes.

This operator was shown to be bounded on L^p and of weak type $(1,1)$ by Hardy and Littlewood [12] in the case $d=1$ and by Wiener [37] for the case $d\geq 1$. In a pioneering work by Fefferman and Stein [11] the following two weights inequality was provided

$$w\left(\left\{x\in\mathbb{R}^d\,:\,Mf(x)>t\right\}\right)\lesssim_d\frac{1}{t}\int_{\mathbb{R}^d}|f(x)|Mw(x)dx. \tag{1.1}$$

Inequality (1.1) is important for several reasons. The first of them is that it was a cornerstone to provide vector valued extensions. Another fundamental reason is that it was a precursor of the theory of weights that was continued later by the seminal work by Muckenhoupt [19]. We recall that in the classical setting $w\in A_1$ if $\left\|\frac{Mw}{w}\right\|_{L^\infty}<\infty$. Since, in general, $w\leq Mw$,

The first and the second author were partially supported by grant PIP (CONICET) 11220130100329CO. The third author was partially supported by ANPCyT PICT 2017-13152 and UNS Grant PGI 24/L115.

this condition implies that actually $w \simeq Mw$. Note that if $w \in A_1$ from (1.1) it readily follows that

$$w\left(\left\{x \in \mathbb{R}^d : Mf(x) > t\right\}\right) \lesssim_d \frac{1}{t}\int_{\mathbb{R}^d} |f(x)|w(x)dx.$$

At this point we would also like to note that Fefferman and Stein noted in [11] that $w \in A_1$ is a necessary condition for this inequality to hold. Since those works, the theory of weights and more in particular Fefferman-Stein inequalities and related variants have been studied in a variety of contexts [23, 36, 2, 27] and for singular integrals [8, 24, 28, 10, 5, 16] and their commutators [25, 26, 15]. See also [17, 3, 7].

The Hardy-Littlewood maximal operator in metric measure spaces has been mainly studied in the doubling setting (see [13]). In the case of non-doubling spaces, results for a suitable modification of the maximal operator were provided in [21, 30, 35]. It is worth mentioning that since Bourgain's seminal work [4] a number of papers, such as [34, 18, 1, 6], have been devoted to the study of discrete versions of operators in harmonic analysis.

In [20] Naor and Tao study the connection between the doubling condition and the maximal function in metric measure spaces. They provide a deep localization theorem for the maximal function and introduce the n-microdoubling property and use it to provide some interesting consequences. Among them they recover the classical result by Strömberg and Stein [33] $\|M\|_{L^1\to L^{1,\infty}} \lesssim n\log(n)$ in the general context of metric spaces satisfying the aforementioned n-microdubling property.

Having in mind results such as the Strömberg and Stein bound mentioned above, one may tend to think that there is always a connection between the doubling condition of the space and the weak type $(1,1)$ of the maximal function. However Naor and Tao show, in some sense, that it is not the case. They provide an example, the infinite rooted k-ary tree, for which even in complete absence of the doubling condition, the weak-type $(1,1)$ for the centered maximal function holds (see Theorem A a few lines below).

Given $k \geq 2$ we will denote by T_k the infinite rooted k-ary tree, namely, the infinite rooted tree such that each vertex has k children. We shall drop k and write just T in case there is no place to confusion. Abusing of notation, we will also use T to denote its vertex set. It is possible to define a metric measure space (T, d, μ) where d is the usual tree metric, namely $d(x, y)$ is the number of edges of the unique path between x and y, and μ is the counting measure defined on parts of the set of vertices. Abusing of notation, given $A \subset T$ we shall denote $|A| = \mu(A)$ and

$$\int_A f(x)dx = \sum_{x\in A} f(x).$$

We will also denote

$$M^\circ f(x) = \sup_{r\geq 0} \frac{1}{|S(x,r)|}\int_{S(x,r)} |f(y)|dy$$

where $S(x, r) = \{y \in T : d(x, y) = r\}$ denotes the sphere with center x and radius r. Note that, in contrast with the standard Euclidean setting, here it makes sense to consider this kind of maximal function because $S(x, r)$ are not sets of measure 0. For $k \geq 2$, we have that $M^\circ \simeq M$ as we will show in Proposition 2.1.

In the infinite rooted k-ary tree setting, covering arguments are essentially unavailable since the doubling condition or even more generally the upper doubling condition on the measure introduced by Hytönen in [14] completely fail. Hence a different approach is required. Via a combinatorial argument, exploiting the "expander" or "non-amenability" properties of the infinite rooted k-ary tree, Naor and Tao managed to settle the following theorem.

Theorem A. If $k \geq 2$, then

$$|\{x \in T_k \,:\, M^\circ f(x) > \lambda\}| \leq \frac{c}{\lambda} \int_T |f(x)| dx$$

with c independent of k.

It is worth noting that this result can be deduced from a work of Rochberg and Taibleson [29], and that it was also established independently by Cowling, Meda, and Setti [9]. For $p > 1$ the strong type estimate was essentially settled by Nevo and Stein [22].

At this point we would like to mention works by Soria and Tradacete [31, 32] in which they study the connection between properties of the maximal function and properties of the underlying graphs. Furthermore [32, Theorem 4.1] is an abstract version of Theorem A.

The main purpose of this work is to get a variant of the Fefferman-Stein estimate for the Hardy-Littlewood maximal function on the infinite rooted k-ary tree, generalizing Theorem A. Most of Fefferman-Stein inequalities in a number of settings rely upon a suitable use of covering lemmas such as Calderón-Zygmund decomposition. In the infinite rooted k-ary tree setting, no regularity nor doubling condition is available, and hence other techniques are required. We will exploit the flexibility in the approach provided in [20] to obtain the following theorem.

Theorem 1.1. *Let $k \geq 2$ and $s > 1$. Then, for every weight $w \geq 0$ on T we have that*

$$w\left(\{x \in T \,:\, Mf(x) > \lambda\}\right) \leq c_s \frac{1}{\lambda} \int_T |f(x)| M_s w(x) dx$$

where $M_s w = M(w^s)^{\frac{1}{s}}$, c_s is independent of k, and $c_s \to +\infty$ when $s \to 1$.

At first sight, having in mind the estimate in the classical setting, one may wonder whether this estimate could be improved to match (1.1). However, this is not the case. Not only it is not possible to choose $s = 1$ but actually it is not even possible to choose any number of iterations of the maximal function for the inequality to hold.

Theorem 1.2. *Let $n \geq 1$. There exists a weight $w \sim Mw$ and a sequence $f_j \in L^1(M^n w)$ such that*

$$w\left(\{Mf_j(x) > 1\}\right) \geq c_j \int_T |f_j(x)| M^n w(x) dx$$

where $M^n = M \circ \overset{n \text{ times}}{\cdots} \circ M$ and $c_j \to \infty$ when $j \to \infty$.

A direct consequence of the preceding theorem is that the fact that a weight w

$$Mw(x) \leq c_w w(x) \qquad \text{for all } x \in T$$

is not sufficient for $\|M\|_{L^1(w) \to L^{1,\infty}(w)} < \infty$.

On the other hand we have the following corollary of Theorem 1.1.

Corollary 1.3. *Let w be a weight such that there exists $s > 1$ for which*

$$M_s w(x) \leq c_{w,s} w(x) \qquad \text{for all } x \in T.$$

Then

$$w\left(\{x \in T : Mf(x) > \lambda\}\right) \lesssim \frac{1}{\lambda} \int_T |f(x)| w(x) dx.$$

We would like to observe that throughout the paper we deal with infinite rooted k-ary trees with $k \geq 2$. It is easy to check that in the case $k = 1$, Theorem 1.1 holds even for $s = 1$, since the the measure on the infinite rooted 1-ary tree is a doubling measure, and hence the classical theory works. Besides that, Proposition 2.1 does not hold for $k = 1$; furthermore, it is not hard to check that M° is not of weak type $(1,1)$ in this case.

The remainder of the paper is organized as follows. Section 2 is devoted to settling Theorem 1.1. In Section 3 we provide examples of non trivial weights that fulfill the assumptions of Corollary 1.3 and we settle Theorem 1.2. Section 4 is devoted to giving some vector valued extensions. We end up the paper with an Appendix devoted to providing a weighted counterpart of [32, Theorem 4.1].

