STUDIA MATHEMATICA 206 (1) (2011)

Compactness properties of weighted summation operators on trees—the critical case

by

Mikhail Lifshits (St. Petersburg) and Werner Linde (Jena)

Abstract. The aim of this paper is to provide upper bounds for the entropy numbers of summation operators on trees in a critical case. In a recent paper [Studia Math. 202 (2011)] we elaborated a framework of weighted summation operators on general trees where we related the entropy of the operator to those of the underlying tree equipped with an appropriate metric. However, the results were left incomplete in a critical case of the entropy behavior, because this case requires much more involved techniques. In the present article we fill this gap. To this end we develop a method, working in the context of general trees and general weighted summation operators, which was recently proposed by the first-named author for a particular critical operator on the binary tree. Those problems appeared in a natural way during the study of compactness properties of certain Volterra integral operators in a critical case.

1. Introduction. Let T be a tree with partial order structure "$\preceq$", i.e., $t \preceq s$ whenever t lies on a path leading from the root of T to s. Suppose we are given two weight functions $\alpha, \sigma : T \to (0,\infty)$ satisfying

$$\kappa := \sup_{s\in T} \Big(\sum_{v\preceq s} \alpha(v)^q \Big)^{1/q} \sigma(s) < \infty \tag{1.1}$$

for some $q \geq 1$. Then the weighted summation operator $V_{\alpha,\sigma}$ is well-defined by

$$(V_{\alpha,\sigma}\mu)(t) := \alpha(t) \sum_{s\succeq t} \sigma(s)\mu(s), \quad t \in T, \tag{1.2}$$

for $\mu \in \ell_1(T)$ and it is bounded from $\ell_1(T)$ to $\ell_q(T)$ with $\|V\| \leq \kappa$.

Our aim is to describe compactness properties of $V_{\alpha,\sigma}$. This turns out to be a challenging problem because those properties do not only depend on α, σ and q but also on the structure of the underlying tree. The motivation for the investigation of those questions stems from [Lif] where compactness

2010 *Mathematics Subject Classification*: Primary 47B06; Secondary 06A06, 05C05.
Key words and phrases: metrics on trees, operators on trees, weighted summation operators, covering numbers, entropy numbers.

DOI: 10.4064/sm206-1-6

properties of certain Volterra integral operators were studied; for the latter subject, see also [Lin]. Probabilistic applications of summation operators on trees may be found in [LL2].

The basic observation in [LL1] is as follows: The weights α and σ generate in a natural way a metric d on T, and covering properties of T by d-balls are tightly related to the degree of compactness of $V_{\alpha,\sigma}$. To be more precise, for $t \preceq s$ we define their distance as

$$d(t,s) := \max_{t \prec v \preceq s} \Big(\sum_{t \prec \tau \preceq v} \alpha(\tau)^q \Big)^{1/q} \sigma(v). \tag{1.3}$$

If σ is non-increasing, i.e., $t \preceq s$ implies $\sigma(t) \geq \sigma(s)$, then, as shown in [LL1], the distance d extends to a metric on the whole tree T.

Let $N(T,d,\varepsilon)$ be the *covering numbers* of (T,d), i.e.,

$$N(T,d,\varepsilon) := \inf \Big\{ n \geq 1 : T = \bigcup_{j=1}^{n} U_\varepsilon(t_j) \Big\}$$

with (open) ε-balls $U_\varepsilon(t_j)$ for certain $t_j \in T$, and let $e_n(V_{\alpha,\sigma})$ be the sequence of *dyadic entropy numbers* of $V_{\alpha,\sigma}$ defined as follows: If X and Y are Banach spaces and $V : X \to Y$ is an operator, then the nth *entropy number* of V is given by

$$e_n(V) := \inf\{\varepsilon > 0 : \{V(x) : \|x\|_X \leq 1\} \text{ is covered by at most } 2^{n-1} \text{ open } \varepsilon\text{-balls in } Y\}.$$

The operator V is compact if and only if $e_n(V) \to 0$ as $n \to \infty$. Thus the behavior of $e_n(V)$ as $n \to 0$ may be viewed as a measure of the degree of compactness of V. We refer to [CS] or [ET] for more information about entropy numbers and their properties.

One of the main results in [LL1] is the following; for simplicity, we only formulate it for $1 < q \leq 2$. Suppose that

$$N(T,d,\varepsilon) \leq c\varepsilon^{-a} |\log \varepsilon|^b$$

for some $a > 0$ and $b \geq 0$. Then

$$e_n(V_{\alpha,\sigma} : \ell_1(T) \to \ell_q(T)) \leq c' n^{-1/a-1/q} (\log n)^{b/a}.$$

Here the order of the right-hand side cannot be improved.

One may ask whether or not a similar relation between $N(T,d,\varepsilon)$ and $e_n(V_{\alpha,\sigma})$ remains valid in the (probably more interesting) case that $N(T,d,\varepsilon)$ tends to infinity exponentially as $\varepsilon \to 0$. In [LL1] we were able to provide a partial answer to this question, but some interesting critical case remained unsolved. More precisely, we proved the following (again we only formulate the result for $1 < q \leq 2$): Assume

$$\log N(T,d,\varepsilon) \leq c\varepsilon^{-a}$$

for some constant $c > 0$ and some $a > 0$. Then

$$e_n(V_{\alpha,\sigma} : \ell_1(T) \to \ell_q(T)) \leq c' n^{-1/q'} (\log n)^{1/q' - 1/a} \tag{1.4}$$

provided that $a < q'$ while for $a > q'$ we have

$$e_n(V_{\alpha,\sigma} : \ell_1(T) \to \ell_q(T)) \leq c' n^{-1/a}. \tag{1.5}$$

Here and throughout, q' denotes the conjugate of q defined by $1/q' = 1 - 1/q$. Again, both estimates are of the best possible order.

The most interesting critical case $a = q'$ has remained open; our methods only led to

$$e_n(V_{\alpha,\sigma} : \ell_1(T) \to \ell_q(T)) \leq c' n^{-1/a} (\log n). \tag{1.6}$$

In view of the available lower estimates and the results in [Lif], where a special but representative case (binary trees, $q = 2$ and some special weights) was handled, we conjectured that the logarithm on the right-hand side of (1.6) is unnecessary.

The main aim of the present paper is to verify this conjecture. We shall prove the following.

Theorem 1.1. *Let T be an arbitrary tree and let α, σ be weights on T satisfying* (1.1) *for some $q \in (1, 2]$. Furthermore, assume σ to be non-decreasing and let d be the metric on T defined via* (1.3)*. Suppose*

$$\log N(T, d, \varepsilon) \leq c_0 \varepsilon^{-q'} \tag{1.7}$$

with some $c_0 > 0$. Then

$$e_n(V_{\alpha,\sigma} : \ell_1(T) \to \ell_q(T)) \leq c c_0^{1-1/q} n^{-(1-1/q)}$$

for some $c = c(q)$ independent of α, σ and T.

