

## Research Article

## Special Issue: In honor of David Jerison

Charles Fefferman\* and Bo'az Klartag

# Linear extension operators for Sobolev spaces on radially symmetric binary trees

<https://doi.org/10.1515/ans-2022-0075>

received February 20, 2023; accepted May 17, 2023

**Abstract:** Let  $1 < p < \infty$  and suppose that we are given a function  $f$  defined on the leaves of a weighted tree. We would like to extend  $f$  to a function  $F$  defined on the entire tree, so as to minimize the weighted  $W^{1,p}$ -Sobolev norm of the extension. An easy situation is when  $p = 2$ , where the harmonic extension operator provides such a function  $F$ . In this note, we record our analysis of the particular case of a radially symmetric binary tree, which is a complete, finite, binary tree with weights that depend only on the distance from the root. Neither the averaging operator nor the harmonic extension operator work here in general. Nevertheless, we prove the existence of a linear extension operator whose norm is bounded by a constant depending solely on  $p$ . This operator is a variant of the standard harmonic extension operator, and in fact, it is harmonic extension with respect to a certain Markov kernel determined by  $p$  and by the weights.

**Keywords:** linear extension operator, binary tree

**MSC 2020:** 46E36

## 1 Introduction

Consider a full binary tree of height  $N$ , whose set of vertices is denoted by

$$V = \bigcup_{k=0}^N \{0, 1\}^k,$$

i.e., the vertices are strings of zeroes and ones of length at most  $N$ . For  $x \in \{0, 1\}^k$  and  $\ell \leq k$ , we write  $\pi_\ell(x) \in \{0, 1\}^\ell$  for the prefix of  $x$  of length  $\ell$ . Thus, for  $k \geq 1$ , the parent of a vertex  $x \in \{0, 1\}^k \subseteq V$  is the vertex  $\pi_{k-1}(x)$ . The set  $\{0, 1\}^0$  is a singleton whose unique element is denoted by  $\emptyset$ , the empty string, which is the root of the tree. The set of leaves of the tree is  $\{0, 1\}^N$ , and all other vertices in  $V$  are internal vertices. A vertex  $x \in \{0, 1\}^k$  is said to have depth

$$d(x) = k,$$

thus leaves have depth  $N$  and the root has depth 0. The collections of bijections from  $V$  to  $V$  that preserve depth and parenthood relations form a group. This group is referred to as the symmetry group of the tree. It has  $2^{2^N-1}$  elements, and it is a 2-Sylow subgroup of the group of all permutations of the leaves. The set

---

Dedicated in friendship to David Jerison.

---

\* **Corresponding author: Charles Fefferman**, Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, New Jersey 08544, USA, e-mail: cf@math.princeton.edu

**Bo'az Klartag:** Department of Mathematics, Weizmann Institute of Science, Rehovot 7610001, Israel, e-mail: boaz.klartag@weizmann.ac.il

$$E = \bigcup_{k=1}^N \{0, 1\}^k = V \setminus \{\emptyset\}$$

is referred to as the set of *edges* of the tree, and the depth of an edge  $e \in \{0, 1\}^k$  is  $d(e) = k$ . That is, we think of  $e \in E$  as an undirected edge connecting the vertex whose string corresponds to  $e$  to its unique parent. Each internal vertex other than the root is connected to three vertices, which are its parent and its two children. Assume that we are given edge weights

$$W_1, W_2, \dots, W_N > 0,$$

where we view  $W_k$  as the weight of all edges of depth  $k$ . For  $1 < p < \infty$ , the associated  $\dot{W}^{1,p}$ -seminorm is defined, for  $F : V \rightarrow \mathbb{R}$ , via

$$\|F\|_{\dot{W}^{1,p}(V)} = \left( \sum_{k=1}^N W_k \sum_{x \in \{0,1\}^k} |F(x) - F(\pi_{k-1}(x))|^p \right)^{1/p}. \quad (1)$$

We write  $\partial V = \{0, 1\}^N \subseteq V$  for the set of leaves of the tree. The trace of the  $\|\cdot\|_{\dot{W}^{1,p}}$ -seminorm is defined, for  $f : \partial V \rightarrow \mathbb{R}$ , via

$$\|f\|_{\dot{W}^{1,p}(\partial V)} = \inf\{\|F\|_{\dot{W}^{1,p}(V)}; F|_{\partial V} = f\}, \quad (2)$$

i.e., the infimum of the  $\dot{W}^{1,p}$ -seminorm over all extensions of  $f$  from the leaves to the entire tree. We write  $\mathbb{R}^V$  for the collection of all functions  $f : V \rightarrow \mathbb{R}$ , and similarly  $\mathbb{R}^{\partial V}$  is the collection of all functions  $f : \partial V \rightarrow \mathbb{R}$ . Our main result is the following:

**Theorem 1.1.** *Let  $1 < p < \infty$  and let  $W_1, \dots, W_N > 0$ . Then there exists a linear operator  $H : \mathbb{R}^{\partial V} \rightarrow \mathbb{R}^V$  with the following properties:*

- (1) *It is a linear extension operator, i.e.,  $(Hf)(x) = f(x)$  for any  $x \in \partial V$  and any function  $f : \partial V \rightarrow \mathbb{R}$ .*
- (2) *Its norm is bounded by a constant  $\bar{C}_p$  depending only on  $p$ , i.e., for any  $f : \partial V \rightarrow \mathbb{R}$ ,*

$$\|Hf\|_{\dot{W}^{1,p}(V)} \leq \bar{C}_p \|f\|_{\dot{W}^{1,p}(\partial V)}.$$

*In fact, we have the bound*

$$\bar{C}_p \leq 4p^{1/p} q^{1/q} \cdot (1 + \max\{(p-1)^{-1/p}, (q-1)^{-1/q}\}) \leq C \cdot \max\left\{\frac{1}{p-1}, \frac{1}{q-1}\right\}, \quad (3)$$

*where  $q = p/(p-1)$  and where  $C > 0$  is a universal constant.*

The proof of Theorem 1.2 is constructive, and the extension operator  $H$  that we construct is in fact a *harmonic extension operator* with respect to a certain random walk defined on the tree. At each step, the random walk jumps from a vertex to one of its neighbors, where of course the neighbors of a vertex are its parent and its children. The Markov kernel corresponding to the random walk is invariant under the symmetries of the tree; thus, the probability to move from a vertex to its neighbor depends only on the weights of the vertex and of its neighbor. The Markov kernel of our random walk is determined by the following requirement: For any  $s = 1, \dots, N$ , the probability that a random walk starting at some vertex of depth  $s$  will reach a leaf before reaching a vertex of depth  $s-1$  equals

$$q_s = \frac{(2^s W_s)^{-1/(p-1)}}{\sum_{k=s}^N (2^k W_k)^{-1/(p-1)}}. \quad (4)$$

Thus, the weights of our random walk typically depend on  $p \in (1, \infty)$ , except for the case where  $W_s$  is proportional to  $2^{-s}$ . This seems inevitable. Indeed, in some examples such as the 3rd example in Section 2, the linear extension operator  $H : \mathbb{R}^{\partial V} \rightarrow \mathbb{R}^V$  that corresponds to the parameter value  $p = p_0$  is not uniformly bounded for any  $p \in (1, \infty) \setminus \{p_0\}$ . When  $p = 2$ , our random walk coincides with the usual random walk corresponding to the given weights on the edges of the binary tree, and thus in this case,  $H$  is the standard harmonic extension operator (hence,  $\bar{C}_p = 1$  for  $p = 2$ ).

There is a certain range of weights where the *averaging operator* yields a uniformly bounded linear extension operator, as proven by Björn et al. [1]. The averaging operator is the extension operator that assigns to each internal vertex  $v$  the average of the function values on the leaves of the subtree whose root is  $v$ . This averaging operator seems natural also from the point of view of Whitney's extension theory, and see the work by Shvartsman [5] on Sobolev extension in  $W^{1,p}(\mathbb{R}^n)$ . However, there are examples of radially-symmetric binary trees where the averaging operator does not provide a uniformly bounded operator, such as the case where  $W_k = 2^{-k}$  for all  $k$ .

What about trees with weights that are not radially symmetric? Suppose that the edge weights are arbitrary positive numbers  $(W_e)_{e \in E}$  that are not necessarily determined by the depth of the edge. When  $p = 2$ , there is still a harmonic extension operator of norm one from  $\dot{W}^{1,p}(\partial V)$  to  $\dot{W}^{1,p}(V)$ . However, for  $p \neq 2$ , the situation seems subtle. We conjecture that in the general case of nonradially-symmetric tree weights, there is no linear extension operator whose norm is bounded by a function of  $p$  alone. This conjecture is closely related to questions about well-complemented subspaces of  $\ell_p^n$  that are beyond the scope of this note.