2. Proof of Theorem 1.1

Before going into the proof of the theorem we present the following proposition.

Proposition 2.1. *Let $k \geq 2$ and $f \in L^1(T_k)$. Then*

$$Mf(x) \leq M^\circ f(x) \leq 2Mf(x).$$

Proof. Since every ball can be written as the disjoint union of spheres, we have the pointwise estimate

$$Mf(x) \leq M^\circ f(x).$$

For the other inequality, let $r \in \mathbb{N}$. We begin observing that

$$\begin{aligned}
\frac{|B(x,r)|}{|S(x,r)|} &= \frac{\sum_{j=0}^r |S(x,j)|}{|S(x,r)|} \\
&\simeq \frac{k^r + k^{r-1} + \cdots + 1}{k^r} = \sum_{j=0}^r \frac{1}{k^j} \leq 2.
\end{aligned}$$

Hence

$$\begin{aligned}
&\frac{1}{|S(x,r)|} \int_{S(x,r)} |f(y)| dy \\
&\leq \frac{|B(x,r)|}{|S(x,r)|} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)| dy \\
&\leq 2 \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)| dy
\end{aligned}$$

and this yields

$$M^\circ f(x) \leq 2Mf(x). \qquad \square$$

From the previous proposition, if we denote $M_s^\circ(w) = M^\circ(w^s)^{\frac{1}{s}}$, it readily follows that to settle Theorem 1.1 it suffices to show that

$$w\left(\{x \in T \,:\, M^\circ f(x) > \lambda\}\right) \leq c_s \frac{1}{\lambda} \int_T |f(x)| M_s^\circ w(x) dx, \tag{2.1}$$

for all $f \in L_1(T)$ and $\lambda > 0$.

We will denote by $\mathbb{1} \otimes w$ the product measure

$$\mathbb{1} \otimes w(A \times B) = |A| w(B) = \sum_{(x,y) \in A \times B} w(y), \qquad A, B \subset T.$$

The proof of Theorem 1.1 follows the scheme provided by Naor and Tao [20]. In particular [20, Lemma 5.1] is a key part of their proof. That lemma is obtained exploiting an expander and combinatorial argument that relies upon the symmetry of the infinite rooted k-ary tree. The role played by the fact that the measure on the space is the counting measure may seem relevant in the proof to provide a suitable "sharp" estimate. However, in the following lemma we overcome that difficulty providing a weighted version that contains a precise enough bound that allows us to push the scheme in [20].

Lemma 2.2. *Let E, F be finite subsets of T, $s > 1$ and let $r \geq 0$ be an integer. Then*

$$\mathbb{1} \otimes w\left(\{(x, y) \in E \times F : \; d(x, y) = r\}\right) \leq c_s k^{r \frac{s'}{s'+1}} w(F)^{\frac{1}{s'+1}} M_s^\circ w(E)^{\frac{s'}{s'+1}}$$

where $s' = \frac{s}{s-1}$ and c_s is a constant depending only on s.

Proof. We split the vertex set of the tree as $T = \bigcup_{j=0}^{\infty} T^j$, where T^j is the generation of the tree at depth j. We split as well accordingly, E and F. We define $E_j = E \cap T^j$ and $F_j = F \cap T^j$. An element in E_j and an element in F_i can be at distance exactly r, if and only if $i = j + r - 2m$ for some $m \in \{0, \dots, r\}$. Hence we can write

$$\begin{aligned} &\mathbb{1} \otimes w\left(\{(x, y) \in E \times F : d(x, y) = r\}\right) \\ &= \sum_{m=0}^{r} \sum_{\substack{i,j \in \mathbb{N} \cup \{0\} \\ i = j + r - 2m}} \mathbb{1} \otimes w\left(\{(x, y) \in E_j \times F_i : d(x, y) = r\}\right). \end{aligned} \tag{2.2}$$

Now we fix $m \in \{0, \dots, r\}$ and $i, j \in \mathbb{N} \cup \{0\}$ such that $i = j + r - 2m$. Note that if $x \in T^j$ and $y \in T_i$ are at distance r in T, then the m^{th} parent of x coincides with the $(r - m)^{th}$ parent of y. This leads to the fact that for each $y \in T^i$ there exist at most k^m elements of $x \in T^j$ with $d(x, y) = r$. From this it readily follows that

$$\mathbb{1} \otimes w\left(\{(x, y) \in E_j \times F_i : d(x, y) = r\}\right) \leq k^m w(F_i).$$

On the other hand note that for each $x \in T^j$ there are at most k^{r-m} elements of $y \in T^i$ with $d(x,y) = r$. Hence we have that for each $s > 1$,

$$\begin{aligned}
&\mathbb{1} \otimes w\left(\{(x,y) \in E_j \times F_i : d(x,y) = r\}\right) \\
&= \sum_{x \in E_j} \sum_{\substack{y \in F_i \\ d(x,y)=r}} w(y) \\
&= \sum_{x \in E_j} w(F_i \cap S(x,r)) \\
&\leq \sum_{x \in E_j} |F_i \cap S(x,r)|^{\frac{1}{s'}} w^s(F_i \cap S(x,r))^{\frac{1}{s}} \\
&= \sum_{x \in E_j} |F_i \cap S(x,r)|^{\frac{1}{s'}} k^{\frac{r}{s}} \frac{1}{k^{\frac{r}{s}}} w^s(F_i \cap S(x,r))^{\frac{1}{s}} \\
&\leq \sum_{x \in E_j} |F_i \cap S(x,r)|^{\frac{1}{s'}} k^{\frac{r}{s}} M_s^\circ(w)(x) \\
&\leq k^{\frac{m-r}{s'}} k^{\frac{r}{s}} M_s^\circ(w)(E_j).
\end{aligned}$$

Thus combining the ideas above

$$\mathbb{1} \otimes w\left(\{(x,y) \in E_j \times F_i : d(x,y) = r\}\right) \leq \min\left\{k^{r-\frac{m}{s'}} M_s^\circ w(E_j), k^m w(F_i)\right\}. \tag{2.3}$$

Taking into account (2.2) and (2.3), to end the proof it suffices to show that

$$\sum_{m=0}^{r} \sum_{\substack{i,j \in \mathbb{N}\cup\{0\} \\ i=j+r-2m}} \min\left\{k^{\frac{r}{s}+\frac{r-m}{s'}} M_s^\circ w(E_j), k^m w(F_i)\right\} \leq c_s k^{r\frac{s'}{s'+1}} w(F_i)^{\frac{1}{s'+1}} M_s^\circ w(E_j)^{\frac{s'}{s'+1}} \tag{2.4}$$

Let us define $c_j = \frac{M_s^\circ w(E_j)}{k^{\frac{j}{s'}}}$ and $d_j = \frac{w(F_j)}{k^j}$ for $j \geq 0$ and $c_j = d_j = 0$ for $j < 0$. Then,

$$\sum_{j=0}^{\infty} k^{\frac{j}{s'}} c_j = M_s^\circ w(E) \quad \text{and} \quad \sum_{j=0}^{\infty} k^j d_j = w(F), \tag{2.5}$$

and we have that whenever $i = j + r - 2m$,

$$\begin{aligned}
\min\left\{k^{\frac{r}{s}+\frac{r-m}{s'}} M_s^\circ w(E_j), k^m w(F_i)\right\} &= \min\left\{k^{\frac{r}{s}+\frac{r-m}{s'}} k^{\frac{j}{s'}} c_j, k^m k^i d_i\right\} \\
&= \min\left\{k^{r-\frac{r}{2s'}} k^{\frac{i+j}{2s'}} c_j, k^{\frac{r}{2}} k^{\frac{i+j}{2}} d_i\right\}.
\end{aligned}$$

Taking the identity above into account, settling (2.4) reduces to show that

$$\sum_{m=0}^{r} \sum_{\substack{i,j \in \mathbb{N}\cup\{0\} \\ i=j+r-2m}} \min\left\{k^{r-\frac{r}{2s'}} k^{\frac{i+j}{2s'}} c_j, k^{\frac{r}{2}} k^{\frac{i+j}{2}} d_i\right\} \leq c_s k^{r\frac{s'}{s'+1}} w(F_i)^{\frac{1}{s'+1}} M_s^\circ w(E_j)^{\frac{s'}{s'+1}}.$$