As an illustration, let us show that the main result of [Lif] is a direct corollary of Theorem 1.1.

Corollary 1.2. *Let T be an infinite binary tree and let the weights on T be defined by*

$$\sigma(t) = 1, \quad \alpha(t) = (|t| + 1)^{-1}, \quad t \in T,$$

where $|\cdot|$ denotes the order of an element in the tree (*cf.* (2.1) *below*)*. Then there exists a finite positive c such that for all positive integers n we have*

$$e_n(V_{\alpha,\sigma} : \ell_1(T) \to \ell_2(T)) \leq c n^{-1/2}.$$

Proof. Take any positive integer k and let T_k be the union of the levels less than or equal to k, i.e., $T_k := \{t \in T : |t| \leq k\}$. Then

$$\# T_k = \sum_{j=0}^{k} 2^j < 2^{k+1}.$$

On the other hand, it is obvious that (with $q = 2$ in the definition of the metric d)

$$\sup_{s\in T}\inf_{t\in T_k} d(t,s)^2 < \sum_{j=k+1}^{\infty} (j+1)^{-2} \le (k+1)^{-1}.$$

It follows that $\log N(T,d,(k+1)^{-1/2}) \le (k+1)\cdot\log 2$, thus Theorem 1.1 applies with $q = 2$ and yields the assertion of the corollary. ∎

Let us briefly explain why the methods of [LL1] are *not* appropriate for the proof of Theorem 1.1 and why, therefore, a completely new approach is needed. A basic step in the proof of (1.4) and (1.5) is an estimate for the entropy of the convex hull of a certain subset of $\ell_q(T)$. Estimates of this type are well-known (see e.g. [CKP] or [CSt]). But here a critical case appears which exactly corresponds to $a = q'$ in our situation. As shown for $q = 2$ in [Ga] and for type q-spaces in [CSt], in this critical case estimates of the entropy of convex hulls give an extra log-term which, in general, cannot be avoided. Thus, in order to prove Theorem 1.1 one has to show that such an extra log-term does not appear for the entropy of convex hulls provided the sets under investigation are related to weighted summation operators on trees. This demands a completely new approach which was for the first time used in [Lif] and which we elaborate here further. The basic idea is to approximate an operator defined on a Banach space X (in our situation we have $X = \ell_1(T)$) by a family of operators depending on the elements in X. To this end one has to control at the same time the entropy numbers of the approximating operators as well as the number of those operators.

2. Trees. Let us recall some basic notation related to trees which will be used later on. In what follows, T always denotes a finite or an infinite tree. We suppose that T has a unique root which we denote by $\mathbf{0}$ and that each element $t \in T$ has a finite number of offsprings. We do not exclude that some elements do not possess any offspring, i.e., the progeny of some elements may "die out". The tree structure leads in a natural way to a partial order "$\preceq$" by letting $t \preceq s$, or $s \succeq t$, provided there are $t = t_0, t_1, \ldots, t_m = s$ in T such that for $1 \le j \le m$ the element t_j is an offspring of t_{j-1}. The strict inequalities have the same meaning with the additional assumption $t \ne s$. Two elements $t, s \in T$ are said to be *comparable* if either $t \preceq s$ or $s \preceq t$.

Given $t \in T$ with $t \ne \mathbf{0}$ we denote by t^- the (unique) parent element of t, i.e., t is supposed to be an offspring of t^-.

For $t, s \in T$ with $t \preceq s$ the order interval $[t,s]$ is defined by

$$[t,s] := \{v \in T : t \preceq v \preceq s\}$$

and in a similar way we define $(t,s]$ and (t,s).

A subset $B \subseteq T$ is said to be a *branch* provided that all elements in B are comparable and, moreover, if $t \preceq v \preceq s$ with $t, s \in B$, then $v \in B$ as well. Of course, finite branches are of the form $[t, s]$ for suitable $t \preceq s$.

A set $B \subseteq T$ is called a *tree* provided it is a tree with respect to the structure of T; in particular, if $r \in B$ is its root, then $[r, s] \subseteq B$ for each $s \in B$ with $s \succeq r$. If $\mathbf{0}$ is the root of the tree, then B is called a *subtree* of T. Given a tree $B \subseteq T$, an element $t \in B$ is said to be *terminal* provided that $s \notin B$ for all offsprings s of t.

Finally, for any $s \in T$ its *order* $|s| \geq 0$ is defined by

$$|s| := \#\{t \in T : t \prec s\}. \tag{2.1}$$

3. Reduction of the problem. An easy scaling argument shows that we may assume that estimate (1.7) holds with $c_0 = 1$. Another quantity that naturally appears in our bounds is κ defined in (1.1). Notice that for any $\varepsilon > \kappa$ we have $N(T, d, \varepsilon) \geq 2$. Therefore, (1.7) yields $\kappa \leq (c_0/\log 2)^{q'}$, hence the scaling of c_0 implies that $\kappa > 0$ is uniformly bounded. After this scaling is done, all the constants appearing in the proof of Theorem 1.1 only depend on q. We will denote them by c without further distinction.

First reduction. A first important simplification of the problem is as follows: Without losing generality we may assume that σ attains only values in $\{2^{-k} : k \in \mathbb{Z}\}$. Although this has been proved in [LL1], let us briefly repeat the argument. Set

$$I_k := \{t \in T : 2^{-k-1} < \sigma(t) \leq 2^{-k}\}$$

and define a new weight $\hat{\sigma}$ by

$$\hat{\sigma} := \sum_{k \in \mathbb{Z}} 2^{-k} \mathbf{1}_{I_k}.$$

Then $e_n(V_{\alpha,\sigma}) \leq e_n(V_{\alpha,\hat{\sigma}})$, and if $\hat{d}$ denotes the metric on T defined via (1.3) with $\hat{\sigma}$ instead of σ, then $N(T, \hat{d}, \varepsilon) \leq N(T, d, 2\varepsilon)$. Moreover, if σ is non-increasing, then so is $\hat{\sigma}$. Clearly, this shows that it suffices to prove Theorem 1.1 for weights σ only attaining values in $\{2^{-k} : k \in \mathbb{Z}\}$.

Second reduction. Suppose that σ is of the special form

$$\sigma = \sum_{k \in \mathbb{Z}} 2^{-k} \mathbf{1}_{I_k} \tag{3.1}$$

for certain disjoint $I_k \subseteq T$, $k \in \mathbb{Z}$. Since σ is assumed to be non-decreasing, the collection $\mathcal{I} := (I_k)_{k \in \mathbb{Z}}$ of subsets in T has the following properties:

(1) The I_k are disjoint and $T = \bigcup_{k \in \mathbb{Z}} I_k$, i.e., $\mathcal{I}$ is a partition of T.
(2) For each $s \in T$ the set $I_k \cap [\mathbf{0}, s]$ is either empty or a branch. Moreover, the I_k are ordered in the right way, i.e., if $l < k$ and $v_l \in I_l \cap [\mathbf{0}, s]$ and $v_k \in I_k \cap [\mathbf{0}, s]$, then $v_l \prec v_k$.