To prove Theorem 1.1, we reformulate the problem in a way that brings us closer to analysis in  $\ell_p$ -spaces. For  $1 < p < \infty$ , the associated  $L^p(E)$ -norm is defined, for  $f: E \rightarrow \mathbb{R}$ , via

$$\|f\|_p = \|f\|_{L^p(E)} = \left( \sum_{k=1}^N W_k \sum_{x \in \{0,1\}^k} |f(x)|^p \right)^{1/p}. \quad (5)$$

We think of a function  $f: E \rightarrow \mathbb{R}$  as the gradient of a function  $\tilde{f}: V \rightarrow \mathbb{R}$ , uniquely determined up to an additive constant. Given  $f: E \rightarrow \mathbb{R}$ , we thus define a function  $\tilde{f}: V \rightarrow \mathbb{R}$  as follows: For any  $x \in V$  other than the root,

$$\tilde{f}(x) = \sum_{i=1}^{d(x)} f(\pi_i x), \quad (6)$$

while  $\tilde{f}(x) = 0$  if  $x = \emptyset$  is the root. The only property of  $\tilde{f}$  that matters is that for all  $x \in E$ ,

$$f(x) = \tilde{f}(x) - \tilde{f}(\pi_{d(x)-1}(x)).$$

We are interested in finding a linear operator  $T: L^p(E) \rightarrow L^p(E)$ , with a uniform bound on its operator norm, which has the following properties:

(1) The operator  $T$  takes the form

$$Tf(x) = \tilde{T}\tilde{f}(x) - \tilde{T}\tilde{f}(\pi_{d(x)-1}(x)) \quad (7)$$

for some linear operator  $\tilde{T}: \mathbb{R}^V \rightarrow \mathbb{R}^V$ . That is,  $\tilde{T}$  takes functions on  $V$  to functions on  $V$ , and  $T$  is induced from  $\tilde{T}$  via formula (7).

- (2) The operator  $\tilde{T}$  is equivariant with respect to the tree symmetries and it satisfies  $\tilde{T}(1) \equiv 1$ , i.e., it maps the constant function 1 to itself.
- (3) The function  $\tilde{T}g$  coincides with the function  $g$  on the leaves of the tree, i.e.,  $(\tilde{T}g)|_{\partial V} = g|_{\partial V}$  for any  $g: V \rightarrow \mathbb{R}$ .
- (4) The function  $\tilde{T}g$  is determined by the values of the function  $g$  on the leaves of the tree.

The operator norm of  $\tilde{T}$  with respect to the  $\dot{W}^{1,p}$ -seminorm equals to the operator norm of  $T$  with respect to the  $L^p(E)$ -norm. Defining  $H(f|_{\partial V}) = \tilde{T}f$ , Theorem 1.1 may thus be reformulated as follows:

**Theorem 1.2.** *Let  $1 < p < \infty$  and let  $W_1, \dots, W_N > 0$ . Then there exists a linear operator  $T: L^p(E) \rightarrow L^p(E)$  with the aforementioned properties, whose operator norm is at most a certain constant  $\bar{C}_p$  depending only on  $p$ . In fact, we have the bound (3) for the constant  $\bar{C}_p$ .*

The proof of Theorem 1.2 occupies the next three sections. In Section 2, we discuss invariant random walks on a full binary tree and describe the corresponding harmonic extension operator. In Section 3, we deal with

the problem of bounding the norm of this operator and use the symmetries of the problem to reduce it to a one-dimensional question. This one-dimensional question is then answered in Section 4 using the Muckenhoupt criterion [4].

When analyzing the binary tree, we use the following notation: We write  $\text{dlca}(x, y)$  for the length of the maximal prefix shared by the strings  $x \in \{0, 1\}^k$  and  $y \in \{0, 1\}^\ell$ , while  $\text{lca}(x, y)$  is the maximal prefix itself. Thus for two vertices  $x, y \in V$ , their least common ancestor is  $\text{lca}(x, y) \in V$  and its depth is  $\text{dlca}(x, y) \in \{0, \dots, N\}$ . Note that for any  $x \in E$  and  $\omega \in \partial V$ ,

$$\text{dlca}(\pi_{d(x)-1}x, \omega) = \min\{d(x) - 1, \text{dlca}(x, \omega)\}. \quad (8)$$

## 2 Invariant random walks

A Markov chain on  $V$  is a sequence of random variables  $R_1, R_2, \dots \in V$  such that the distribution of  $R_{i+1}$  conditioned on  $R_1, \dots, R_i$  is the same as its distribution conditioned on  $R_i$ . A Markov chain is time-homogeneous if for any  $x, y \in V$ , the probability that  $R_{i+1} = x$  conditioned on the event  $R_i = y$  does not depend on  $i$ . A random walk on  $V$  is a time-homogeneous Markov chain  $R_1, R_2, \dots \in V$  such that  $R_{i+1}$  is a neighbor of  $R_i$  with probability one.

We say that the random walk is *invariant* if the probability to jump from a vertex  $x$  to a vertex  $y$  depends only on the depths  $d(x)$  and  $d(y)$ . Our random walk will be invariant, and it will stop when it reaches a leaf, i.e., we have the stopping time

$$\tau = \min\{i \geq 1; R_i \text{ is a leaf}\}.$$

For  $s \geq 1$ , we define  $q_s$  to be the probability of the following event: Assuming that  $R_1$  is a vertex of depth  $s$ , the event is that  $R_i$  will remain at the subtree whose root is  $R_1$  for all  $1 \leq i \leq \tau$ . Equivalently, define

$$X_i = d(R_i) \in \{0, \dots, N\}.$$

Then  $X_1, X_2, \dots \in \{0, \dots, N\}$  is a random walk, since  $|X_{i+1} - X_i| = 1$  for all  $i$ . Furthermore,

$$q_s = \mathbb{P}(\forall 1 \leq i \leq \tau, X_i \geq s \mid X_1 = s). \quad (9)$$

Clearly,

$$q_0 = q_N = 1. \quad (10)$$

For  $r, s \in \{0, \dots, N\}$  with  $r \leq s$ , we set

$$p_{s,r} = \mathbb{P}(\min\{X_i; i \leq \tau\} = r \mid X_1 = s).$$

That is, the number  $p_{s,r}$  is the probability that  $r$  is the minimal node that the walker visits when starting from node  $s$ , before reaching the terminal node  $N$ . Clearly,  $\sum_{r=0}^s p_{s,r} = 1$ .

**Lemma 2.1.** For  $0 \leq r \leq s \leq N - 1$ ,

$$p_{s,r} = q_r \cdot \prod_{k=r+1}^s (1 - q_k), \quad (11)$$

where an empty product equals 1. Moreover,  $p_{N,r} = \delta_{rN}$ , where  $\delta_{rN}$  is the Kronecker delta.

**Proof.** The expression on the right-hand side of (11) is the probability to ever reach  $s - 1$  when starting from  $X_1 = s$ , and from  $s - 1$  to ever reach  $s - 2$ , etc. until we finally reach  $r$ , yet from  $r$  we require to *never* reach  $r - 1$ . Alternatively, when  $0 \leq r \leq s, s \geq 1$ , we have the recurrence relation:

$$p_{s,r} = q_s \delta_{s,r} + (1 - q_s) \cdot p_{s-1,r}. \quad (12)$$

This recurrence relation leads to another proof of (11).  $\square$

Suppose that our random walk  $R_1, R_2, \dots \in V$  begins at a vertex  $R_1 = x$  with  $d(x) = s$ . Consider a leaf  $y \in \partial V$  with  $\text{dlca}(x, y) = r$ . What is the probability that our random walk will reach the leaf  $y$ ? We claim that this probability is

$$b_{s,r} := \mathbb{P}(R_\tau = y) = \sum_{k=0}^r 2^{k-N} p_{s,k}. \quad (13)$$

Indeed, conditioning on the value of  $k = \min_{i \leq \tau} d(R_i)$ , by symmetry, we know that  $R_\tau$  is distributed uniformly among the  $2^{N-k}$  leaf descendants of the vertex  $\pi_k(x)$ . When  $k \leq r$ , exactly one of these leaf-descendants is the leaf  $y$ , since the vertex of minimal depth that  $(R_i)$  visits must be the vertex  $\pi_k(x)$ , which is a prefix of  $y$  as  $k \leq r = \text{dlca}(x, y)$ . Hence, the probability that  $R_\tau = y$ , conditioning on the value of  $k$ , equals to  $1/2^{N-k}$  when  $k \leq r$ , and it vanishes otherwise. By using the definition of  $p_{s,k}$  and the complete probability formula, we obtain (13). The *harmonic extension operator* associated with our invariant random walk is given by

$$\tilde{T}g(x) = \sum_{\omega \in \{0,1\}^N} b_{d(x), \text{dlca}(x, \omega)} \cdot g(\omega) \quad (x \in V). \quad (14)$$

The operator  $T$  is induced from  $\tilde{T}$  via formula (7). Requirements 1, ..., 4 from Section 1 are clearly satisfied.