To prove this inequality, we fix a real parameter α to be chosen later, and argue as follows:

$$\begin{aligned}
&\sum_{m=0}^{r} \sum_{\substack{i,j\in\mathbb{N}\cup\{0\}\\ i=j+r-2m}} \min\left\{k^{r-\frac{r}{2s'}}k^{\frac{i+j}{2s'}}c_j, k^{\frac{r}{2}}k^{\frac{i+j}{2}}d_i\right\}\\
&\le \sum_{\substack{i,j\in\mathbb{N}\cup\{0\}\\ i<j+\alpha}} k^{r-\frac{r}{2s'}}k^{\frac{i+j}{2s'}}c_j + \sum_{\substack{i,j\in\mathbb{N}\cup\{0\}\\ i\ge j+\alpha}} k^{\frac{r}{2}}k^{\frac{i+j}{2}}d_i\\
&= k^{r-\frac{r}{2s'}}\sum_{j=0}^{\infty}\sum_{i\in\mathbb{N}\cup\{0\}:i<j+\alpha} k^{\frac{i+j}{2s'}}c_j + k^{\frac{r}{2}}\sum_{i=0}^{\infty}\sum_{j\in\mathbb{N}\cup\{0\}:j\le i-\alpha} k^{\frac{i+j}{2}}d_i\\
&\lesssim c_s k^{r-\frac{r}{2s'}}\sum_{j=0}^{\infty} k^{\frac{j}{s'}+\frac{\alpha}{2s'}}c_j + k^{\frac{r}{2}}\sum_{i=0}^{\infty}k^{i-\frac{\alpha}{2}}d_i\\
&\le c_s\left(k^{r-\frac{r}{2s'}}k^{\frac{\alpha}{2s'}}M_s^{\circ}w(E_j) + k^{\frac{r}{2}}k^{-\frac{\alpha}{2}}w(F_i)\right).
\end{aligned}$$

Now we provide some hints about how to optimize on α. Let $f_{a,b}(\alpha) = k^{\frac{\alpha}{2s'}}a + k^{-\frac{\alpha}{2}}b$ for $a, b > 0$. Note that $f_{a,b}$ reaches its absolute minimum at $\frac{2\log_k(\frac{b}{a})}{1+\frac{1}{s'}}$. Hence choosing $a_0 = k^{r-\frac{r}{2s'}}M_s^{\circ}w(E_j)$ and $b_0 = k^{\frac{r}{2}}w(F_i)$ and $\alpha_0 = \frac{2\log_k(\frac{b_0}{a_0})}{1+\frac{1}{s'}}$ we have that

$$\begin{aligned}
&\sum_{m=0}^{r} \sum_{\substack{i,j\in\mathbb{N}\cup\{0\}\\ i=j+r-2m}} \min\left\{k^{r-\frac{r}{2s'}}k^{\frac{i+j}{2s'}}c_j, k^{\frac{r}{2}}k^{\frac{i+j}{2}}d_i\right\}\\
&\qquad\le c_s f_{a_0,b_0}(\alpha_0)\\
&\qquad\le c_s\left(k^{r\frac{s'}{s'+1}}w(F_i)^{\frac{1}{s'+1}}M_s^{\circ}w(E_j)^{\frac{s'}{s'+1}} + k^{r\frac{s'}{s'+1}}w(F_i)^{\frac{1}{s'+1}}M_s^{\circ}w(E_j)^{\frac{s'}{s'+1}}\right)\\
&\qquad\le c_s k^{r\frac{s'}{s'+1}}w(F_i)^{\frac{1}{s'+1}}M_s^{\circ}w(E_j)^{\frac{s'}{s'+1}}
\end{aligned}$$

and hence we are done. □

For each $r \ge 0$, we denote by A_r° the spherical averaging operator

$$A_r^{\circ}f(x) = \frac{1}{|S(x,r)|}\sum_{y\in S(x,r)}|f(y)|.$$

Hence $M^{\circ}f(x) = \sup_{r\ge 0} A_r^{\circ}f(x)$. We can use Lemma 2.2 to obtain a distributional estimate on A_r°.

Lemma 2.3. *Let $r \ge 1$ and $\lambda > 0$. Then*

$$w\left(\{A_r^{\circ}f \ge \lambda\}\right) \lesssim c_s \sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\le 2^n\le 2k^r}} \left(\frac{2^n}{k^r}\right)^{\frac{1}{2s'}} 2^n M_s^{\circ}w\left(\left\{|f| \ge 2^{n-1}\lambda\right\}\right)$$

where c_s depends only on s and $c_s \to \infty$ when $s \to 1$.

Remark 2.4. We would like to note that the decay $\left(\frac{2^n}{k^r}\right)^{\frac{1}{2s'}}$ will be fundamental for our purposes. Note that in the case $s=1$ then we would not have this decay and, as we will see later, in the absence of that decay we would not be able to settle Theorem 1.1. At this point we would like to note as well that this inequality with a good enough decay in $\frac{2^n}{k^r}$ and $s=1$ cannot hold since that would contradict Theorem 1.2.

Proof of Lemma 2.3. We can assume without loss of generality f to be non-negative and $\lambda=1$. We bound

$$f \leq \frac{1}{2} + \sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\leq 2^n\leq k^r}} 2^n \chi_{E_n} + f\chi_{\{f\geq \frac{1}{2}k^r\}}, \tag{2.6}$$

where E_n is the sublevel set

$$E_n = \left\{2^{n-1} \leq f < 2^n\right\}. \tag{2.7}$$

Hence

$$A_r^\circ f \leq \frac{1}{2} + \sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\leq 2^n\leq k^r}} 2^n A_r^\circ\left(\chi_{E_n}\right) + A_r^\circ\left(f\chi_{\{f\geq \frac{1}{2}k^r\}}\right). \tag{2.8}$$

Since $|S(x,r)|\leq k^r$ we see that

$$\begin{aligned} w\left(A_r^\circ\left(f\chi_{\{f\geq \frac{1}{2}k^r\}}\right)\neq 0\right) &\leq w\left(\bigcup_{y\in\{f\geq\frac{1}{2}k^r\}} S(y,r)\right)\\ &\leq \sum_{y\in\{f\geq\frac{1}{2}k^r\}} w(S(y,r)) = |S(x,r)| \sum_{y\in\{f\geq\frac{1}{2}k^r\}} \frac{w(S(y,r))}{|S(x,r)|}\\ &\leq k^r M^\circ w\left(\left\{f\geq \frac{1}{2}k^r\right\}\right). \end{aligned} \tag{2.9}$$

Thus we have that combining the estimates above

$$\begin{aligned} w\left(A_r^\circ f \geq 1\right) &\leq w\left(\sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\leq 2^n\leq k^r}} 2^n A_r^\circ\left(\chi_{E_n}\right) \geq \frac{1}{2}\right) + w\left(A_r^\circ\left(f\chi_{\{f\geq\frac{1}{2}k^r\}}\right)\neq 0\right)\\ &\leq w\left(\sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\leq 2^n\leq k^r}} 2^n A_r^\circ\left(\chi_{E_n}\right) \geq \frac{1}{2}\right) + k^r M^\circ w\left(\left\{f\geq\frac{1}{2}k^r\right\}\right). \end{aligned}$$

Let β be a real parameter such that $0<\beta<1$ to be chosen later. Note that if

$$\sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\leq 2^n\leq k^r}} 2^n A_r^\circ\left(\chi_{E_n}\right) \geq \frac{1}{2}$$

then we necessarily have some $n \in \mathbb{N}$ such that $1 \le 2^n \le k^r$ for which

$$A_r^\circ(\chi_{E_n}) \ge \frac{2^\beta - 1}{2^{n+2}} \left(\frac{2^n}{k^r}\right)^\beta.$$