(3) Since σ is bounded, it follows of course that $I_k = \emptyset$ whenever $k \le k_0$ for a certain $k_0 \in \mathbb{Z}$.

Using this partition $\mathcal{I}$ of T we construct an operator $W : \ell_1(T) \to \ell_q(T)$ which may be viewed as a localization of $V_{\alpha,\sigma}$. It is defined as follows: If $\mu \in \ell_1(T)$, then

$$(3.2)\qquad (W\mu)(t) := \alpha(t) \sum_{\substack{s\succeq t\\ s\in I_k}} \sigma(s)\mu(s) = \alpha(t)\, 2^{-k} \sum_{\substack{s\succeq t\\ s\in I_k}} \mu(s), \qquad t \in I_k.$$

Note that for each $k \in \mathbb{Z}$ and $t \in I_k$ the value of $(W\mu)(t)$ depends only on the values of $\mu(s)$ for $s \succeq t$ and $s \in I_k$. This is in complete contrast to the definition of $V_{\alpha,\sigma}$ because here the value of $(V_{\alpha,\sigma}\mu)(t)$ depends on the values of $\mu(s)$ for *all* $s \succeq t$. Nevertheless, as shown in [LL1, Proposition 4.3] the following is valid.

Proposition 3.1. *We have*

$$e_n(V_{\alpha,\sigma} : \ell_1(T) \to \ell_q(T)) \le 2e_n(W : \ell_1(T) \to \ell_q(T)).$$

As a consequence, it suffices to estimate the entropy numbers of W suitably.

Third reduction. Given $\varepsilon > 0$ a set $S \subset T$ is said to be an *ε-order net* if for each $t \in T$ there is an $s \in S$ with $s \preceq t$ and $d(s,t) < \varepsilon$. Let

$$\tilde{N}(T,d,\varepsilon) := \inf\{\#S : S \text{ is an } \varepsilon\text{-order net of } T\}$$

be the corresponding order covering numbers. Clearly,

$$N(T,d,\varepsilon) \le \tilde{N}(T,d,\varepsilon).$$

But, surprisingly, also a reverse estimate holds as shown in [LL1, Proposition 3.3]. More precisely, we always have

$$\tilde{N}(T,d,2\varepsilon) \le N(T,d,\varepsilon).$$

Summing up, it follows that it suffices to prove the following variant of Theorem 1.1:

Theorem 3.2. *Suppose that σ is non-increasing and of the special form* (3.1). *Define W as in* (3.2) *with respect to the partition $\mathcal{I}$ generated by σ. If*

$$\log \tilde{N}(T,d,\varepsilon) \le \varepsilon^{-q'}$$

for some $q \in (1,2]$, then

$$e_n(W : \ell_1(T) \to \ell_q(T)) \le cn^{-(1-1/q)}.$$

4. Proof of Theorem 3.2. We start by explaining the strategy for proving Theorem 3.2. This is done by a quite general approximation procedure, which, to our knowledge, for the first time appeared in [Lif].

PROPOSITION 4.1. *Let V be a bounded linear operator between the Banach spaces X and Y and let $\{V_\gamma : \gamma \in \Gamma\}$ be a (finite) collection of operators from X to Y. Set $M := [\log_2(\#\Gamma)] + 1$. Then, for each $k \geq 1$,*

$$e_{k+M}(V) \leq \sup_{\gamma\in\Gamma} e_k(V_\gamma) + \sup_{\|x\|_X\leq 1} \inf_{\gamma\in\Gamma} \|Vx - V_\gamma x\|_Y. \tag{4.1}$$

How this proposition is applied? Let $a > 0$ and suppose for each $n \geq 1$ there exist operators $\{V_\gamma^n : \gamma \in \Gamma_n\}$ from X to Y such that $\log(\#\Gamma_n) \leq c_1 n$ and $e_{[\rho n]}(V_\gamma^n) \leq c_2 n^{-a}$ for some $\rho \geq 1$. If, furthermore, for each $x \in X$ with $\|x\|_X \leq 1$ there is a $\gamma = \gamma(x) \in \Gamma_n$ with

$$\|Vx - V_\gamma^n x\|_Y \leq c_3 n^{-a},$$

then an application of Proposition 4.1 with $k = [\rho n]$ immediately leads to $e_{c_4 n}(V) \leq c_5 n^{-a}$ for $n \in \mathbb{N}$, hence by the monotonicity of entropy numbers, $e_n(V) \leq c n^{-a}$ for $n \geq 1$.

Thus, in order to apply this general approximation scheme to W defined in (3.2) and with $a = 1 - 1/q$, for each $n \geq 1$ we have to construct a suitable collection $\{W_\gamma^n : \gamma \in \Gamma_n\}$ of operators from $\ell_1(T)$ to $\ell_q(T)$ with $\log(\#\Gamma_n) \leq c_1 n$,

$$\inf_{\gamma\in\Gamma_n} \|W\mu - W_\gamma^n \mu\|_q \leq c_3 n^{-(1-1/q)}, \qquad \|\mu\|_1 \leq 1, \tag{4.2}$$

such that

$$e_{[\rho n]}(W_\gamma^n) \leq c_2 n^{-(1-1/q)}, \qquad n \geq 1, \tag{4.3}$$

for a certain $\rho \geq 1$.

Let us briefly describe the strategy of this quite involved construction. In a first step, we build an auxiliary structure on the tree T. Namely, we construct a system $(\mathcal{B}_m)_{m\geq 0}$ of refining tree partitions of T based on the weights α and σ. This is done in Subsections 4.1 and 4.2.

Next, given $n \in \mathbb{N}$, we construct a set $\mathbb{L}_n$ of partitions of T that will play the role of the parameter set Γ_n mentioned above. Namely, for any $\mu \in \ell_1(T)$ with $\|\mu\|_1 \leq 1$ we construct a special partition $\mathcal{L}_\mu = \mathcal{L}_\mu(n)$ of T. Any element of the partition $\mathcal{L}_\mu$ belongs to a suitable partition $\mathcal{B}_m$ built in the first step. This construction is presented in Subsection 4.3. We let $\mathbb{L}_n = \{\mathcal{L}_\mu : \mu \in \ell_1(T),\ \|\mu\|_1 \leq 1\}$. The size of $\mathbb{L}_n$ has an exponential bound as required above.

Furthermore, each partition $\mathcal{L} \in \mathbb{L}_n$ generates a representation $W = \sum_{i=1}^4 W_{\mathcal{L}}^i$, as explained in (4.22) below. We show that the sums $\sum_{i=1}^3 W_{\mathcal{L}}^i$ can be used as approximating operators as in (4.2) and they admit the bound for the entropy numbers as in (4.3). Algebraic properties of the entropy numbers imply that it suffices to verify $e_n(W_{\mathcal{L}}^i) \leq c_i n^{-(1-1/q)}$ for $i = 1, 2, 3$. The proof of these estimates will be presented in Subsection 4.4. Surprisingly, each of the three operators must be treated by a different method.