We stipulate that the collection of descendants of a vertex  $x \in V$ , denoted by  $D(x) \subseteq V$ , includes the vertex  $x$  itself. Abbreviate  $a \wedge b = \min\{a, b\}$  and  $a \vee b = \max\{a, b\}$ . The operator  $T$  takes the form

$$Tf(x) = \sum_{y \in E} K(x, y) f(y), \quad (15)$$

where the kernel  $K$  is described next.

**Proposition 2.2.** *Let  $x, y \in E$  and denote  $s = d(x)$ ,  $t = d(y)$ ,  $r = \text{dlca}(x, y)$ . Then the followings hold: If  $x \notin D(y)$  and  $y \notin D(x)$ , then  $r \leq s \wedge t - 1$  and*

$$K(x, y) = -q_s \cdot 2^{-t} \cdot \sum_{k=0}^r 2^k p_{s-1,k} \leq 0. \quad (16)$$

*Otherwise, i.e., if  $y \in D(x)$  or if  $x \in D(y)$ , then  $r = s \wedge t$  and*

$$K(x, y) = q_s \cdot 2^{-t} \cdot \sum_{k=0}^{r-1} (2^r - 2^k) p_{s-1,k} \geq 0. \quad (17)$$

**Proof.** By (7) and (14), we have, for any  $x \in E$ ,

$$\begin{aligned} Tf(x) &= \tilde{T}\tilde{f}(x) - \tilde{T}\tilde{f}(\pi_{d(x)-1}x) \\ &= \sum_{\omega \in \{0,1\}^N} b_{d(x), \text{dlca}(x, \omega)} \tilde{f}(\omega) - \sum_{\omega \in \{0,1\}^N} b_{d(x)-1, \text{dlca}(\pi_{d(x)-1}x, \omega)} \tilde{f}(\omega) \\ &= \sum_{\omega \in \{0,1\}^N} a_{d(x), \text{dlca}(x, \omega)} \tilde{f}(\omega), \end{aligned} \quad (18)$$

where for any  $0 \leq r \leq s$ , by (8) and (13),

$$a_{s,r} = b_{s,r} - b_{s-1, \min\{s-1, r\}} = \sum_{k=0}^r 2^{k-N} p_{s,k} - \sum_{k=0}^{\min\{s-1, r\}} 2^{k-N} p_{s-1,k}. \quad (19)$$

Hence, by (6) and (18),

$$\begin{aligned} Tf(x) &= \sum_{\omega \in \{0,1\}^N} a_{d(x), \text{dlca}(x, \omega)} \sum_{i=1}^N f(\pi_i \omega) \\ &= \sum_{y \in E} \left[ \sum_{\omega \in \{0,1\}^N; \pi_{d(y)} \omega = y} a_{d(x), \text{dlca}(x, \omega)} \right] f(y) = \sum_{y \in E} K(x, y) f(y), \end{aligned}$$

where the kernel  $K$  of the operator  $T$  satisfies, for any  $x, y \in E$ ,

$$K(x, y) = \sum_{\omega \in \{0,1\}^N; \pi_{d(y)}\omega = y} a_{d(x), \text{dlca}(x, \omega)}. \quad (20)$$

Fix  $x, y \in E$  with  $s = d(x)$ ,  $t = d(y)$  and  $r = \text{dlca}(x, y)$ . Let us consider first the case where  $x$  is *not* a descendant of  $y$ . This means that the prefix of  $y$  that is shared by  $x$ , is not the entire string  $y$ . Hence, for any  $\omega \in \{0, 1\}^N$  with  $\pi_{d(y)}\omega = y$ , we have  $\text{dlca}(x, \omega) = \text{dlca}(x, y) \leq d(y) - 1$ . Therefore, from (20) and (19),

$$K(x, y) = 2^{N-d(y)} a_{d(x), \text{dlca}(x, y)} = 2^{N-t} [b_{s,r} - b_{s-1, (s-1) \wedge r}] = \sum_{k=0}^r 2^{k-t} p_{s,k} - \sum_{k=0}^{(s-1) \wedge r} 2^{k-t} p_{s-1,k},$$

as  $b_{s,r} = \sum_{k=0}^r 2^{k-N} p_{s,k}$ . Since  $p_{s,s} = q_s$ , we have

$$K(x, y) = \delta_{rs} 2^{s-t} q_s + \sum_{k=0}^{(s-1) \wedge r} 2^{k-t} [p_{s,k} - p_{s-1,k}] = q_s \cdot \left[ \delta_{rs} 2^{s-t} - \sum_{k=0}^{(s-1) \wedge r} 2^{k-t} p_{s-1,k} \right], \quad (21)$$

where we used relation (12), which implies that when  $k \leq s-1$ ,

$$p_{s,k} - p_{s-1,k} = -q_s \cdot p_{s-1,k}. \quad (22)$$

We may now prove the conclusion of the proposition in the case where  $x \notin D(y)$  and  $y \notin D(x)$ . Indeed, in this case,  $r \leq s \wedge t - 1$  and formula (21) applies. Since  $\delta_{rs} = 0$  in this case, we deduce formula (16) from (21).

The next case we consider is the case where  $r = s \leq t - 1$ , or equivalently, where  $y \in D(x) \setminus \{x\}$ . Thus,  $x \notin D(y)$  and formula (21) applies. Recalling that  $\sum_{k=0}^{s-1} p_{s-1,k} = 1$ , we obtain from (21) that

$$K(x, y) = q_s \cdot \sum_{k=0}^{s-1} (2^{s-t} - 2^{k-t}) p_{s-1,k} = q_s \cdot 2^{-t} \cdot \sum_{k=0}^{s-1} (2^r - 2^k) p_{s-1,k},$$

proving formula (17) in the case  $y \in D(x) \setminus \{x\}$ .

We move on to the case where  $x \in D(y) \setminus \{y\}$ , thus  $\text{dlca}(x, y) = t \leq s - 1$ . In this case, by applying (20), (19), (13) and then (22),

$$\begin{aligned} K(x, y) &= \sum_{\omega \in \{0,1\}^N; \pi_{d(y)}\omega = y} a_{d(x), \text{dlca}(x, \omega)} = 2^{N-d(x)} a_{d(x), d(x)} + \sum_{k=d(y)}^{d(x)-1} 2^{N-k-1} a_{d(x), k} \\ &= 2^{N-s} a_{s,s} + \sum_{k=t}^{s-1} 2^{N-k-1} a_{s,k} = 2^{N-s} (b_{s,s} - b_{s-1, s-1}) + \sum_{k=t}^{s-1} 2^{N-k-1} (b_{s,k} - b_{s-1,k}) \\ &= 2^{N-s} \left[ \sum_{k=0}^s 2^{k-N} p_{s,k} - \sum_{k=0}^{s-1} 2^{k-N} p_{s-1,k} \right] + \sum_{k=t}^{s-1} 2^{N-k-1} \sum_{\ell=0}^k 2^{\ell-N} (p_{s,\ell} - p_{s-1,\ell}) \\ &= q_s - q_s \sum_{k=0}^{s-1} 2^{k-s} p_{s-1,k} - q_s \sum_{\ell=0}^{s-1} \sum_{k=\ell \vee t}^{s-1} 2^{\ell-k-1} p_{s-1,\ell} \\ &= q_s \left( 1 - \sum_{k=0}^{s-1} 2^{k-s} p_{s-1,k} - \sum_{\ell=0}^{s-1} [2^{\ell-\ell \vee t} - 2^{\ell-s}] p_{s-1,\ell} \right) \\ &= q_s \left( 1 - \sum_{k=0}^{s-1} 2^{k-k \vee t} p_{s-1,k} \right) = q_s \cdot \left( 1 - \sum_{k=0}^{t-1} 2^{k-t} p_{s-1,k} - \sum_{k=t}^{s-1} p_{s-1,k} \right) \\ &= q_s \cdot \sum_{k=0}^{t-1} (1 - 2^{k-t}) p_{s-1,k} = q_s \cdot 2^{-t} \cdot \sum_{k=0}^{t-1} (2^t - 2^k) p_{s-1,k}. \end{aligned}$$