Indeed, otherwise we have that

$$\begin{aligned}
\frac{1}{2} &\le \sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\le 2^n\le k^r}} 2^n A_r^\circ(\chi_{E_n}) \le \frac{2^\beta - 1}{4k^{r\beta}} \sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\le 2^n\le k^r}} 2^{\beta n} \\
&\le \frac{(2^\beta - 1)}{4k^{r\beta}} \frac{(2^{\beta(\log_2 k^r + 1)} - 1)}{(2^\beta - 1)} = \frac{2^\beta k^{r\beta} - 1}{4k^{r\beta}} < \frac{2^\beta}{4} < \frac{1}{2}
\end{aligned}$$

which is a contraction. Thus

$$w(A_r^\circ f \ge 1) \le \sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\le 2^n\le k^r}} w(F_n) + k^r M^\circ w\left(\left\{f \ge \frac{1}{2}k^r\right\}\right)$$

where

$$F_n = \left\{A_r^\circ(\chi_{E_n}) \ge \frac{2^\beta - 1}{2^{n+2}} \left(\frac{2^n}{k^r}\right)^\beta\right\}.$$

Note that F_n is finite and observe that, since A_r° is a selfadjoint operator,

$$\begin{aligned}
&\frac{1}{k^r}\mathbb{1} \otimes w(\{(x,y) \in E_n \times F_n : d(x,y) = r\}) \\
&= \frac{1}{k^r} \sum_{x\in E_n} \sum_{\substack{y\in F_n\\ d(x,y)=r}} w(y) \simeq \int_T \chi_{E_n} A_r^\circ(w\chi_{F_n})(x)dx = \int_{F_n} w A_r^\circ(\chi_{E_n})(y)dy \\
&\ge w(F_n)\frac{2^\beta - 1}{2^{n+2}} \left(\frac{2^n}{k^r}\right)^\beta.
\end{aligned}$$

Now, using Lemma 2.2,

$$\frac{1}{k^r}\mathbb{1} \otimes w(\{(x,y) \in E_n \times F_n : d(x,y) = r\}) \le c_s k^{-\frac{r}{s'+1}} w(F_n)^{\frac{1}{s'+1}} M_s^\circ w(E_n)^{\frac{s'}{s'+1}}.$$

Hence

$$\begin{aligned}
&w(F_n)\frac{2^\beta - 1}{2^{n+2}} \left(\frac{2^n}{k^r}\right)^\beta \le c_s k^{-\frac{r}{s'+1}} w(F_n)^{\frac{1}{s'+1}} M_s^\circ w(E_n)^{\frac{s'}{s'+1}} \\
&\qquad \iff w(F_n)^{1-\frac{1}{s'+1}} \le c_s \frac{1}{2^\beta - 1} k^{\beta r - \frac{r}{s'+1}} 2^{(1-\beta)n} M_s^\circ w(E_n)^{\frac{s'}{s'+1}} \\
&\qquad \iff w(F_n)^{\frac{s'}{s'+1}} \le c_s \frac{1}{2^\beta - 1} k^{\beta r - \frac{r}{s'+1}} 2^{(1-\beta)n} M_s^\circ w(E_n)^{\frac{s'}{s'+1}} \\
&\qquad \iff w(F_n) \le c_s \frac{1}{(2^\beta - 1)^{\frac{s'+1}{s'}}} k^{r((s'+1)\beta - 1)\frac{1}{s'}} 2^{\frac{s'+1}{s'}(1-\beta)n} M_s^\circ w(E_n)
\end{aligned}$$

Choosing $\beta = \frac{1}{2(s'+1)}$ we have that

$$w(F_n) \leq c_s k^{-\frac{r}{2s'}} 2^{\frac{n}{2s'}} 2^n M_s^\circ w(E_n) \leq c_s \left(\frac{2^n}{k^r}\right)^{\frac{1}{2s'}} 2^n M_s^\circ w\left(\left\{f \geq 2^{n-1}\right\}\right).$$

Therefore

$$w(\{A_r^\circ f \geq 1\}) \leq c_s \sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\leq 2^n\leq k^r}} \left(\frac{2^n}{k^r}\right)^{\frac{1}{2s'}} 2^n M_s^\circ w\left(\left\{f \geq 2^{n-1}\right\}\right) + k^r M^\circ w\left(\left\{f \geq \frac{1}{2}k^r\right\}\right).$$

Since, in the right-hand side, the second term is dominated by the last term of the summation in the fist term, this yields the desired conclusion. □

Combining the ingredients above we are in the position to settle Theorem 1.1.

Proof. As we argued above, it suffices to settle (2.1). Since $M^\circ f = \sup_{r\geq 0} A_r^\circ f$, Lemma 2.3 implies that

$$\begin{aligned} w\left(M^\circ f \geq \lambda\right) &\leq \sum_{r=0}^{\infty} w\left(A_r^\circ f \geq \lambda\right) \\ &\leq c_s \sum_{r=0}^{\infty} \sum_{\substack{n\in\mathbb{N}\cup\{0\}\\ 1\leq 2^n\leq 2k^r}} \left(\frac{2^n}{k^r}\right)^{\frac{1}{2s'}} 2^n M_s^\circ w\left(|f| \geq 2^{n-1}\lambda\right) \\ &= c_s \sum_{x\in T} \sum_{n=0}^{\infty} \left(\sum_{\substack{r\in\mathbb{N}\cup\{0\}\\ k^r\geq 2^{n-1}}} \frac{1}{k^{\frac{r}{2s'}}} \right) 2^{n+\frac{n}{2s'}} \chi_{\{|f(x)|\geq 2^{n-1}\lambda\}}(x) M_s^\circ w(x) \\ &\lesssim c_s \sum_{x\in T} \sum_{n=0}^{\infty} 2^n \chi_{\{|f(x)|\geq 2^{n-1}\lambda\}} M_s^\circ w(x) \lesssim c_s \sum_{x\in T} \frac{1}{\lambda} |f(x)| M_s^\circ w(x). \end{aligned}$$

Hence (2.1) holds and the proof of Theorem 1.1 is complete. □

3. Examples of non trivial weights and the failure of the classical Fefferman-Stein estimate

3.1. **Radial weights.** A natural way to define a radial weight on the infinite rooted k-ary tree is the following. Let us consider

$$T = \bigcup_{j=0}^{\infty} T^j$$

where T^0 is the set whose only element is the root of the tree, T^1 is the set of vertices that are children of the root, and analogously T^j is the set of vertices that are children of vertices in T^{j-1}. Given this splitting, a radial weight can be defined as follows

$$w(x) = \sum_j c_j \chi_{T^j}(x) \qquad c_j \geq 0.$$

A natural question is trying to find choices of c_j such that

$$Mw(x) \lesssim w(x).$$

First of all, note that since $Mw(x) \simeq M^{\circ}w(x)$ it suffices to study the estimate for the latter. The problem would be to prove

$$\frac{1}{|S(x,r)|}\sum_{y\in S(x,r)} w(y) \leq \kappa w(x)$$

for some $\kappa > 0$ uniformly on r and on x. First note that $|S(x,r)| \simeq k^r$. Now arguing as Naor and Tao [20] we note that given $x \in T^j$ a vertex at distance exactly r belongs to T^i if and only if $i = j + r - 2m$ where $m \in \{0, \dots r\}$ and there are exactly k^{r-m} of such vertices. Hence if $x \in T^j$

$$\frac{1}{|S(x,r)|}\sum_{y\in S(x,r)} w(y) \simeq \frac{1}{k^r}\sum_{y\in S(x,r)} w(y) = \frac{1}{k^r}\sum_{m=0}^{r} c_i k^{r-m} = \frac{1}{k^r}\sum_{m=0}^{r} c_{j+r-2m}k^{r-m}.$$

Given the fact that spheres in this tree grow exponentially, a first natural choice could be studying the behaviour in the case

$$c_j = k^{j\beta}.$$

We shall call

$$w_\beta(x) = \sum_j k^{j\beta}\chi_{T^j}(x).$$

Note that

$$\frac{1}{k^r}\sum_{m=0}^{r} c_{j+r-2m}k^{r-m} = \frac{1}{k^r}\sum_{m=0}^{r} k^{(j+r-2m)\beta}k^{r-m} = k^{j\beta}k^{r\beta}\sum_{m=0}^{r} k^{m(-2\beta-1)}.$$

Hence we would need to show that

$$k^{r\beta}\sum_{m=0}^{r} k^{m(-2\beta-1)} \leq c_\beta$$

uniformly on r. Note that if

$$-2\beta - 1 > 0 \iff \beta < -\frac{1}{2}$$

we have that

$$k^{r\beta}\sum_{m=0}^{r} k^{m(-2\beta-1)} \simeq c_\beta k^{r\beta + r(-2\beta-1)} = c_\beta k^{(-\beta-1)r}.$$

Hence if $\beta \in \left[-1, -\frac{1}{2}\right)$, $k^{r\beta}\sum_{m=0}^{r} k^{m(-2\beta-1)} \leq c_\beta$. If $-\frac{1}{2} \leq \beta < 0$ we have that

$$k^{r\beta}\sum_{m=0}^{r} k^{m(-2\beta-1)} \leq k^{r\beta}\sum_{m=0}^{r} k^{m(-2\beta-1)} \leq k^{r\beta} r \leq 2^{r\beta} r c_\beta.$$

The case $\beta = 0$ is trivial, since it corresponds with having no weight.