Hence, let us start with the investigation of a special type of partition of T.

4.1. Tree partitions. Suppose we are given a subset $R \subseteq T$ with $\mathbf{0} \in R$. If $r \in R$, set

$$B_r := \{s \in T : s \succeq r \text{ and } (r, s] \cap R = \emptyset\}.$$

Then B_r is a tree in T with root $r \in R$. Letting $\mathcal{B} := \{B_r : r \in R\}$, the family $\mathcal{B}$ is a partition of T where each partition element is a tree. We call $\mathcal{B}$ a *tree partition* of T. Notice that each tree partition of T may be represented as described before with R being the set of roots of $B \in \mathcal{B}$.

Given two tree partitions $\mathcal{B}_1$ and $\mathcal{B}_2$, we say that $\mathcal{B}_2$ refines $\mathcal{B}_1$ provided that each $B_2 \in \mathcal{B}_2$ is contained in a suitable $B_1 \in \mathcal{B}_1$. Clearly, this is equivalent to $R_1 \subseteq R_2$ with generating (or root) sets R_1 and R_2 of $\mathcal{B}_1$ and $\mathcal{B}_2$, respectively.

Suppose now that $(\mathcal{B}_m)_{m \geq 0}$ is a sequence of tree partitions satisfying $\mathcal{B}_0 = \{T\}$ such that for each $m \geq 1$ the partition $\mathcal{B}_m$ refines $\mathcal{B}_{m-1}$. Of course, this is equivalent to

$$\{\mathbf{0}\} = R_0 \subseteq R_1 \subseteq \cdots$$

with the R_m being the corresponding root sets. In order to distinguish the sets in different levels let us write

$$\mathcal{B}_m = \{B_{r,m} : r \in R_m\}$$

where $B_{r,m}$ is an element of $\mathcal{B}_m$ with the root $r \in R_m$. Finally, set

$$\mathcal{B}_\infty := \{B_{r,m} : r \in R_m,\, m \geq 0\}. \tag{4.4}$$

Given $B_{r,m}$ and $B_{r',m'}$ in $\mathcal{B}_\infty$ we say that the latter tree is an offspring of the former one provided that $m' = m + 1$ and $B_{r',m'} \subseteq B_{r,m}$. In that way $\mathcal{B}_\infty$ becomes a tree with root $B_{\mathbf{0},0} = \{T\}$. If we denote the generated partial order in $\mathcal{B}_\infty$ by "$\trianglelefteq$" (and by "$\triangleleft$" the strict order), then $B_{r,m} \trianglelefteq B_{r',m'}$ if and only if $m' \geq m$ and $B_{r',m'} \subseteq B_{r,m}$. Notice a minor abuse of notation: We may have $B_{r,m} = B_{r,m'}$ as sets but they are equal in $\mathcal{B}_\infty$ only if $m = m'$, i.e., the same set may appear in different levels and is then treated as a multitude of different elements in $\mathcal{B}_\infty$.

In particular, all notions concerning trees apply to $\mathcal{B}_\infty$. For example, if we define the order of an element in $\mathcal{B}_\infty$ as in (2.1), then we have

$$\{B \in \mathcal{B}_\infty : |B| = m\} = \mathcal{B}_m.$$

Let us mention a special property of $\mathcal{B}_\infty$. Suppose that $m' \geq m$. Then, if $B, B' \in \mathcal{B}_\infty$ have orders m, m', respectively, then either $B \trianglelefteq B'$ or $B \cap B' = \emptyset$.

4.2. Construction of tree partitions. Suppose we are given weights α and σ on T where σ is assumed to be as in (3.1) with partition $\mathcal{I} = (I_k)_{k \in \mathbb{Z}}$. Let d be the metric constructed by (1.3) with respect to α, σ and q. One of the main difficulties is that the metric d and the partition $\mathcal{I}$ do not match. For example, if $t \preceq s$, $t \in I_{k-1}$ and $s \in I_k$, then we get

$$d(t,s) = \max\Big\{2^{-(k-1)}\Big(\sum_{t \prec v \prec \lambda(s)} \alpha(v)^q\Big)^{1/q}, 2^{-k}\Big(\sum_{t \prec v \preceq s} \alpha(v)^q\Big)^{1/q}\Big\} \tag{4.5}$$

where $\lambda(s)$ is defined by $I_k \cap [\mathbf{0}, s] = [\lambda(s), s]$. Since we do not have any information about the inner sums, this expression is difficult to handle. Observe that for general $t, s \in T$ with $t \preceq s$ expression (4.5) becomes even more complicated. Therefore we modify d in a way better suited to $\mathcal{I}$ and set

$$d_{\mathcal{I}}(t,s) := \begin{cases} \min\{d(\lambda(s), s), d(t,s)\}, & t \preceq s, \\ +\infty, & \text{otherwise}, \end{cases}$$

with $\lambda(s)$ as before. Let us reformulate this expression. To this end we define an equivalence relation on T by setting $t \equiv s$ provided that there is a $k \in \mathbb{Z}$ with $t, s \in I_k$. Then, if $t \preceq s$, we may write $d_{\mathcal{I}}(t,s)$ as

$$d_{\mathcal{I}}(t,s) = \sigma(s)\Big(\sum_{\substack{t \prec v \preceq s \\ v \equiv s}} \alpha(v)^q\Big)^{1/q}.$$

Thus $d_{\mathcal{I}}$ may be viewed as a localization of d. Although in general $d_{\mathcal{I}}$ is not a metric on T, we will use it later on to measure some "distances".

We suppose now that T is a tree satisfying

$$\log \tilde{N}(T, d, \varepsilon) \le \varepsilon^{-q'}. \tag{4.6}$$

The next objective is to construct tree partitions suited to our problem. We do so by defining the corresponding root sets. For each $m \ge 1$ set

$$\varepsilon_m := (m \log 2)^{-(1-1/q)}.$$

In view of (4.6) this choice immediately provides

$$\tilde{N}(T, d, \varepsilon_m) \le 2^m. \tag{4.7}$$

Proposition 4.2. *Suppose* (4.6). *Then there are subsets* $(R_m)_{m \ge 0}$ *of* T *with the following properties:*

(1) $\{\mathbf{0}\} = R_0 \subseteq R_1 \subseteq \cdots$.