Since  $r = t$  in this case, we have proved formula (17) in the case where  $y \in D(x) \setminus \{x\}$ . Finally, the last case that remains is when  $x = y$ . In this case,  $r = s = t$  and

$$\begin{aligned}
K(x, y) &= \sum_{\omega \in \{0,1\}^N; \pi_{d(y)} \omega = y} a_{d(x), \text{dlca}(x, \omega)} = 2^{N-d(x)} a_{d(x), d(x)} = 2^{N-s} a_{s,s} = 2^{N-s} (b_{s,s} - b_{s-1,s-1}) \\
&= 2^{N-s} \left[ \sum_{k=0}^s 2^{k-N} p_{s,k} - \sum_{k=0}^{s-1} 2^{k-N} p_{s-1,k} \right] = q_s - q_s \sum_{k=0}^{s-1} 2^{k-s} p_{s-1,k} \\
&= q_s \cdot \sum_{k=0}^{s-1} (1 - 2^{k-s}) p_{s-1,k},
\end{aligned}$$

completing the proof of formula (17).  $\square$

*Some examples.*

- (1) The simplest example is when  $q_s = 1$  for all  $s \geq 1$ . In this case, the operator  $\tilde{T}$  is the familiar averaging operator. That is, the extension operator  $\tilde{T}$  is the operator that assigns to each vertex the average of the values at the leaves of its subtree. In this case,

$$p_{s,r} = \delta_{s,r}.$$

- (2) Consider the case where the invariant random walk is such that  $X_i := d(R_i)$  is a symmetric random walk on  $\{0, \dots, N\}$ , i.e., the probability to jump from  $i$  to  $i+1$  is exactly  $1/2$  for  $i = 1, \dots, N-1$ . Recall that  $q_s$  is the probability to never leave the subtree when starting at a vertex of depth  $s$ . We claim that in this example, for  $s = 1, \dots, N$ ,

$$q_s = \frac{1}{N-s+1}. \quad (23)$$

Indeed, the function  $f(i) = i$  is harmonic on  $\{0, 1, \dots, N\}$ , and hence,  $f(X_i)$  is a martingale. Thus, for any stopping time  $\tilde{\tau}$ , we have  $f(X_1) = \mathbb{E}f(X_{\tilde{\tau}})$ . We pick the stopping time

$$\tilde{\tau} = \min\{i; X_i \in \{s-1, N\}\}$$

and obtain (23) since

$$N \cdot q_s + (s-1) \cdot (1 - q_s) = s.$$

Next, we use Lemma 2.1 and find a formula for  $p_{s,r}$ . Since formula (23) is valid for any  $s \geq 1$ , we conclude that for any  $r \geq 0$  and  $s \geq r+1$ ,

$$\prod_{k=r+1}^s (1 - q_k) = \prod_{k=r+1}^s \frac{N-k}{N-k+1} = \frac{N-s}{N-r}. \quad (24)$$

Formula (24) is actually valid for any  $0 \leq r \leq s$ , since an empty product equals one. Recall that  $q_0 = 1$ . We thus conclude from Lemma 2.1 that for  $s = 0, \dots, N-1$ ,

$$p_{s,r} = \frac{N-s}{N-r} \cdot \begin{cases} 1 & r = 0 \\ \frac{1}{N-r+1} & 1 \leq r \leq s, \end{cases}$$

while  $p_{N,r} = \delta_{N,r}$ .

- (3) Let  $0 < \delta < 1$ , and consider the case where  $(X_i)$  is a random walk on  $\{0, \dots, N\}$  such that the probability to jump from  $k$  to  $k+1$  equals  $1/2$  if  $k < N-1$ , and it equals  $\delta$  if  $k = N-1$ . A harmonic function here is

$$f(k) = \begin{cases} k & k \leq N-1 \\ N-2+1/\delta & k = N. \end{cases}$$

Therefore, for  $s = 1, \dots, N-1$ ,

$$(N-2+1/\delta) \cdot q_s + (s-1)(1 - q_s) = s.$$

Thus,  $q_0 = q_N = 1$ , while for  $s = 1, \dots, N-1$ ,

$$q_s = \frac{1}{N-s+1/\delta-1}.$$

Hence, for any  $s \leq N - 1$  and  $r \leq s - 1$ ,

$$\prod_{k=r+1}^s (1 - q_k) = \prod_{k=r+1}^s \frac{N - k + \delta^{-1} - 2}{N - k + \delta^{-1} - 1} = \frac{N - s + \delta^{-1} - 2}{N - r + \delta^{-1} - 2}.$$

We conclude from Lemma 2.1 that for  $s = 0, \dots, N - 1$  and  $0 \leq r \leq s$ ,

$$p_{s,r} = q_r \cdot \prod_{k=r+1}^s (1 - q_k) = \frac{N - s + \delta^{-1} - 2}{N - r + \delta^{-1} - 2} \cdot \begin{cases} 1 & r = 0 \\ \frac{1}{N - r + 1/\delta - 1} & 1 \leq r \leq s, \end{cases}$$

while  $p_{N,r} = \delta_{N,r}$ .

We conclude this section with the following:

**Lemma 2.3.** *For any numbers  $q_1, \dots, q_{N-1} \in (0, 1)$ , there exists a random walk*

$$X_1, X_2, \dots \in \{0, \dots, N\}$$

*satisfying (9) with  $\tau = \min\{i \geq 1; X_i = N\}$  for  $s = 1, \dots, N - 1$ .*

**Proof.** Write  $x_k$  for the probability that the random walk jumps from  $k$  to  $k + 1$ . Then for  $s = 0, \dots, N - 1$ ,

$$q_s = x_s(q_{s+1} + (1 - q_{s+1})x_s), \quad (25)$$

where we set  $q_0 = q_N = 1$ . We claim that  $x_s \in [0, 1]$  is determined by equation (25). Indeed, the right-hand side of (25) is a continuous, increasing function of  $x_s$  in the interval  $[0, 1]$ , which maps this interval to itself.  $\square$

### 3 The ancestral and nonancestral parts of the kernel

We need to bound the operator norm in  $L^p(E)$  of the operator  $T$  whose kernel is described in Proposition 2.2. Let us consider first the *nonancestral* part of the operator, given by the kernel

$$K_0(x, y) = K(x, y) \cdot 1_{\{x \notin D(y), y \notin D(x)\}} = -1_{\{r \leq s \wedge t-1\}} \cdot q_s \cdot 2^{-t} \cdot \sum_{k=0}^r 2^k p_{s-1,k}. \quad (26)$$

Here, as usual  $s = d(x)$ ,  $t = d(y)$ , and  $r = \text{dlca}(x, y)$ . Write  $T_0 : L^p(E) \rightarrow L^p(E)$  for the operator whose kernel is  $K_0$ . A function  $f : E \rightarrow \mathbb{R}$  is invariant under the symmetries of the tree, or *invariant* in short, if it takes the form

$$f(x) = F(d(x))$$

for some function  $F : \{1, \dots, N\} \rightarrow \mathbb{R}$ . The operator  $T_0$  is equivariant under the symmetries of the tree. Therefore, if  $f(x) = F(d(x))$  is an invariant function, then so is  $T_0 f$ . In fact, in the case where  $f(x) = F(d(x))$ , we can write

$$T_0 f(x) = \sum_{t=1}^N L_0(d(x), t) F(t) \quad (27)$$

for a certain kernel  $L_0(s, t)$  defined for  $s, t = 1, \dots, N$ .

**Lemma 3.1.** *For  $s, t = 1, \dots, N$ ,*

$$L_0(s, t) = -q_s \cdot \sum_{k=0}^{m-1} (1 - 2^{k-m}) p_{s-1,k} \leq 0.$$

**Proof.** Let  $r \leq \min\{t, s\} - 1$ . A moment of reflection reveals that for  $x \in E$  with  $d(x) = s$ ,



$$n(t; s, r) = \#\{y \in E; d(y) = t, \text{dlca}(x, y) = r\} = 2^{t-r-1}.$$

By (26), (27), and the definition of  $T_0$ , for any  $x \in E$  with  $d(x) = s$ ,

$$L_0(s, t) = \sum_{y \in E; d(y)=t} K_0(x, y) = - \sum_{r=0}^{t \wedge s-1} n(t; s, r) \cdot q_s \cdot 2^{-t} \cdot \sum_{k=0}^r 2^k p_{s-1, k}.$$

Denote  $m = s \wedge t$ . Then for  $s, t = 1, \dots, N$ ,

$$L_0(s, t) = -q_s \cdot \sum_{r=0}^{t \wedge s-1} \sum_{k=0}^r 2^{k-r-1} p_{s-1, k} = -q_s \cdot \sum_{k=0}^{m-1} \sum_{r=k}^{m-1} 2^{k-r-1} p_{s-1, k} = -q_s \cdot \sum_{k=0}^{m-1} (1 - 2^{k-m}) p_{s-1, k}. \quad \square$$

Write  $\Omega_N = \{1, \dots, N\}$  and for  $F : \Omega_N \rightarrow \mathbb{R}$  define

$$\|F\|_p = \|F\|_{L^p(\Omega_N)} = \left( \sum_{k=1}^N 2^k \cdot W_k \cdot |F(k)|^p \right)^{1/p}. \quad (28)$$

Observe that  $\|f\|_{L^p(E)} = \|F\|_p$  if  $f(x) = F(d(x))$ . Let

$$S_0 F(s) = \sum_{t=1}^N L_0(s, t) F(t)$$

so that by (27),

$$T_0(F \circ d) = (S_0 F) \circ d. \quad (29)$$

For  $f, g : E \rightarrow \mathbb{R}$ , we consider the scalar product

$$\langle f, g \rangle = \sum_{x \in E} W_{d(x)} f(x) g(x),$$

while for  $F, G : \Omega_N \rightarrow \mathbb{R}$ , we set

$$\langle F, G \rangle = \langle F \circ d, G \circ d \rangle = \sum_{k=1}^N 2^k W_k F(k) G(k).$$

The adjoint operators  $T_0^*$  and  $S_0^*$  are defined with respect to these scalar products. The following lemma is probably well known to experts (see, e.g., Howard and Schep [2] for a related argument), and its proof is provided for completeness.