In the remainder of the cases, namely if $\beta \in \mathbb{R}\backslash[-1, 0]$, the claimed uniform estimate is not available.

The discusion above can be summarized in the following theorem

Theorem 3.1. *The radial weight*

$$w_\beta(x) = \sum_j k^{j\beta} \chi_{T^j}(x)$$

satisfies $Mw \simeq w$ *iff* $\beta \in [-1, 0]$.

Remark 3.2. Note that if $\beta \in (-1, 0]$ by the argument above there exists $s_\beta > 1$ such that

$$M_{s_\beta} w_\beta(x) \lesssim w_\beta(x)$$

and w_β satisfies the assumptions of Corollary 1.3.

Remark 3.3. The argument used above proves as well that in the case $\beta = -1$ the inequality

$$M_\gamma w_{-1}(x) \lesssim w_{-1}(x)$$

does not hold for any $\gamma > 1$. Indeed, if $\gamma > 1$ we have that

$$\frac{1}{|S(x,r)|} \int_{S(x,r)} w_{-1}(x)^\gamma dx \simeq \frac{1}{k^r} \sum_{m=0}^{r} k^{-(j+r-2m)\gamma} k^{r-m} \simeq k^{-j\gamma} c_\gamma k^{(\gamma-1)r}$$

and on the other hand

$$\left(\frac{1}{|S(x,r)|} \int_{S(x,r)} w_{-1}(x) dx \right)^\gamma \simeq \left(\frac{1}{k^r} \sum_{m=0}^{r} k^{-(j+r-2m)} k^{r-m} \right)^\gamma \simeq k^{-j\gamma}.$$

3.2. **The classical Fefferman-Stein estimate does not hold.** In this section we give our proof of Theorem (1.2). Let $w(x) = \sum_{j=0}^{\infty} \frac{1}{k^j} \chi_{T^j}(x)$. As we showed in the preceding section, for this weight we know that

$$M^n w \simeq_n w.$$

Let

$$f_j(x) = 3\chi_{T^j}(x).$$

First we observe that, since $\int_T f_j(x) w(x) dx = 3$,

$$\int_T f_j(x) M^n w(x) dx \simeq c_n. \tag{3.1}$$

On the other hand, if $x \in T^i$ for $i \leq j$,

$$\begin{aligned}
M^\circ f_j(x) &= \sup_{r \geq 0} \frac{1}{|S(x,r)|} \sum_{y \in B(x,r)} f_j(x) \\
&= 3 \sup_{r \geq 0} \frac{|T^j \cap B(x,r)|}{|S(x,r)|} \\
&\geq 3 \frac{|T^j \cap B(x, j-i)|}{|S(x, j-i)|} \geq 3 \frac{k^{j-i}}{2k^{j-i}} = \frac{3}{2} > 1
\end{aligned}$$

which in turn implies that

$$\bigcup_{i=0}^{j} T^i \subset \{M^\circ f_j(x) > 1\}.$$

Hence, since $w(T^i) = 1$ for every $0 \leq i \leq j$, we have that

$$j \leq \sum_{i=0}^{j} w(T^i) = w\left(\bigcup_{i=0}^{j} T^i\right) \leq w\left(\{M^\circ f_j(x) > 1\}\right). \tag{3.2}$$

The desired conclusion readily follows combining (3.1) and (3.2).

3.3. $Mw \lesssim w$ **is necessary.** We end up this section settling the following result.

Theorem 3.4. *Assume that for a weight w the following estimate holds*

$$w\left(\{x \in T \,:\, M^\circ f(x) > \lambda\}\right) \leq c_w \frac{1}{\lambda} \sum_{x\in T} |f(x)| w(x).$$

Then $Mw \lesssim w$.

Proof. Let us fix $x_0 \in T$ and $r > 0$. We are going to show that

$$\frac{1}{k^r} w(S(x_0, r)) \lesssim w(x_0).$$

We begin noting that

$$S(x_0, r) \subset \left\{x \in T \,:\, M^\circ(\delta_{x_0}) > \frac{1}{2k^r}\right\}$$

Indeed, note that if $x \in S(x_0, r)$ then

$$M^\circ(\delta_{x_0})(x) \simeq \sup_{s\geq 0} \frac{1}{k^s} \sum_{z\in S(x,s)} \delta_{x_0}(z) = \frac{1}{k^r}.$$

Hence

$$w(S(x_0, r)) \leq w\left(\left\{x \in T \,:\, M^\circ(\delta_{x_0}) > \frac{1}{2k^r}\right\}\right) \leq c_w 2k^r \sum_{x\in T} \delta_{x_0}(x) w(x)$$

and consequently

$$\frac{w(S(x_0, r))}{k^r} \leq 2c_w w(x_0)$$

which readily implies that

$$M^\circ w(x) \lesssim w(x) \qquad \text{for all } x \in T.$$ $\square$

4. Vector valued extensions

An interesting application that Fefferman and Stein found in [11] for their two weights estimate were bounds for the following vector valued extensions

$$\left(\sum_{j=1}^{\infty} M(f_j)^q\right)^{\frac{1}{q}} \qquad 1 < q < \infty.$$

Those estimates, besides being an extension of the maximal function, were a generalization of the Marcinkiewicz operator that consists in choosing each f_j to be a suitable characteristic function.

In our case we will also be able to provide some vector valued extensions. First we provide L^p versions of our endpoint Fefferman-Stein estimate that are a direct consequence of the fact that

$$\|Mf\|_{L^\infty(w)} \lesssim \|f\|_{L^\infty(M_s w)}$$

combined with Theorem 1.1 and the Marcinkiewicz interpolation theorem.

Theorem 4.1. *Let $1 < p, s < \infty$ and w a weight. Then*

$$\left(\int_T (Mf)^p\, wdx\right)^{\frac{1}{p}} \leq c_s \left(\int_T |f|^p\, M_s w dx\right)^{\frac{1}{p}} \tag{4.1}$$

where $c_s \to \infty$ when $s \to 1$.

At this point we would like to note that this Theorem can be regarded as a generalization of Nevo and Stein [22] where the case $w = 1$ was essentially settled.

In our next Theorem we provide some vector valued extensions following classical ideas in [11].