(2) *For each* $m \ge 0$ *one has*

$$\# R_m \le 2^{m+1}. \tag{4.8}$$

(3) *The sets* R_m *are* ε_m*-order nets with respect to* $d_{\mathcal{I}}$*, i.e.*

$$\sup_{s \in T} \min_{r \in R_m} d_{\mathcal{I}}(r, s) < \varepsilon_m. \tag{4.9}$$

(4) *The sets* R_m *are minimal with respect to the order in* T *in the following sense: Whenever* $\tau \in R_m \setminus R_{m-1}$*, hence* $\tau \ne \mathbf{0}$*, thus* τ^- *is well-defined, then* $R_m^\tau := (R_m \setminus \{\tau\}) \cup \{\tau^-\}$ *no longer satisfies* (4.9).

Proof. We construct the sets R_m by induction. Take $R_0 := \{\mathbf{0}\}$ and suppose that for some $m \geq 1$ we have already defined $R_0, \ldots, R_{m-1}$ with the desired properties. Let

$$\mathcal{A}_m := \Big\{ A \subseteq T : \#A \leq 2^m \text{ and } \sup_{s \in T} \min_{r \in A \cup R_{m-1}} d_{\mathcal{I}}(r,s) < \varepsilon_m \Big\}.$$

First of all, we establish that $\mathcal{A}_m \neq \emptyset$. Indeed, due to (4.7) there exist ε_m-order nets of cardinality less than or equal to 2^m. Moreover, any such net belongs to $\mathcal{A}_m$, due to the definition of order nets and since $d_{\mathcal{I}}(r,s) \leq d(r,s)$ whenever $r \preceq s$.

Next define

$$\mathcal{A}_m^0 := \{A \in \mathcal{A}_m : \#A \text{ is minimal}\}$$

and distinguish between the following two cases:

First case: $\emptyset \in \mathcal{A}_m^0$. This happens whenever R_{m-1} satisfies (4.9) not only for ε_{m-1} but also for ε_m. In that case we set $R_m := R_{m-1}$. Of course, the first three properties are satisfied and the fourth one holds trivially.

Second case: $\emptyset \notin \mathcal{A}_m^0$. Then all sets in $\mathcal{A}_m^0$ have the same positive cardinality $p \leq 2^m$. For any $A \in \mathcal{A}_m^0$ let $F(A) := \sum_{\tau \in A} |\tau|$ and choose a set $A_m^* \in \mathcal{A}_m^0$ such that

$$F(A_m^*) = \min_{A \in \mathcal{A}_m^0} F(A).$$

Set

$$R_m := R_{m-1} \cup A_m^*.$$

Clearly, $R_{m-1} \subseteq R_m$. Next,

$$\#R_m = \#R_{m-1} + \#A_m^* \leq 2^m + p \leq 2^m + 2^m = 2^{m+1},$$

as asserted in property (2). Since $A_m^* \in \mathcal{A}_m$, condition (4.9) required in property (3) holds as well.

It remains to check property (4). Fix any $\tau \in A_m^*$. Note that necessarily $\tau \notin R_{m-1}$ because otherwise, by dropping τ, we could diminish the cardinality of the set and stay in $\mathcal{A}_m$. In particular, we get $\tau \neq \mathbf{0}$.

Next, set $A_m^\tau := (A_m^* \setminus \{\tau\}) \cup \{\tau^-\}$. Clearly, $F(A_m^\tau) < F(A_m^*)$. Since F attains its minimum on $\mathcal{A}_m^0$ at A_m^*, we infer that $A_m^\tau \notin \mathcal{A}_m^0$. Moreover, since $\#A_m^\tau = \#A_m^* = p$, it follows that $A_m^\tau \notin \mathcal{A}_m$. Consequently, because of $A_m^* = R_m \setminus R_{m-1}$, property (4) of the proposition holds. This completes the proof. ∎

Before proceeding further, let us recall that the R_m constructed above lead to tree partitions $\mathcal{B}_m$ of T with $\mathcal{B}_m = \{B_{r,m} : r \in R_m\}$ where $s \in B_{r,m}$ if and only if $s \succeq r$ and $(r,s] \cap R_m = \emptyset$. We have

$$d_{\mathcal{I}}(r,s) < \varepsilon_m \quad \text{whenever} \quad s \in B_{r,m}. \tag{4.10}$$

Moreover, each partition $\mathcal{B}_m$ refines $\mathcal{B}_{m-1}$.

Next we turn to the crucial estimate for the tree partitions constructed in Proposition 4.2. Here property (4) plays an important role.

PROPOSITION 4.3. *Let the R_m be as in Proposition 4.2 with corresponding tree partitions $\mathcal{B}_m = \{B_{r,m} : r \in R_m\}$. Fix $m \geq 1$ and let $r \in R_{m-1}$, $\tau \in R_m$ with $r \prec \tau$ and $\tau \in B_{r,m-1}$. Then*

$$\sigma(\tau)\Big(\sum_{\substack{r \prec v \preceq \tau^- \\ v \equiv \tau}} \alpha(v)^q\Big)^{1/q} \leq cm^{-1}.$$

Proof. First note that $\tau \in B_{r,m-1}$ yields always $\tau \notin R_{m-1}$, so $\tau \in R_m \setminus R_{m-1}$ and we are in the situation of property (4) of Proposition 4.2. Hence there is an $s \in T$ such that with $R_m^\tau := (R_m \setminus \{\tau\}) \cup \{\tau^-\}$ we get

$$\min_{v \in R_m^\tau} d_{\mathcal{I}}(v,s) \geq \varepsilon_m. \tag{4.11}$$

On the other hand, by property (3) of Proposition 4.2,

$$\min_{v \in R_m} d_{\mathcal{I}}(v,s) < \varepsilon_m. \tag{4.12}$$

Since R_m and R_m^τ differ only by one point, namely, by τ or τ^-, the two minima in (4.11) and (4.12) may only be different if τ is the maximal element in $[\mathbf{0}, s] \cap R_m$. This implies $s \in B_{\tau,m}$ as well as

$$d_{\mathcal{I}}(\tau^-, s) \geq \varepsilon_m \quad \text{and} \quad d_{\mathcal{I}}(\tau, s) < \varepsilon_m. \tag{4.13}$$

Furthermore, $\tau \equiv s$. Indeed, if they were not equivalent, then this would imply

$$d_{\mathcal{I}}(\tau, s) = d(\lambda(s), s) = d_{\mathcal{I}}(\tau^-, s),$$

which contradicts (4.13). Recall that for $s \in I_k$ we denoted by $\lambda(s)$ the minimal element in $[\mathbf{0}, s] \cap I_k$. Now from $s \equiv \tau$ we derive

$$\varepsilon_m^q \leq d_{\mathcal{I}}(\tau^-, s)^q = \sigma(s)^q \sum_{\tau^- \prec v \preceq s} \alpha(v)^q.$$

On the other hand, we have $s \in B_{\tau,m} \subseteq B_{r,m-1}$, hence by (4.10),

$$\sigma(s)^q \sum_{\substack{r \prec v \preceq s \\ v \equiv s}} \alpha(v)^q = d_{\mathcal{I}}(r,s)^q < \varepsilon_{m-1}^q.$$