**Lemma 3.2.** *Let  $1 < p < \infty$ . Then the norm of the operator  $T_0 : L^p(E) \rightarrow L^p(E)$  is attained at an invariant, nonnegative function  $f$ , and it equals to the norm of the operator  $S_0 : L^p(\Omega_n) \rightarrow L^p(\Omega_n)$ .*

**Proof.** Denote momentarily  $\mathcal{T} = -T_0$  and  $S = -S_0$ . By Lemma 3.1, the kernel  $-L_0$  of the operator  $S$  is nonnegative, and by (26), the kernel of the operator  $\mathcal{T}$  is nonnegative as well. By approximation, we may assume that these two kernels are strictly positive, while keeping condition (29), thus

$$\mathcal{T}(F \circ d) = (SF) \circ d. \quad (30)$$

We deduce that  $\mathcal{T}^*(F \circ d) = (S^*F) \circ d$ . By compactness,

$$\sup_{0 \neq F \in L^p(\Omega_N)} \frac{\|SF\|_p}{\|F\|_p}$$

is attained at some function  $F$ . Since the kernel of  $S$  is nonnegative, we may assume that the extremal function  $F$  is nonnegative. By the Lagrange multipliers theorem, the function  $F$  satisfies a certain eigenvalue equation, and in fact, there exists  $\lambda \in \mathbb{R}$  such that

$$S^*(SF)^{p-1} = \lambda F^{p-1}. \quad (31)$$

Since the kernel of  $S$  is positive and  $F$  is nonnegative and not identically zero, it follows from (31) that  $F$  is actually positive. The norm of the operator  $S : L^p(\Omega_n) \rightarrow L^p(\Omega_n)$  equals  $\lambda^{1/p} > 0$ , since

$$\|SF\|_p^p = \langle (SF)^{p-1}, SF \rangle = \langle S^*(SF)^{p-1}, F \rangle = \lambda \langle F^{p-1}, F \rangle = \lambda \|F\|_p^p.$$

Denoting  $f = F \circ d$ , we find from (30) that  $f$  is a positive invariant function satisfying

$$\mathcal{T}^*(\mathcal{T}f)^{p-1} = \lambda f^{p-1}.$$

Since the kernel of  $\mathcal{T}$  is nonnegative, we have the pointwise Hölder inequality

$$\mathcal{T}(uv) \leq \mathcal{T}(u^p)^{1/p} \cdot \mathcal{T}(v^q)^{1/q},$$

valid for any nonnegative functions  $u, v \in L^p(E)$ , where  $q = p/(p-1)$ . The operator  $\mathcal{T}$  has a nonnegative kernel, and hence, its norm is attained at a nonnegative function  $u \in L^p(E)$ . By the pointwise Hölder inequality,

$$(\mathcal{T}u)^p \leq \mathcal{T}(u^p f^{-p/q}) \cdot (\mathcal{T}f)^{p/q} = \mathcal{T}(u^p f^{1-p}) \cdot (\mathcal{T}f)^{p-1}.$$

Therefore,

$$\|\mathcal{T}u\|_{L^p(E)}^p \leq \langle \mathcal{T}(u^p f^{1-p}), (\mathcal{T}f)^{p-1} \rangle = \langle u^p f^{1-p}, \mathcal{T}^*(\mathcal{T}f)^{p-1} \rangle = \lambda \langle u^p f^{1-p}, f^{p-1} \rangle = \lambda \|u\|_{L^p(E)}^p.$$

Thus, the norm of  $\mathcal{T} : L^p(E) \rightarrow L^p(E)$  is at most  $\lambda^{1/p}$ , which is the norm of  $S : L^p(\Omega_N) \rightarrow L^p(\Omega_N)$ . The two norms must therefore be equal, since the operator  $S$  is equivalent to the restriction of  $\mathcal{T}$  to the space of invariant functions.  $\square$

We move on to the ancestral part of the operator, which according to Proposition 2.2 is given by

$$K_1(x, y) = K(x, y) - K_0(x, y) = 1_{\{r=s \wedge t\}} \cdot q_s \cdot 2^{-t} \cdot \sum_{k=0}^{r-1} (2^r - 2^k) p_{s-1, k} \geq 0. \quad (32)$$

Write  $T_1$  for the operator whose kernel is  $K_1$ . As before, for an invariant function  $f(x) = F(d(x))$ , we may write

$$T_1 f(x) = \sum_{t=1}^N L_1(d(x), t) F(t) \quad (33)$$

for a certain kernel  $L_1(s, t)$  defined for  $s, t = 1, \dots, N$ . We also write

$$S_1 F(s) = \sum_{t=1}^N L_1(s, t) F(t).$$

It is possible to use formula (7) for the operator  $T$  and the definition (14) of the harmonic extension operator  $\tilde{T}$  and deduce that

$$S_0 + S_1 \equiv 0, \quad (34)$$

essentially because the only invariant, harmonic function on the vertices of the tree is the constant function. An alternative, more direct proof of (34) is provided in the following:

**Lemma 3.3.** *Let  $1 < p < \infty$ . Then the norm of the operator  $T_1 : L^p(E) \rightarrow L^p(E)$  is equal to the norm of  $S_1 : L^p(\Omega_n) \rightarrow L^p(\Omega_n)$ . In addition, for  $s, t = 1, \dots, N$  with  $m = s \wedge t$ , we have*

$$L_1(s, t) = q_s \cdot \sum_{k=0}^{m-1} (1 - 2^{k-m}) p_{s-1, k} = -L_0(s, t).$$

**Proof.** The first assertion of the lemma follows from the fact that the kernel of  $T_1$  is nonnegative and invariant under the symmetries of the tree, as in Lemma 3.2. For the second part, let  $s, t = 1, \dots, N$  and denote  $m = s \wedge t$ . We claim that for  $x \in E$  with  $d(x) = s$ ,

$$n(t; s, m) = \#\{y \in E; d(y) = t, \text{dlca}(x, y) = m\} = \max\{1, 2^{t-s}\}. \quad (35)$$

Indeed, assume first that  $t \geq s$ . How many  $y$ s are there with  $d(y) = t$  and  $\text{dlca}(x, y) = m$ ? Since  $m = s$ , the answer is  $2^{t-s}$ . Next, if  $s \geq t$ , then the number of such  $y$ s is one. This proves (35). Therefore,

$$\begin{aligned} L_1(s, t) &= \sum_{y \in E; d(y)=t} K_1(x, y) = n(t; s, m) \cdot q_s \cdot 2^{-t} \cdot \sum_{k=0}^{m-1} (2^m - 2^k) p_{s-1, k} \\ &= \max\{2^{-t}, 2^{-s}\} \cdot q_s \cdot \sum_{k=0}^{m-1} (2^m - 2^k) p_{s-1, k} \\ &= 2^{-m} \cdot q_s \cdot \sum_{k=0}^{m-1} (2^m - 2^k) p_{s-1, k}. \end{aligned} \quad \square$$

**Corollary 3.4.** *We have*

$$\|T\|_{L^p(E) \rightarrow L^p(E)} \leq 2\|S_0\|_{L^p(\Omega_N) \rightarrow L^p(\Omega_N)}.$$

**Proof.** This follows from the fact that  $T = T_0 + T_1$  together with the facts that  $\|T_0\| = \|S_0\|$  and  $\|T_1\| = \|S_1\|$ , while  $S_1 = -S_0$ .  $\square$

In view of Corollary 3.4, we are interested in bounds for the norm of the operator  $S = -S_0 = S_1 : L^p(\Omega_N) \rightarrow L^p(\Omega_N)$  whose nonnegative kernel is

$$L(s, t) = q_s \cdot \sum_{k=0}^{s \wedge t - 1} (1 - 2^{k-s \wedge t}) p_{s-1, k} \leq q_s \cdot \sum_{k=0}^{s \wedge t - 1} p_{s-1, k}. \quad (36)$$

From Lemma 2.1, we know that  $p_{s, r} = q_r \cdot \prod_{k=r+1}^s (1 - q_k)$  for  $s \leq N - 1$ . A little exercise in probability shows that for any  $1 \leq m \leq s \leq N$ ,

$$\sum_{k=0}^{m-1} p_{s-1, k} = \prod_{k=m}^{s-1} (1 - q_k). \quad (37)$$

Alternatively, (37) holds true for  $m = 0$  as  $q_0 = 1$ , and it may be proven by induction on  $m$  since

$$p_{s-1, m} + \prod_{k=m}^{s-1} (1 - q_k) = q_m \cdot \prod_{k=m+1}^{s-1} (1 - q_k) + \prod_{k=m}^{s-1} (1 - q_k) = \prod_{k=m+1}^s (1 - q_k).$$

From (36) and (37), we thus obtain

**Corollary 3.5.** *For  $s, t = 1, \dots, N$ , with  $m = \min\{s, t\}$ ,*

$$0 \leq L(s, t) \leq q_s \prod_{k=m}^{s-1} (1 - q_k),$$

where an empty product equals one.