Theorem 4.2. *Let $1 < q \leq p < \infty$. Then*

$$\left\|\left(\sum_{j=1}^{\infty} M(f_j)^q\right)^{\frac{1}{q}}\right\|_{L^p(T)} \leq c_{p,q} \left\|\left(\sum_{j=1}^{\infty} |f_j|^q\right)^{\frac{1}{q}}\right\|_{L^p(T)}.$$

Proof. If $p = q$ the proof is straightforward, hence we omit it. For the case $q < p$ we argue by duality.

$$\left\|\left(\sum_{j=1}^{\infty} M(f_j)^q\right)^{\frac{1}{q}}\right\|_{L^p(T)}^q = \sup_{\|g\|_{L^{\left(\frac{p}{q}\right)'}(T)}=1} \left|\int_T \sum_{j=1}^{\infty} M(f_j)^q g dx\right|.$$

Note that using (4.1)

$$\begin{aligned}
\left|\int_T \sum_{j=1}^{\infty} M(f_j)^q g dx\right| &\leq \sum_{j=1}^{\infty} \int_T M(f_j)^q\, |g|\, dx \\
&\leq c_s^q \int_T \sum_{j=1}^{\infty} |f_j|^q M_s(g) dx \\
&\leq c_s^q \left\|\left(\sum_{j=1}^{\infty} |f_j|^q\right)^{\frac{1}{q}}\right\|_{L^p(T)}^q \|M_s g\|_{L^{\left(\frac{p}{q}\right)'}(T)}.
\end{aligned}$$

Now choosing $s < \left(\frac{p}{q}\right)'$ we have that

$$\begin{aligned}&\left\|\left(\sum_{j=1}^{\infty}|f_j|^q\right)^{\frac{1}{q}}\right\|_{L^p(T)}^q \|M_s g\|_{L^{\left(\frac{p}{q}\right)'}(T)}\\ &\leq c_{\left(\frac{p}{q}\right)'}\left\|\left(\sum_{j=1}^{\infty}|f_j|^q\right)^{\frac{1}{q}}\right\|_{L^p(T)}^q \|g\|_{L^{\left(\frac{p}{q}\right)'}(T)}.\end{aligned}$$

Combining the estimates above we are done. □

Appendix A. A weighted version of Soria-Tradacete result for infinite trees

In this appendix we provide a weighted version of [32, Theorem 4.1]. Analogously to the case of the infinite rooted k-ary tree the spherical maximal function on any tree T can be defined as follows:

$$M^{\circ}f(x) = \sup_{r\geq 0}\frac{1}{|S(x,r)|}\sum_{y\in S(x,r)}|f(y)|,$$

where $S(x,r)$ is the sphere

$$S(x,r) = \{y \in T : d(x,y) = r\}.$$

Let $\alpha \in (0,1)$ and $s \in (1,\infty)$. We define

$$\mathcal{E}_T^w(s,r,\alpha) = \sup_{\substack{E,F\subset G\\ |E|,|F|<\infty}}\frac{1}{w(F)^{\alpha}M_s^{\circ}w(E)^{1-\alpha}}\sum_{x\in E}\frac{w\left(F\cap S(x,r)\right)}{|S(x,r)|}$$

and

$$S_T(r) = \sup_{x\in T}|S(x,r)|.$$

With these quantities at our disposal we are ready to settle our weighted version of [32, Theorem 4.1].

Theorem A.1. *For every weight w on a tree T we have that*

$$w\left(\{x\in T\,:\,M^{\circ}f(x)>\lambda\}\right)\lesssim \Gamma_{T,w,r,\alpha,s}\frac{1}{\lambda}\int_T |f(x)|M_s^{\circ}w(x)dx$$

where

$$\Gamma_{T,w,r,\alpha,s} = c_{\alpha}\sup_{n\in\mathbb{N}}\left\{\sum_{S_T(r)\geq 2^{n-1}}^{\infty}\mathcal{E}_T^w(s,r,\alpha)^{\frac{1}{1-\alpha}}S_T(r)^{\frac{1}{2}\frac{\alpha}{1-\alpha}}2^{n\frac{1}{2}\frac{\alpha}{1-\alpha}}\right\}$$

and $c_{\alpha}\to+\infty$ when $\alpha\to 0$.

Again, as we did for the infinite rooted k-ary tree, given any infinite tree T we can define the average operator over the tree T as

$$A_r^\circ f(x) = \frac{1}{|S(x,r)|}\sum_{y\in S(x,r)} |f(y)|.$$

Our next lemma contains the key estimate required to settle Theorem A.1.

Lemma A.2. *Let $r, s > 0$ and $\lambda > 0$. Then*

$$w\left(\{A_r^\circ f \geq \lambda\}\right) \lesssim c_\alpha \sum_{n=0}^{n(r)} 2^{\frac{n}{2}\frac{\alpha}{1-\alpha}} \mathcal{E}_T^w(s,r,\alpha)^{\frac{1}{1-\alpha}} S_T(r)^{\frac{1}{2}\frac{\alpha}{1-\alpha}} 2^n M_s^\circ w\left(\left\{|f| \geq 2^{n-1}\lambda\right\}\right)$$

where $n(r)$ is an integer such that $2^{n(r)} \leq S_T(r) < 2^{n(r)+1}$ and $c_\alpha \to +\infty$ when $\alpha \to 0$.

Proof. We can assume without loss of generality f to be non-negative and $\lambda = 1$. We bound

$$f \leq \frac{1}{2} + \sum_{n=0}^{n(r)} 2^n \chi_{E_n} + f\chi_{\{f\geq \frac{1}{2}S_T(r)\}}, \tag{A.1}$$

where E_n is the sublevel set

$$E_n = \left\{2^{n-1} \leq f < 2^n\right\}. \tag{A.2}$$

Hence

$$A_r^\circ f \leq \frac{1}{2} + \sum_{n=0}^{n(r)} 2^n A_r^\circ\left(\chi_{E_n}\right) + A_r^\circ\left(f\chi_{\{f\geq\frac{1}{2}S_T(r)\}}\right). \tag{A.3}$$

First we note that

$$\begin{aligned} w\left(A_r^\circ\left(f\chi_{\{f\geq\frac{1}{2}S_T(r)\}}\right) \neq 0\right) &\leq w\left(\bigcup_{y\in\{f\geq\frac{1}{2}S_T(r)\}} S(y,r)\right) \\ &\leq \sum_{y\in\{f\geq\frac{1}{2}S_T(r)\}} w(S(y,r)) \leq S_T(r)\sum_{y\in\{f\geq\frac{1}{2}S_T(r)\}}\frac{w(S(y,r))}{|S(y,r)|} \\ &\leq S_T(r)M^\circ w\left(\left\{f \geq \frac{1}{2}S_T(r)\right\}\right) \end{aligned} \tag{A.4}$$

Thus we have that combining the estimates above

$$\begin{aligned} w\left(A_r^\circ f \geq 1\right) &\leq w\left(\sum_{n=0}^{n(r)} 2^n A_r^\circ\left(\chi_{E_n}\right) \geq \frac{1}{2}\right) + w\left(A_r^\circ\left(f\chi_{\{f\geq\frac{1}{2}S_T(r)\}}\right) \neq 0\right) \\ &\leq w\left(\sum_{n=0}^{n(r)} 2^n A_r^\circ\left(\chi_{E_n}\right) \geq \frac{1}{2}\right) + S_T(r)M^\circ w\left(\left\{f \geq \frac{1}{2}S_T(r)\right\}\right) \end{aligned}$$

Let γ be a real parameter such that $0<\gamma<1$ to be chosen later. Note that if

$$\sum_{n=0}^{n(r)} 2^n A_r^\circ\left(\chi_{E_n}\right) \geq \frac{1}{2}$$

then we necessarily have some $n\in\mathbb{N}$, such that $1\leq 2^n\leq 2^{n(r)}$, for which

$$A_r^\circ\left(\chi_{E_n}\right) \geq \frac{2^\gamma-1}{2^{n+2}}\left(\frac{2^n}{S_T(r)}\right)^\gamma .$$

Indeed, otherwise we have that

$$\begin{aligned}
\frac{1}{2} &\leq \sum_{n=0}^{n(r)} 2^n A_r^\circ\left(\chi_{E_n}\right) \leq \frac{(2^\gamma-1)}{4S_T(r)^\gamma}\sum_{n=0}^{n(r)} 2^{\gamma n} \\
&\leq \frac{(2^\gamma-1)}{4S_T(r)^\gamma}\frac{(2^{\gamma(n(r)+1)}-1)}{(2^\gamma-1)} \leq \frac{2^\gamma S_T(r)^\gamma-1}{4S_T(r)^\gamma} < \frac{2^\gamma}{4} < \frac{1}{2}
\end{aligned}$$

which is a contraction. Thus

$$w\left(A_r^\circ f\geq 1\right) \leq \sum_{n=0}^{n(r)} w(F_n) + S_T(r)M^\circ w\left(\left\{f\geq \frac{1}{2}S_T(r)\right\}\right)$$

where

$$F_n = \left\{A_r^\circ\left(\chi_{E_n}\right) \geq \frac{2^\gamma-1}{2^{n+2}}\left(\frac{2^n}{S_T(r)}\right)^\gamma\right\}.$$