By subtraction we obtain

$$\begin{aligned}
\sigma(\tau)^q \sum_{\substack{r \prec v \preceq \tau^- \\ v \equiv \tau}} \alpha(v)^q &= \sigma(s)^q \sum_{\substack{r \prec v \preceq \tau^- \\ v \equiv s}} \alpha(v)^q \\
&= \sigma(s)^q \Big(\sum_{\substack{r \prec v \preceq s \\ v \equiv s}} \alpha(v)^q - \sum_{\tau^- \prec v \preceq s} \alpha(v)^q \Big) \\
&\leq \varepsilon_{m-1}^q - \varepsilon_m^q = c((m-1)^{-(q-1)} - m^{-(q-1)}) \leq cm^{-q},
\end{aligned}$$

as required. ■

4.3. Heavy and light domains. Suppose we are given a sequence $\mathcal{B}_m$ of tree partitions as before, i.e., $\mathcal{B}_0 = \{T\}$ and $\mathcal{B}_m$ refines $\mathcal{B}_{m-1}$. For the moment those tree partitions may be quite general, but later on we will take the special partitions constructed via Proposition 4.2.

We fix a number $n \geq 1$; everything done now will depend on this number n (although we do not reflect this dependence in the notation). Let $\mu \in \ell_1(T)$ be an arbitrary non-zero element with $\|\mu\|_1 \leq 1$. Define $\mathcal{B}_\infty$ as in (4.4). A subset $B \in \mathcal{B}_\infty$ is said to be *heavy* (with respect to μ) provided that

$$|\mu|(B) > |B|/n;$$

here and later on $|\mu|(B) := \sum_{t\in B} |\mu(t)|$. Recall that $|B| = m$ means that $B \in \mathcal{B}_m$. Otherwise, i.e., if

$$|\mu|(B) \leq |B|/n, \tag{4.14}$$

we call $B \in \mathcal{B}_\infty$ *light*. If

$$\mathcal{B}_\mu^\bullet := \{B \in \mathcal{B}_\infty : B \text{ is heavy with respect to } \mu\}, \tag{4.15}$$

it follows that $\mathcal{B}_\mu^\bullet \subseteq \{B \in \mathcal{B}_\infty : |B| \leq n\}$. We have $B_{\mathbf{0},0} \in \mathcal{B}_\mu^\bullet$, and moreover, whenever $B \in \mathcal{B}_\mu^\bullet$ and $B' \in \mathcal{B}_\infty$ satisfy $B' \trianglelefteq B$, then this implies $B' \in \mathcal{B}_\mu^\bullet$ as well. In other words, $\mathcal{B}_\mu^\bullet$ is a subtree of $\mathcal{B}_\infty$ and we call it, in accordance with the terminology of [Lif], the *essential tree* in $\mathcal{B}_\infty$ (with respect to μ and n).

Among the light subsets of $\mathcal{B}_\infty$ we choose the extremal ones as follows:

$$\mathcal{L}_\mu := \{L \in \mathcal{B}_\infty : L \text{ is light and all } B \in \mathcal{B}_\infty \text{ with } B \triangleleft L \text{ are heavy}\}. \tag{4.16}$$

In other words, a set $B_{r,m} \in \mathcal{B}_\infty$ belongs to $\mathcal{L}_\mu$ if and only if $B_{r,m}$ is light and each $B_{r',m'} \in \mathcal{B}_\infty$ with $m' < m$ and $B_{r,m} \subseteq B_{r',m'}$ is heavy. Of course, it suffices if this property is valid for $m' = m - 1$.

PROPOSITION 4.4. *The set $\mathcal{L}_\mu \subseteq \mathcal{B}_\infty$ is a tree partition of T.*

Proof. First we show that for $L, L' \in \mathcal{L}_\mu$ we have either $L = L'$ or $L \cap L' = \emptyset$. Thus suppose $L \neq L'$. If $L \in \mathcal{B}_m$ and $L' \in \mathcal{B}_{m'}$, then $m = m'$ yields $L \cap L' = \emptyset$ and we are done. Assume now $m' < m$. In that case either $L \cap L' = \emptyset$ or $L' \triangleleft L$. But since $L \in \mathcal{L}_\mu$ and L' is light, the latter case is impossible and this shows $L \cap L' = \emptyset$ as asserted.

Take now an arbitrary $s \in T$ and let $B_m(s) \in \mathcal{B}_m$ be the unique set in $\mathcal{B}_m$ with $s \in B_m(s)$. Then

$$T = B_0(s) \trianglelefteq B_1(s) \trianglelefteq \cdots .$$

Since $B_0(s)$ is heavy, there is a smallest $m_0 = m_0(s)$ such that $B_{m_0}(s)$ is light. Of course, $B_{m_0}(s) \in \mathcal{L}_\mu$, showing that $T = \bigcup_{L\in\mathcal{L}_\mu} L$. This completes the proof. ■

REMARK. Observe that $\mathcal{L}_\mu$ is completely determined by $\mathcal{B}_\mu^\bullet$. Indeed, we have $L \in \mathcal{L}_\mu$ if and only if $L \notin \mathcal{B}_\mu^\bullet$ and there is a $B \in \mathcal{B}_\mu^\bullet$ such that L is an offspring of B. In other words, given $\mu_1, \mu_2 \in \ell_1(T)$, we have $\mathcal{L}_{\mu_1} = \mathcal{L}_{\mu_2}$ if and only if $\mathcal{B}_{\mu_1}^\bullet = \mathcal{B}_{\mu_2}^\bullet$.

The size of each essential tree and the total number of all possible essential trees (if we let μ vary) are strongly bounded. For completeness, let us repeat the corresponding arguments from [Lif].

LEMMA 4.5. *Let $\mu \in \ell_1(T)$ with $\|\mu\|_1 \le 1$ and denote by $\mathcal{Q} \subseteq \mathcal{B}_\infty$ the set of terminal domains of the subtree $\mathcal{B}_\mu^\bullet$. Then*

$$\sum_{B\in\mathcal{Q}} |B| < n. \tag{4.17}$$

Moreover,

$$\#\mathcal{B}_\mu^\bullet \le n. \tag{4.18}$$

Proof. Since all terminal domains are disjoint and they are all heavy, we have

$$1 \ge \|\mu\|_1 \ge \sum_{B\in\mathcal{Q}} |\mu|(B) > \sum_{B\in\mathcal{Q}} \frac{|B|}{n}.$$

It follows that $\sum_{B\in\mathcal{Q}} |B| < n$, as asserted in (4.17).