*Some examples (parallel to the ones discussed in Section 2).*

(1) For the averaging operator, where  $q_s = 1$  and  $p_{s, r} = \delta_{s, r}$ , we have

$$L(s, t) = -\frac{1}{2} \cdot 1_{\{s \leq t\}},$$

i.e., this is the matrix whose entries equal 0 below the diagonal and  $-1/2$  on and above the diagonal. This is a rather simple matrix, and it is bounded with respect to the weighted  $L_p$ -norm for quite a few sequences of weights.

(2) For the symmetric random walk matrix, we have  $q_0 = 1$  while for  $1 \leq s \leq N$ ,

$$q_s = \frac{1}{N - s + 1}.$$

Hence, in view of Corollary 3.5, with  $m = \min\{s, t\}$ ,

$$0 \leq L(s, t) \leq \frac{1}{N-s+1} \prod_{k=m}^{s-1} \frac{N-k}{N-k+1} = \frac{1}{N-m+1}.$$

(3) In the case where

$$q_s = \frac{1}{N-s+1/\delta-1}$$

for some  $0 < \delta < 1$ , we have

$$0 \leq L(s, t) \leq \frac{1}{N-s+1/\delta-1} \prod_{k=m}^{s-1} \frac{N-k+1/\delta-2}{N-k+1/\delta-1} = \frac{1}{N-m+1/\delta-1}.$$

All that remains is to bound the  $L_p(\Omega_N)$ -norm of the operator whose kernel is discussed in Corollary 3.5. Recall from Lemma 2.3 that we have the freedom to choose the parameters  $q_1, \dots, q_{N-1} \in (0, 1)$  as we please.

How should we choose these parameters? Since  $L(N, t) \leq \prod_{k=t}^{N-1} (1 - q_k)$  and we are looking for upper bounds for the norm, the  $q_s$  should not be too tiny. On the other hand,  $L(s, t) \leq q_s$  for  $s \leq N-1$ , and hence, it is beneficial to choose  $q_s$  rather small. We would therefore need some balance for the  $q_s$ , which is the subject of the next section.

## 4 One-dimensional analysis

Let  $1 < p < \infty$ . It will be slightly more convenient to denote

$$\mathcal{K}(s, t) = L(N+1-s, N+1-t) \quad \text{and} \quad Q_s = q_{N+1-s}.$$

Recalling from (10) that  $q_N = 1$ , we see that

$$Q_1 = 1.$$

From Corollary 3.5, we know that for  $s, t = 1, \dots, N$ ,

$$\mathcal{K}(s, t) = L(N+1-s, N+1-t) \leq q_{N+1-s} \prod_{k=N+1-\max\{s,t\}}^{N-s} (1 - q_k) = Q_s \prod_{k=s+1}^{\max\{s,t\}} (1 - Q_k).$$

From Corollary 3.5, we know that  $L \geq 0$ . Consequently, for  $s, t = 1, \dots, N$ ,

$$0 \leq \mathcal{K}(s, t) \leq \begin{cases} Q_s & t \leq s \\ Q_s \cdot \prod_{k=s+1}^t (1 - Q_k) & t \geq s+1, \end{cases} \quad (38)$$

Recall that we are given edge weights  $W_1, \dots, W_N > 0$ , and that the associated  $L_p(E)$ -norm is given by (5). Denote

$$w_s := 2^{N+1-s} \cdot W_{N+1-s} > 0.$$

Consider the weighted  $\ell_p$ -norm

$$\|f\|_{p,w} = \left( \sum_{k=1}^N w_k |f(k)|^p \right)^{1/p}, \quad (39)$$

and the operator  $\mathcal{T}$  whose kernel is  $\mathcal{K}(s, t)$ . We are allowed to choose the weights  $q_1, \dots, q_{N-1} \in (0, 1)$  as we please, or equivalently, we have the freedom to determine  $Q_2, \dots, Q_N \in (0, 1)$ . We must keep  $Q_1 = 1$ . On the basis of the considerations related to the *Muckenhoupt criterion* discussed below, we set

$$Q_s = \frac{w_s^{-1/(p-1)}}{\sum_{k=1}^s w_k^{-1/(p-1)}}. \quad (40)$$

It is clear that  $Q_1 = 1$  and that  $Q_s \in (0, 1)$  for all  $s \geq 2$ . Recall that  $q = p/(p-1)$ .

**Lemma 4.1.** *To prove Theorem 1.2, it suffices to show that the operator norm of  $\mathcal{T}$  with respect to the  $\|\cdot\|_{p,w}$ -norm is bounded by a constant  $\hat{C}_p > 0$  depending only on  $p$ , where in fact*

$$\hat{C}_p \leq 2p^{1/p}q^{1/q} \cdot (1 + \max\{(p-1)^{-1/p}, (q-1)^{-1/q}\}). \quad (41)$$

**Proof.** In view of Corollary 3.4, it suffices to bound the operator norm of  $-S_0$ , whose kernel is  $L$ , with respect to the  $L^p(\Omega_N)$ -norm defined in (28). Under the transformation

$$s \mapsto N+1-s,$$

the operator  $-S_0$  whose kernel is  $L$  transforms to the operator  $\mathcal{T}$  whose kernel is  $\mathcal{K}$ . The  $L^p(\Omega_N)$ -norm from (28) transforms to the  $\|\cdot\|_{p,w}$ -norm defined in (39). Hence, Theorem 1.2 would follow once we obtain the bound (41), where  $\bar{C}_p \leq 2\hat{C}_p$  by Corollary 3.4.  $\square$

The remainder of this section is devoted to the proof of the following:

**Proposition 4.2.** *The operator norm of  $\mathcal{T}$  with respect to the norm (39) is bounded by a number  $\hat{C}_p$  depending only on  $p \in (1, \infty)$ . In fact, we have the bound (41) for the constant  $\hat{C}_p$ .*

Our main tool in the proof of Proposition 4.2 is the Muckenhoupt criterion [4], which is an indispensable tool for proving one-dimensional inequalities of the Poincaré-Sobolev type. For the reader's convenience, we include here a statement and a proof of a straightforward modification of the Muckenhoupt criterion, with sums in place of integrals:

**Theorem 4.3.** (Muckenhoupt) *Let  $1 < p < \infty$  and write  $\Omega_N = \{1, \dots, N\}$ . Let  $U, V : \Omega_N \rightarrow (0, \infty)$  and let  $A > 0$  be such that for all  $r = 1, \dots, N$ ,*

$$\left( \sum_{k=r}^N |U(k)|^p \right)^{1/p} \leq A \left( \sum_{k=1}^r |V(k)|^{-q} \right)^{-1/q}. \quad (42)$$

*Then for any function  $f : \Omega_N \rightarrow \mathbb{R}$ ,*

$$\left( \sum_{k=1}^N \left| U(k) \sum_{\ell=1}^k f(\ell) \right|^p \right)^{1/p} \leq C_p A \left( \sum_{k=1}^N |V(k)f(k)|^p \right)^{1/p}, \quad (43)$$

*with  $C_p = p^{1/p}q^{1/q}$ .*

By continuity, the analog of Theorem 4.3 for  $p = 1, \infty$  holds true with  $C_1 = C_\infty = 1$ . We remark that as in [4], this criterion is tight, in the sense that the infimum over all  $A > 0$  satisfying (42) is equivalent to the best constant in inequality (43). For the proof of Theorem 4.3, we require the following:

**Lemma 4.4.** *For  $1 < p < \infty$ ,  $q = p/(p-1)$ ,  $\alpha_1, \dots, \alpha_N > 0$  and  $r = 1, \dots, N$ ,*

$$\sum_{k=1}^r \alpha_k \left( \sum_{\ell=1}^k \alpha_\ell \right)^{-1/p} \leq q \cdot \left( \sum_{k=1}^r \alpha_k \right)^{1/q}. \quad (44)$$

**Proof.** We will use the simple inequality

$$(a+b)^{1/q} - a^{1/q} \geq b \cdot \min_{\xi \in (a, a+b)} \frac{\xi^{1/q-1}}{q} = \frac{1}{q} (a+b)^{1/q-1} \cdot b, \quad (45)$$

valid for any  $a, b \geq 0$  with  $a+b > 0$ . From (45), for  $k = 1, \dots, N$ ,

$$\left(\sum_{\ell=1}^k \alpha_{\ell}\right)^{1/q} - \left(\sum_{\ell=1}^{k-1} \alpha_{\ell}\right)^{1/q} \geq \frac{1}{q} \cdot \alpha_k \cdot \left(\sum_{\ell=1}^k \alpha_{\ell}\right)^{1/q-1},$$

where an empty sum equals zero. By summing this for  $k = 1, \dots, r$ , we obtain (44).  $\square$