Note that F_n is finite and observe that since A_r° is a selfadjoint operator,

$$\begin{aligned}
w(F_n)\frac{2^\gamma-1}{2^{n+2}}\left(\frac{2^n}{S_T(r)}\right)^\gamma &\leq \int_{F_n} wA_r^\circ(\chi_{E_n})(y)dy = \int_{E_n} A_r^\circ(w\chi_{F_n})(y)dy \\
&= \sum_{x\in E_n} A_r^\circ(w\chi_{F_n}) = \sum_{x\in E_n}\frac{1}{|S(x,r)|}\sum_{\substack{y\in F_n\\ d(x,y)=r}} w(y) \\
&= \sum_{x\in E_n}\frac{w(F_n\cap S(x,r))}{|S(x,r)|} \\
&\leq \mathcal{E}_T^w(s,r,\alpha)w(F_n)^\alpha M_s^\circ w(E_n)^{1-\alpha}.
\end{aligned}$$

Now we observe that

$$\begin{aligned}
& w(F_n)\frac{2^\gamma-1}{2^{n+2}}\left(\frac{2^n}{S_T(r)}\right)^\gamma \leq \mathcal{E}_T^w(s,r,\alpha)w(F_n)^\alpha M_s^\circ w(E_n)^{1-\alpha} \\
&\iff w(F_n)^{1-\alpha} \lesssim \frac{1}{2^\gamma-1}2^{n-\gamma n}\mathcal{E}_T^w(s,r,\alpha)S_T(r)^\gamma M_s^\circ w(E_n)^{1-\alpha} \\
&\iff w(F_n) \lesssim \frac{1}{(2^\gamma-1)^{\frac{1}{1-\alpha}}}2^{n\frac{1-\gamma}{1-\alpha}}\mathcal{E}_T^w(s,r,\alpha)^{\frac{1}{1-\alpha}}S_T(r)^{\frac{\gamma}{1-\alpha}}M_s^\circ w(E_n).
\end{aligned}$$

If we choose $\gamma=\frac{\alpha}{2}$ then

$$w(F_n) \lesssim c_\alpha 2^{n\frac{1}{2}\frac{\alpha}{1-\alpha}}\mathcal{E}_T^w(s,r,\alpha)^{\frac{1}{1-\alpha}}S_T(r)^{\frac{1}{2}\frac{\alpha}{1-\alpha}}M_s^\circ w(E_n)2^n$$

with $c_\alpha \to \infty$ when $\alpha \to 0$. This leads to the desired conclusion as in the end of the proof of Lemma 2.3. □

Having the above lemma at our disposal we are in the position to settle Theorem A.1.

Proof. Since $M^\circ f = \sup_{r\geq 0} A_r^\circ f$, Lemma A.2 implies that

$$\begin{aligned}
& w\left(M^\circ f \geq \lambda\right) \\
& \leq \sum_{r=0}^{\infty} w\left(A_r^\circ f \geq \lambda\right) \\
& \lesssim c_\alpha \sum_{r=0}^{\infty} \sum_{n=0}^{n(r)} \mathcal{E}_T^w(s,r,\alpha)^{\frac{1}{1-\alpha}} S_T(r)^{\frac{1}{2}\frac{\alpha}{1-\alpha}} 2^{n\frac{1}{2}\frac{\alpha}{1-\alpha}} 2^n M_s^\circ w\left(\left\{|f| \geq 2^{n-1}\lambda\right\}\right) \\
& \lesssim c_\alpha \sum_{n=0}^{\infty} \sum_{S_T(r)\geq 2^{n-1}}^{\infty} \mathcal{E}_T^w(s,r,\alpha)^{\frac{1}{1-\alpha}} S_T(r)^{\frac{1}{2}\frac{\alpha}{1-\alpha}} c_s 2^{n\frac{1}{2}\frac{\alpha}{1-\alpha}} 2^n M_s^\circ w\left(\left\{|f| \geq 2^{n-1}\lambda\right\}\right) \\
& \lesssim c_\alpha \sup_{n\in\mathbb{N}} \left\{ \sum_{S_T(r)\geq 2^{n-1}}^{\infty} \mathcal{E}_T^w(s,r,\alpha)^{\frac{1}{1-\alpha}} S_T(r)^{\frac{1}{2}\frac{\alpha}{1-\alpha}} c_s 2^{n\frac{1}{2}\frac{\alpha}{1-\alpha}} \right\} \sum_{x\in T} \sum_{n=0}^{\infty} 2^n \chi_{\{|f(x)|\geq 2^{n-1}\lambda\}} M_s^\circ w(x) \\
& \lesssim c_\alpha \sup_{n\in\mathbb{N}} \left\{ \sum_{S_T(r)\geq 2^{n-1}}^{\infty} \mathcal{E}_T^w(s,r,\alpha)^{\frac{1}{1-\alpha}} S_T(r)^{\frac{1}{2}\frac{\alpha}{1-\alpha}} c_s 2^{n\frac{1}{2}\frac{\alpha}{1-\alpha}} \right\} \sum_{x\in T} \frac{1}{\lambda} |f(x)| M_s^\circ w(x).
\end{aligned}$$

This proves Theorem A.1. □

Remark. We would like to note that computing $\mathcal{E}_T^w(s,r,\alpha)$ and $S_T(r)$ and choosing a suitable α can be very difficult. However in certain cases such as for the infinite rooted k-ary tree T it is possible. First we recall that as was noted in [32],

$$S_T(r) \simeq k^r.$$

Now, from Lemma 2.2 it follows that

$$\sum_{x\in E} \frac{w(F\cap S(x,r))}{|S(x,r)|} \lesssim c_s S_T(r)^{\left(\frac{s'}{s'+1}-1\right)} w(F_i)^{\frac{1}{s'+1}} M_s^\circ w(E_j)^{\frac{s'}{s'+1}}$$

and this yields

$$\frac{1}{w(F_i)^{\frac{1}{s'+1}} M_s^\circ w(E_j)^{\frac{s'}{s'+1}}} \sum_{x\in E} \frac{w(F\cap S(x,r))}{|S(x,r)|} \lesssim c_s S_T(r)^{-\frac{1}{s'+1}}.$$

Hence, choosing $\alpha = \frac{1}{s'+1}$ and consequently $1-\alpha = \frac{s'}{s'+1}$ we have that

$$\mathcal{E}_T^w(s,r,\alpha) \leq c_s k^{-\frac{r}{s'+1}}.$$

Then, since $\frac{\alpha}{1-\alpha} = \frac{1}{s'}$,

$$\mathcal{E}_T^w(s,r,\alpha)^{\frac{1}{1-\alpha}} S_T(r)^{\frac{1}{2}\frac{\alpha}{1-\alpha}} 2^{n\frac{1}{2}\frac{\alpha}{1-\alpha}} \lesssim c_s k^{-\frac{r}{s'}} k^{r\frac{1}{2}\frac{1}{s'}} 2^{\frac{n}{2}\frac{1}{s'}} \simeq c_s \left(\frac{2^n}{k^r}\right)^{\frac{1}{2s'}}$$

and consequently, Theorem A.1 recovers the estimate in Theorem 1.1. Note that, by the definition of α, in this case c_α is actually $\tilde{c}_s$ and $\tilde{c}_s \to \infty$ when $s \to 1$ since the latter implies that $\alpha \to 0$.