Since any node in a finite tree precedes at least one terminal node, we have

$$\begin{aligned}\#\mathcal{B}_\mu^\bullet &= 1 + \sum_{B\in\mathcal{B}_\mu^\bullet,\,|B|>0} 1 \le 1 + \sum_{B\in\mathcal{B}_\mu^\bullet,\,|B|>0} \#\{B' \in \mathcal{Q} : B \trianglelefteq B'\} \\ &= 1 + \sum_{B'\in\mathcal{Q}} \#\{B : |B| > 0, B \trianglelefteq B'\} = 1 + \sum_{B'\in\mathcal{Q}} |B'|,\end{aligned}$$

thus (4.18) follows from (4.17). ∎

Till now the sequence $(\mathcal{B}_m)_{m\ge0}$ of tree partitions could be quite general, i.e. we only assumed $\mathcal{B}_0 = \{T\}$ and that $\mathcal{B}_m$ refines $\mathcal{B}_{m-1}$ for $m \ge 1$. To proceed we have to know something about the size of the sets $\mathcal{B}_m$. In particular, this is the case if the $\mathcal{B}_m$ are constructed from root sets R_m with the properties of Proposition 4.2. Thus let us deal with those special tree partitions.

LEMMA 4.6. *The number of subtrees of $\mathcal{B}_\infty$ whose terminal set $\mathcal{Q}$ satisfies* (4.17) *does not exceed* $(8e)^n$.

Proof. Since a subtree is entirely defined by its terminal set, we have to find out how many sets $\mathcal{Q}$ satisfy (4.17). Denote $q_m = \#\{\mathcal{Q} \cap \mathcal{B}_m\}$. Then

(4.17) reads

$$\sum_m m q_m < n. \tag{4.19}$$

Since $q_m < n/m$, the number of non-negative integer solutions of this inequality does not exceed

$$\prod_{m=1}^{n-1}\left(1+\frac{n}{m}\right) \le \prod_{m=1}^{n-1}\frac{2n}{m} = \frac{(2n)^n}{n!} \le \frac{(2n)^n}{(n/e)^n} = (2e)^n. \tag{4.20}$$

Recall the bound for the size levels (4.8), which yields

$$\#\mathcal{B}_m = \#R^m \le 2^{m+1}.$$

Thus, for a given sequence q_m, while constructing a terminal set $\mathcal{Q}$, on the mth level $\mathcal{B}_m$ we have to choose q_m elements from at most 2^{m+1} elements of $\mathcal{B}_m$. Therefore, because of (4.19), the number of possible sets does not exceed

$$\prod_{m=1}^{n-1}\binom{2^{m+1}}{q_m} \le \prod_{m=1}^{n-1}(2^{m+1})^{q_m} = 2^{\sum_{m=1}^{n-1}(m+1)q_m} \le 2^{2n}. \tag{4.21}$$

Combining (4.20) and (4.21) leads to the desired estimate. ∎

4.4. Approximating operators. As mentioned at the beginning of Section 4, our objective is to find families of operators from $\ell_1(T)$ into $\ell_q(T)$ approximating W in a pointwise sense and such that we are able to control their entropy numbers. We are going to construct those families now.

Let $(\mathcal{B}_m)_{m\ge 0}$ be a sequence of tree partitions with root sets $(R_m)_{m\ge 0}$ as in Proposition 4.2. Fix $n \ge 1$ and, given $\mu \in \ell_1(T)$ with $\|\mu\|_1 \le 1$, define the tree partition $\mathcal{L}_\mu$ and the subtree $\mathcal{B}_\mu^\bullet$ as in (4.16) and (4.15), respectively. Set

$$\mathbb{L}_n := \{\mathcal{L}_\mu : \mu \in \ell_1(T),\ \|\mu\|_1 \le 1\}$$

(recall that n plays an important role in the construction of heavy and light domains, thus the $\mathcal{L}_\mu$ really depend on n) and observe that by Lemma 4.6,

$$\#\mathbb{L}_n \le (8e)^n.$$

Fix $\mathcal{L} \in \mathbb{L}_n$. We are going to define elements $(r_L^\circ)_{L\in\mathcal{L}}$, $(r_L^-)_{L\in\mathcal{L}}$ and $(r_L^\bullet)_{L\in\mathcal{L}}$ as follows: Take $L \in \mathcal{L}$ which may be represented as $L = B_{r,m}$ for some $m \ge 1$ and $r \in R_m$. Then we set

(1) $r_L^\circ := r$, i.e., r_L° is the root of the tree L.
(2) If $r_L^\circ \ne \mathbf{0}$, then by r_L^- we denote the parent element of r_L°.
(3) Finally, let $B_{r',m-1}$ be the parent element (in $\mathcal{B}_\infty$) of $L = B_{r,m}$. Then we put $r_L^\bullet := r'$.

Two cases may appear.

Generic case. A set $L \in \mathcal{L}$ is called *generic* provided that $r_L^\bullet \prec r_L^\circ$. Note that then even $r_L^\bullet \preceq r_L^- \prec r_L^\circ$.

Degenerate case. A set $L \in \mathcal{L}$ is called *degenerate* if $r_L^\bullet = r_L^\circ$, i.e., L and its parent element (in $\mathcal{B}_\infty$) have the same root.

Fix $\mathcal{L} \in \mathbb{L}_n$. We now define four operators $W_\mathcal{L}^1, \dots, W_\mathcal{L}^4$ depending on $\mathcal{L}$ and acting from $\ell_1(T)$ to $\ell_q(T)$ so that

$$W = \sum_{i=1}^{4} W_\mathcal{L}^i. \tag{4.22}$$

Given $s \in T$ we denote by $\delta_s \in \ell_1(T)$ the unit vector at s, i.e., $\delta_s(t) = 0$ if $t \neq s$ and $\delta_s(s) = 1$. Then the operator W defined in (3.2) is completely described by

$$(W\delta_s)(t) = \sigma(s)\alpha(t)\, \mathbf{1}_{\{t \equiv s\}}\, \mathbf{1}_{[\mathbf{0},s]}(t), \qquad s,t \in T. \tag{4.23}$$

The representation of W as a sum is related to a splitting of the branch $[\mathbf{0}, s]$ (and of the corresponding indicator $\mathbf{1}_{[0,s]}$ which appears in (4.23)) into four pieces as described below.