**Proof of Theorem 4.3.** (Muckenhoupt) By the Hölder inequality, for any function  $f: \Omega_N \rightarrow \mathbb{R}$  and weights  $h: \Omega_N \rightarrow (0, \infty)$ ,

$$\begin{aligned} \sum_{k=1}^N \left| U(k) \sum_{\ell=1}^k f(\ell) \right|^p &\leq \sum_{k=1}^N U^p(k) \cdot \sum_{\ell=1}^k |f(\ell)V(\ell)h(\ell)|^p \cdot \left( \sum_{j=1}^k |V(j)h(j)|^{-q} \right)^{p/q} \\ &= \sum_{\ell=1}^N |f(\ell)V(\ell)h(\ell)|^p \sum_{k=\ell}^N U^p(k) \left( \sum_{j=1}^k |V(j)h(j)|^{-q} \right)^{p/q}. \end{aligned} \quad (46)$$

Set  $h(k) = (\sum_{\ell=1}^k V(\ell)^{-q})^{1/(pq)}$ . By applying Lemma 4.4 with  $\alpha_k = V(k)^{-q}$ , we obtain

$$\sum_{k=1}^r |V(k)h(k)|^{-q} = \sum_{k=1}^r \alpha_k \left( \sum_{\ell=1}^k \alpha_{\ell} \right)^{-1/p} \leq q \cdot \left( \sum_{k=1}^r \alpha_k \right)^{1/q} = q \cdot \left( \sum_{k=1}^r V(k)^{-q} \right)^{1/q}.$$

Hence, for any  $f: \Omega_N \rightarrow \mathbb{R}$ , the expression in (46) is at most

$$q^{p/q} \cdot \sum_{\ell=1}^N |f(\ell)V(\ell)h(\ell)|^p \sum_{k=\ell}^N U^p(k) \left( \sum_{j=1}^k |V(j)|^{-q} \right)^{p/q^2}. \quad (47)$$

By applying (42) and then Lemma 4.4 with  $\alpha_k = U^p(N+1-k)$  and with  $p \in (1, \infty)$  playing the rôle of  $q \in (1, \infty)$ , we see that

$$\begin{aligned} &\sum_{k=\ell}^N |U(k)|^p \left( \sum_{j=1}^k |V(j)|^{-q} \right)^{p/q^2} \\ &\leq A^{p/q} \sum_{k=\ell}^N |U(k)|^p \left( \sum_{j=k}^N |U(j)|^p \right)^{-1/q} \\ &= A^{p/q} \sum_{k=1}^{N+1-\ell} \alpha_k \left( \sum_{j=1}^k \alpha_j \right)^{-1/q} \leq p \cdot A^{p/q} \left( \sum_{k=1}^{N+1-\ell} \alpha_k \right)^{1/p} = p \cdot A^{p/q} \left( \sum_{k=\ell}^N U^p(k) \right)^{1/p}. \end{aligned}$$

Hence, the expression in (47) is at most

$$p \cdot (qA)^{p/q} \cdot \sum_{\ell=1}^N |f(\ell)V(\ell)h(\ell)|^p \cdot \left( \sum_{k=\ell}^N U^p(k) \right)^{1/p}.$$

By applying (42) again, we bound the last expression from above by

$$p \cdot (qA)^{p/q} \cdot A \cdot \sum_{\ell=1}^N |f(\ell)V(\ell)h(\ell)|^p \cdot \left( \sum_{k=1}^{\ell} V^{-q}(k) \right)^{-1/q} = pq^{p/q} A^p \sum_{\ell=1}^N |f(\ell)V(\ell)|^p,$$

hence completing the proof.  $\square$

**Corollary 4.5.** Let  $1 < p < \infty$  and  $\Omega_N = \{1, \dots, N\}$ . Let  $U, V: \Omega_N \rightarrow (0, \infty)$  and let  $A > 0$  be such that for all  $r = 1, \dots, N$ ,

$$\left( \sum_{k=1}^r |U(k)|^p \right)^{1/p} \leq A \left( \sum_{k=r}^N |V(k)|^{-q} \right)^{-1/q}. \quad (48)$$

Then for any  $f: \Omega_N \rightarrow \mathbb{R}$ ,

$$\left( \sum_{k=1}^N \left| U(k) \sum_{\ell=k}^N f(\ell) \right|^p \right)^{1/p} \leq C_p A \left( \sum_{k=1}^N |V(k)f(k)|^p \right)^{1/p}, \quad (49)$$

with  $C_p = p^{1/p} q^{1/q}$ .

**Proof.** Denote  $\tilde{U}(k) = U(N+1-k)$  and  $\tilde{V}(k) = V(N+1-k)$ . Then from (48), for all  $r = 1, \dots, N$ ,

$$\left( \sum_{k=r}^N |\tilde{U}(k)|^p \right)^{1/p} \leq A \left( \sum_{k=1}^r |\tilde{V}(k)|^{-q} \right)^{-1/q}.$$

By Theorem 4.3, this implies that for any  $g: \Omega_N \rightarrow \mathbb{R}$ , denoting  $f(r) = g(N+1-r)$ ,

$$\left( \sum_{k=1}^N \left| \tilde{U}(k) \sum_{\ell=1}^k f(N+1-\ell) \right|^p \right)^{1/p} \leq C_p A \left( \sum_{k=1}^N |\tilde{V}(k)f(N+1-k)|^p \right)^{1/p},$$

or equivalently,

$$\left( \sum_{k=1}^N \left| \tilde{U}(N+1-k) \sum_{\ell=k}^N f(\ell) \right|^p \right)^{1/p} \leq C_p A \left( \sum_{k=1}^N |\tilde{V}(N+1-k)f(k)|^p \right)^{1/p}.$$

This implies (49). □

**Proposition 4.6.** For  $f: \Omega_n \rightarrow \mathbb{R}$  and  $s = 1, \dots, N$  denote

$$\mathcal{T}_0 f(s) = Q_s \sum_{t=1}^s f(t),$$

where  $Q_s$  is defined in (40). Then the operator norm of  $\mathcal{T}_0$  with respect to the norm  $\|\cdot\|_{p,w}$  defined in (39) is bounded by a number  $\tilde{C}_p$  depending only on  $p \in (1, \infty)$ . In fact,  $\tilde{C}_p \leq 2^{1/p} p^{1/p} q^{1/q} \cdot \max\{1, (p-1)^{-1/p}\}$ .

The proof of Proposition 4.6 requires the following:

**Lemma 4.7.** For  $1 < p < \infty$ ,  $\alpha_1, \dots, \alpha_N > 0$  and  $r = 1, \dots, N$ ,

$$\sum_{k=r}^N \alpha_k \left( \sum_{\ell=1}^k \alpha_\ell \right)^{-p} \leq \frac{2}{\min\{1, p-1\}} \cdot \left( \sum_{k=1}^r \alpha_k \right)^{-p+1}. \quad (50)$$

**Proof.** We use the inequality

$$(p-1) \cdot b(a+b)^{-p} \leq a^{-p+1} - (a+b)^{-p+1},$$

which is valid for any  $a, b > 0$ . Then for  $k = 2, \dots, N$ ,

$$(p-1) \cdot \alpha_k \left( \sum_{\ell=1}^k \alpha_\ell \right)^{-p} \leq \left( \sum_{\ell=1}^{k-1} \alpha_\ell \right)^{-p+1} - \left( \sum_{\ell=1}^k \alpha_\ell \right)^{-p+1}.$$

By summing this for  $k = r+1, \dots, N$ , we obtain

$$(p-1) \cdot \sum_{k=r+1}^N \alpha_k \left( \sum_{\ell=1}^k \alpha_\ell \right)^{-p} \leq \left( \sum_{k=1}^r \alpha_k \right)^{-p+1} - \left( \sum_{\ell=1}^N \alpha_\ell \right)^{-p+1} \leq \left( \sum_{k=1}^r \alpha_k \right)^{-p+1},$$

where an empty sum equals zero. We conclude (50) by summing this with the trivial inequality

$$\min\{1, p-1\} \cdot \alpha_r \left( \sum_{\ell=1}^r \alpha_\ell \right)^{-p} \leq \left( \sum_{k=1}^r \alpha_k \right)^{-p+1}.$$

□

**Proof of Proposition 4.6.** Define

$$U(k) = w_k^{1/p} \cdot Q_k \quad \text{and} \quad V(k) = w_k^{1/p}.$$

Let us verify the condition of the Muckenhoupt criterion. We need to find  $A > 0$  such that for all  $r = 1, \dots, N$  inequality (42) holds true, that is,

$$\sum_{k=r}^N w_k Q_k^p \leq A^p \left( \sum_{k=1}^r w_k^{-q/p} \right)^{-p/q}. \quad (51)$$

Recall that  $p/q = p-1$ . From the definition (40) of  $Q_k$ , we need

$$\sum_{k=r}^N w_k^{-1/(p-1)} \left( \sum_{\ell=1}^k w_\ell^{-1/(p-1)} \right)^{-p} \leq A^p \left( \sum_{k=1}^r w_k^{-1/(p-1)} \right)^{-(p-1)}.$$

By setting  $\alpha_k = w_k^{-1/(p-1)}$  and using Lemma 4.7, we see that (51) holds true with

$$A = 2^{1/p} \cdot \max\{1, (p-1)^{-1/p}\}.$$

From Theorem 4.3, we thus conclude that for any  $f: \Omega_N \rightarrow \mathbb{R}$ ,

$$\left( \sum_{k=1}^N w_k \left| Q_k \sum_{\ell=1}^k f(\ell) \right|^p \right)^{1/p} \leq p^{1/p} q^{1/q} A \cdot \left( \sum_{k=1}^N w_k |f(k)|^p \right)^{1/p}.$$

This implies the required bound for the operator norm of  $f$ .