Remark. We would like to end this Appendix observing that the definition of $\mathcal{E}_T^w(s,r,\alpha)$ could be stated as follows

$$\mathcal{E}_T^w(\tilde{M},r,\alpha) = \sup_{\substack{E,F\subset G\\ |E|,|F|<\infty}} \frac{1}{w(F)^\alpha \tilde{M}w(E)^{1-\alpha}} \sum_{x\in E} \frac{w\left(F\cap S(x,r)\right)}{|S(x,r)|}.$$

where $\tilde{M}$ is some maximal operator such that $M^\circ g \lesssim \tilde{M}g$ for any function $g \in L^1(T)$. Then, exactly the same argument given above allows us to prove the following more general estimate

$$w\left(\{x \in T \,:\, M^\circ f(x) > \lambda\}\right) \lesssim \Gamma_{T,w,r,\alpha,\tilde{M}} \frac{1}{\lambda} \int_T |f(x)|\tilde{M}w(x)dx.$$

where

$$\Gamma_{T,w,r,\alpha,\tilde{M}} = c_\alpha \sup_{n\in\mathbb{N}} \left\{ \sum_{S_T(r)\geq 2^{n-1}}^{\infty} \mathcal{E}_T^w(\tilde{M},r,\alpha)^{\frac{1}{1-\alpha}} S_T(r)^{\frac{1}{2}\frac{\alpha}{1-\alpha}} c_s 2^{n\frac{1}{2}\frac{\alpha}{1-\alpha}} \right\}$$

with $c_\alpha \to \infty$ as $\alpha \to 0$.

Acknowledgement. The first author would like to thank Javier Soria for sharing with him some insights about endpoint estimates for the maximal function on graphs that were one of the motivations to carry out this work.

References

[1] D. Aalto and J. Kinnunen. The discrete maximal operator in metric spaces. *J. Anal. Math.*, 111:369–390, 2010.

[2] H. Aimar, A. Bernardis, and L. Nowak. Dyadic Fefferman-Stein inequalities and the equivalence of Haar bases on weighted Lebesgue spaces. *Proc. Roy. Soc. Edinburgh Sect. A*, 141(1):1–21, 2011.

[3] D. Beltran. A Fefferman-Stein inequality for the Carleson operator. *Rev. Mat. Iberoam.*, 34(1):221–244, 2018.

[4] J. Bourgain. On the maximal ergodic theorem for certain subsets of the integers. *Israel J. Math.*, 61(1):39–72, 1988.

[5] M. Caldarelli, A. K. Lerner, and S. Ombrosi. On a counterexample related to weighted weak type estimates for singular integrals. *Proc. Amer. Math. Soc.*, 145(7):3005–3012, 2017.

[6] E. Carneiro and K. Hughes. On the endpoint regularity of discrete maximal operators. *Math. Res. Lett.*, 19(6):1245–1262, 2012.

[7] M. J. Carro, C. Pérez, F. Soria, and J. Soria. Maximal functions and the control of weighted inequalities for the fractional integral operator. *Indiana Univ. Math. J.*, 54(3):627–644, 2005.

[8] A. Cordoba and C. Fefferman. A weighted norm inequality for singular integrals. *Studia Math.*, 57(1):97–101, 1976.

[9] M. G. Cowling, S. Meda, and A. G. Setti. A weak type $(1,1)$ estimate for a maximal operator on a group of isometries of a homogeneous tree. *Colloq. Math.*, 118(1):223–232, 2010.

[10] C. Domingo-Salazar, M. T. Lacey, and G. Rey. Borderline weak-type estimates for singular integrals and square functions. *Bull. Lond. Math. Soc.*, 48(1):63–73, 2016.

[11] C. Fefferman and E. M. Stein. Some maximal inequalities. *Amer. J. Math.*, 93:107–115, 1971.

[12] G. H. Hardy and J. E. Littlewood. A maximal theorem with function-theoretic applications. *Acta Math.*, 54(1):81–116, 1930.

[13] J. Heinonen. *Lectures on analysis on metric spaces*. Universitext. Springer-Verlag, New York, 2001.

[14] T. Hytönen. A framework for non-homogeneous analysis on metric spaces, and the RBMO space of Tolsa. *Publ. Mat.*, 54(2):485–504, 2010.
[15] A. K. Lerner, S. Ombrosi, and I. P. Rivera-Ríos. On pointwise and weighted estimates for commutators of Calderón-Zygmund operators. *Advances in Mathematics*, 319:153 – 181, 2017.
[16] K. Li, C. Pérez, I. P. Rivera-Ríos, and L. Roncal. Weighted norm inequalities for rough singular integral operators. *J. Geom. Anal.*, 29(3):2526–2564, 2019.
[17] T. Luque and I. Parissis. The endpoint Fefferman-Stein inequality for the strong maximal function. *J. Funct. Anal.*, 266(1):199–212, 2014.
[18] A. Magyar, E. M. Stein, and S. Wainger. Discrete analogues in harmonic analysis: spherical averages. *Ann. of Math. (2)*, 155(1):189–208, 2002.
[19] B. Muckenhoupt. Weighted norm inequalities for the Hardy maximal function. *Trans. Amer. Math. Soc.*, 165:207–226, 1972.
[20] A. Naor and T. Tao. Random martingales and localization of maximal inequalities. *J. Funct. Anal.*, 259(3):731–779, 2010.
[21] F. Nazarov, S. Treil, and A. Volberg. Weak type estimates and Cotlar inequalities for Calderón-Zygmund operators on nonhomogeneous spaces. *Internat. Math. Res. Notices*, (9):463–487, 1998.
[22] A. Nevo and E. M. Stein. A generalization of Birkhoff's pointwise ergodic theorem. *Acta Math.*, 173(1):135–154, 1994.
[23] J. Orobitg and C. Pérez. A_p weights for nondoubling measures in $\mathbf{R}^n$ and applications. *Trans. Amer. Math. Soc.*, 354(5):2013–2033, 2002.
[24] C. Pérez. Weighted norm inequalities for singular integral operators. *J. London Math. Soc. (2)*, 49(2):296–308, 1994.
[25] C. Pérez. Sharp estimates for commutators of singular integrals via iterations of the Hardy-Littlewood maximal function. *J. Fourier Anal. Appl.*, 3(6):743–756, 1997.
[26] C. Pérez and G. Pradolini. Sharp weighted endpoint estimates for commutators of singular integrals. *Michigan Math. J.*, 49(1):23–37, 2001.
[27] M. Rapicki. Fefferman-Stein inequalities for the dyadic-like maximal operators. *Arch. Math. (Basel)*, 113(1):81–93, 2019.
[28] M. C. Reguera and C. Thiele. The Hilbert transform does not map $L^1(Mw)$ to $L^{1,\infty}(w)$. *Math. Res. Lett.*, 19(1):1–7, 2012.
[29] R. Rochberg and M. Taibleson. Factorization of the Green's operator and weak-type estimates for a random walk on a tree. volume 35, pages 187–207. 1991. Conference on Mathematical Analysis (El Escorial, 1989).
[30] Y. Sawano. Sharp estimates of the modified Hardy-Littlewood maximal operator on the nonhomogeneous space via covering lemmas. *Hokkaido Math. J.*, 34(2):435–458, 2005.
[31] J. Soria and P. Tradacete. Best constants for the Hardy-Littlewood maximal operator on finite graphs. *J. Math. Anal. Appl.*, 436(2):661–682, 2016.
[32] J. Soria and P. Tradacete. Geometric properties of infinite graphs and the Hardy-Littlewood maximal operator. *J. Anal. Math.*, 137(2):913–937, 2019.
[33] E. M. Stein and J.-O. Strömberg. Behavior of maximal functions in $\mathbf{R}^n$ for large n. *Ark. Mat.*, 21(2):259–269, 1983.
[34] E. M. Stein and S. Wainger. Discrete analogues in harmonic analysis. II. Fractional integration. *J. Anal. Math.*, 80:335–355, 2000.
[35] K. Stempak. Modified Hardy-Littlewood maximal operators on nondoubling metric measure spaces. *Ann. Acad. Sci. Fenn. Math.*, 40(1):443–448, 2015.
[36] S. A. Tozoni. Weighted norm inequalities for vector-valued singular integrals on homogeneous spaces. *Studia Math.*, 161(1):71–97, 2004.
[37] N. Wiener. The ergodic theorem. *Duke Math. J.*, 5(1):1–18, 1939.

Departamento de Matemática, Universidad Nacional del Sur (UNS), Bahía Blanca, Argentina and INMABB, Universidad Nacional del Sur (UNS)-CONICET, Bahía Blanca, Argentina.

E-mail address: `sombrosi@uns.edu.ar` (Sheldy Ombrosi)
E-mail address: `israel.rivera@uns.edu.ar` (Israel P. Rivera-Ríos)
E-mail address: `msafe@uns.edu.ar` (Martín D. Safe)