□

**Proposition 4.8.** For  $f: \Omega_n \rightarrow \mathbb{R}$  and  $s = 1, \dots, N$  denote

$$\mathcal{T}_1 f(s) = \sum_{t=s+1}^N \left[ Q_s \prod_{k=s+1}^t (1 - Q_k) \right] f(t),$$

where  $Q_s$  is as in (40). Then the operator norm of  $\mathcal{T}_1$  with respect to the norm  $\|\cdot\|_{p,w}$  defined in (39) is bounded by  $2^{1/q} p^{1/p} q^{1/q} \cdot \max\{1, (q-1)^{-1/q}\}$ .

**Proof.** Denote  $\alpha_k = w_k^{-1/(p-1)}$  and recall from (40) that  $Q_k = \alpha_k / \sum_{\ell=1}^k \alpha_\ell$ . We claim that for any  $t \geq s+1$ ,

$$Q_s \prod_{k=s+1}^t (1 - Q_k) = \frac{\alpha_s}{\sum_{k=1}^t \alpha_k}. \quad (52)$$

Indeed, (52) holds true for  $t = s$ , since an empty product equals one, and for  $t \geq s+1$ , it is proven by an easy induction on  $t$ . Consequently,

$$\mathcal{T}_1 f(s) = \alpha_s \sum_{t=s+1}^N \frac{1}{\sum_{j=1}^t \alpha_j} f(t).$$

Since  $p, q \in (1, \infty)$ , the elementary inequality of Lemma 4.7 is valid also when  $p$  is replaced by  $q$ . It implies that for  $r = 1, \dots, N$ ,

$$\left( \sum_{k=1}^r \alpha_k \right)^{-q+1} \geq \frac{\min\{1, q-1\}}{2} \cdot \sum_{k=r}^N \alpha_k \left( \sum_{j=1}^k \alpha_j \right)^{-q}. \quad (53)$$

Set  $A = 2^{1/q} \min\{1, q-1\}^{-1/q}$ . Since  $\alpha_k = w_k^{-1/(p-1)}$  and  $q/p = q-1 = 1/(p-1)$ , it follows from (53) that for  $r = 1, \dots, N$ ,



$$\left( \sum_{k=1}^r w_k \alpha_k^p \right)^{1/p} \leq A \left( \sum_{k=r}^N w_k^{-q/p} \left( \sum_{j=1}^k \alpha_j \right)^{-q} \right)^{-1/q}.$$

This is precisely the Muckenhoupt criterion from Corollary 4.5, with

$$U(k) = w_k^{1/p} \alpha_k \quad \text{and} \quad V(k) = w_k^{1/p} \sum_{j=1}^k \alpha_j.$$

Thus, by Corollary 4.5, for any  $g : \Omega_N \rightarrow \mathbb{R}$ ,

$$\sum_{k=1}^N w_k \left| \alpha_k \sum_{\ell=k}^N g(\ell) \right|^p \leq (C_p A)^p \sum_{k=1}^N w_k \left| \left( \sum_{j=1}^k \alpha_j \right) g(k) \right|^p, \quad (54)$$

with  $C_p = p^{1/p} q^{1/q}$ . By restricting attention to nonnegative functions  $g$  in (54), we may alter (54) and replace  $\sum_{\ell=k}^N$  by the shorter sum  $\sum_{\ell=k+1}^N$ . Inequality (54) remains correct, for nonnegative  $g$ , also after this modification. By denoting  $g(k) = f(k)/\sum_{j=1}^k \alpha_j$ , we conclude that for any nonnegative function  $f : \Omega_N \rightarrow \mathbb{R}$ ,

$$\sum_{k=1}^N w_k \left| \alpha_k \sum_{\ell=k+1}^N \frac{1}{\sum_{j=1}^{\ell} \alpha_j} f(\ell) \right|^p \leq (C_p A)^p \sum_{k=1}^N w_k |f(k)|^p. \quad (55)$$

Since the kernel of  $\mathcal{T}_1$  is nonnegative, its operator norm is attained at a nonnegative function  $f : \Omega_N \rightarrow \mathbb{R}$ . Therefore, (55) implies the required bound for the operator norm of  $\mathcal{T}_1$ .  $\square$

**Proof of Proposition 4.2.** The kernel of the operator  $\mathcal{T}$  is given in (38). It is a nonnegative kernel, and therefore, the operator norm of  $\mathcal{T}$  is at most the operator norm of the operator whose kernel is the expression on the right-hand side of (38). The latter operator equals

$$\mathcal{T}_0 + \mathcal{T}_1$$

with  $\mathcal{T}_0$  from Proposition 4.6 and  $\mathcal{T}_1$  from Proposition 4.8. From these two propositions, it follows that the operator norm of  $\mathcal{T}$  is at most

$$2^{1/p} p^{1/p} q^{1/q} \cdot \max\{1, (p-1)^{-1/p}\} + 2^{1/q} p^{1/p} q^{1/q} \cdot \max\{1, (q-1)^{-1/q}\}. \quad \square$$

Theorem 1.2 follows from Lemma 4.1 and Proposition 4.2.

#### Remark.

- (1) In this article, we have left open several natural questions, including the existence of linear extension operators for  $\dot{W}^{1,p}(T)$  for weighted trees  $T$  in the extreme cases  $p = 1, p = \infty$ , as well as the analog of our result for the inhomogeneous Sobolev space  $W^{1,p}(T)$  in place of  $\dot{W}^{1,p}(T)$ .
- (2) The problem of existence of linear Sobolev extension operators for weighted trees arose in connection with an extension problem for  $W^{2,p}(\mathbb{R}^2)$ . More precisely, given  $E \subseteq \mathbb{R}^2$ , let  $\dot{W}^{2,p}(E)$  denote the space of restrictions to  $E$  of functions in  $\dot{W}^{2,p}(\mathbb{R}^2)$ , endowed with the natural seminorm. Does there exist a linear extension operator from  $\dot{W}^{2,p}(E)$  to  $\dot{W}^{2,p}(\mathbb{R}^2)$ ? The answer is affirmative for  $p > 2$ ; see the study by Israel [3]. For  $1 < p < 2$ , the answer is unknown. For a particular class of examples  $E$ , the problem reduces to the question answered by Theorem 1.1.

**Acknowledgements:** We would like to thank Jacob Carruth, Arie Israel, Anna Skorobogatova and Ignacio Uriarte-Tuero for helpful conversations. This research was conducted while BK was visiting Princeton University's Department of Mathematics; he is grateful for their gracious hospitality.

**Funding information:** Charles Fefferman was supported by the Air Force Office of Scientific Research (Grant number FA9950-18-1-0069) and the National Science Foundation (NSF) (Grant number DMS-1700180). Bo'az Klartag was supported by a grant from the Israel Science Foundation (ISF).

**Conflict of interest:** The authors state no conflict of interest.

## References

- [1] A. Björn, J. Björn, J. T. Gill, and N. Shanmugalingam, *Geometric analysis on Cantor sets and trees*, J. Reine Angew. Math. **725** (2017), 63–114.
- [2] R. Howard and A. R. Schep, *Norms of positive operators on  $L^p$ -spaces*, Proc. Amer. Math. Soc. **109** (1990), no. 1, 135–146.
- [3] A. Israel, *A bounded linear extension operator for  $L^{2,p}(\mathbb{R}^2)$* , Ann. Math. two, **178** (2013), no. 1, 183–230.
- [4] B. Muckenhoupt, *Hardy's inequality with weights*, Studia Math. **44** (1972), 31–38.
- [5] P. Shvartsman, *Sobolev  $W_p^1$ -spaces on closed subsets of  $\mathbb{R}^n$* , Adv. Math. **220** (2009), no. 6, 1842–1922.