For each $s \in T$ choose the unique element $L \in \mathcal{L}$ such that $s \in L$. If this light set L with $s \in L$ is generic, then we split $[\mathbf{0}, s]$ as follows:

$$[\mathbf{0}, s] = [\mathbf{0}, r_L^\bullet] \cup (r_L^\bullet, r_L^-] \cup \{r_L^\circ\} \cup (r_L^\circ, s]. \tag{4.24}$$

Accordingly, we let

$$(W_\mathcal{L}^1\, \delta_s)(t) := \sigma(s)\alpha(t)\, \mathbf{1}_{\{t \equiv s\}}\, \mathbf{1}_{[\mathbf{0},r_L^\bullet]}(t), \qquad t \in T; \tag{4.25}$$

$$(W_\mathcal{L}^2\, \delta_s)(t) := \sigma(s)\alpha(t)\, \mathbf{1}_{\{t \equiv s\}}\, \mathbf{1}_{(r_L^\bullet,r_L^-]}(t), \qquad t \in T; \tag{4.26}$$

$$(W_\mathcal{L}^3\, \delta_s)(t) := \sigma(s)\alpha(t)\, \mathbf{1}_{\{t \equiv s\}}\, \mathbf{1}_{\{r_L^\circ\}}(t), \qquad t \in T; \tag{4.27}$$

$$(W_\mathcal{L}^4\, \delta_s)(t) := \sigma(s)\alpha(t)\, \mathbf{1}_{\{t \equiv s\}}\, \mathbf{1}_{(r_L^\circ,s]}(t), \qquad t \in T.$$

But if $L \in \mathcal{L}$ with $s \in L$ is degenerate, i.e., $r_L^\bullet = r_L^\circ$, then we simply have

$$[\mathbf{0}, s] = [\mathbf{0}, r_L^\circ] \cup (r_L^\circ, s] = [\mathbf{0}, r_L^\bullet] \cup (r_L^\circ, s].$$

Accordingly, we define $W_\mathcal{L}^1\delta_s$ and $W_\mathcal{L}^4\delta_s$ as in the generic case, while now $W_\mathcal{L}^2\delta_s = W_\mathcal{L}^3\delta_s = 0$. The representation (4.22) is straightforward.

Setting

$$W_\mathcal{L} := \sum_{i=1}^{3} W_\mathcal{L}^i$$

we get $W - W_\mathcal{L} = W_\mathcal{L}^4$, hence in view of $\log(\#\mathbb{L}_n) \leq cn$ it suffices to prove that for $\mu \in \ell_1(T)$ with $\|\mu\|_1 \leq 1$ always

$$\inf_{\mathcal{L} \in \mathbb{L}_n} \|W_\mathcal{L}^4 \mu\|_q \leq cn^{-(1-1/q)} \tag{4.28}$$

as well as

$$\sup_{\mathcal{L}\in\mathbb{L}_n} e_{[\rho n]}(W_{\mathcal{L}}) \le c n^{-(1-1/q)} \tag{4.29}$$

for a certain $\rho \ge 1$.

We begin by proving the first assertion.

Proposition 4.7. *There is a $c = c(q)$ such that* (4.28) *holds for each $\mu \in \ell_1(T)$ with $\|\mu\|_1 \le 1$.*

Proof. In a first step we estimate $\|W^4_{\mathcal{L}}\mu\|_q$ for *arbitrary* $\mathcal{L} \in \mathbb{L}_n$ and $\mu \in \ell_1(T)$. Here we have

$$\|W^4_{\mathcal{L}}\mu\|_q^q = \Big\|\sum_{s\in T}\mu(s)W^4_{\mathcal{L}}\delta_s\Big\|_q^q = \Big\|\sum_{L\in\mathcal{L}}\sum_{s\in L}\mu(s)W^4_{\mathcal{L}}\delta_s\Big\|_q^q.$$

Notice that the interior sums represent elements of $\ell_q(T)$ with disjoint supports (each sum is supported by the corresponding domain L). Hence, using the definitions of $W^4_{\mathcal{L}}$ and of $d_{\mathcal{I}}$ it follows that

$$\begin{aligned}\|W^4_{\mathcal{L}}\mu\|_q^q &= \sum_{L\in\mathcal{L}}\Big\|\sum_{s\in L}\mu(s)W^4_{\mathcal{L}}\delta_s\Big\|_q^q \le \sum_{L\in\mathcal{L}}\Big[\sum_{s\in L}|\mu(s)|\ \|W^4_{\mathcal{L}}\delta_s\|_q\Big]^q \\ &= \sum_{L\in\mathcal{L}}\Big[\sum_{s\in L}|\mu(s)|\ d_{\mathcal{I}}(r^\circ_L, s)\Big]^q \le \sum_{L\in\mathcal{L}}|\mu|(L)^q \sup_{s\in L} d_{\mathcal{I}}(r^\circ_L, s)^q \\ &\le \sum_{L\in\mathcal{L}}|\mu|(L)^q\varepsilon^q_{|L|}\end{aligned} \tag{4.30}$$

where in the last step we have used (4.10).

Estimate (4.31) holds for any $\mathcal{L} \in \mathbb{L}_n$ and $\mu \in \ell_1(T)$. Next, given $\mu \in \ell_1(T)$ with $\|\mu\|_1 \le 1$, we specify $\mathcal{L}$ by taking $\mathcal{L} = \mathcal{L}_\mu$ for the given μ. By the construction of $\mathcal{L}_\mu$ this yields $|\mu|(L) \le |L|/n$ for each $L \in \mathcal{L}_\mu$ (cf. (4.14) and (4.16)). Then (4.31) can be further estimated as follows:

$$\begin{aligned}\|W^4_{\mathcal{L}_\mu}\mu\|_q^q &\le \sum_{L\in\mathcal{L}_\mu}|\mu|(L)^q\varepsilon^q_{|L|} = \sum_{L\in\mathcal{L}_\mu}[|\mu|(L)\cdot|\mu|(L)^{q-1}\varepsilon^q_{|L|}] \\ &\le \sum_{L\in\mathcal{L}_\mu}|\mu|(L)\cdot\sup_{L\in\mathcal{L}_\mu}|\mu|(L)^{q-1}\varepsilon^q_{|L|} \\ &\le \|\mu\|_1\cdot\sup_{L\in\mathcal{L}_\mu}(|L|/n)^{q-1}(\log 2\,|L|)^{-(q-1)} \\ &\le cn^{-(q-1)}.\end{aligned}$$

Thus our calculations result in

$$\|W^4_{\mathcal{L}_\mu}\mu\|_q \le cn^{-(1-1/q)},$$

which, of course, implies (4.28) and completes the proof. ∎

Our next objective is to verify (4.29). Recall that $W_{\mathcal{L}} = W^1_{\mathcal{L}} + W^2_{\mathcal{L}} + W^3_{\mathcal{L}}$ with $W^i_{\mathcal{L}}$, $i = 1, 2, 3$, defined in (4.25), (4.26) and (4.27), respectively. By the additivity of the entropy numbers this implies

$$e_{3n-2}(W_{\mathcal{L}}) \le e_n(W^1_{\mathcal{L}}) + e_n(W^2_{\mathcal{L}}) + e_n(W^3_{\mathcal{L}}).$$

Thus, if we are able to verify $e_n(W^i_{\mathcal{L}}) \le c_i n^{-(1-1/q)}$ for $i = 1, 2, 3$, then this leads to

$$e_{3n-2}(W_{\mathcal{L}}) \le c n^{-(1-1/q)},$$

hence (4.29) is valid with $\rho = 3$. Consequently, it suffices to estimate $e_n(W^i_{\mathcal{L}})$ for $i = 1, 2, 3$ separately. We start by estimating $e_n(W^1_{\mathcal{L}})$.

PROPOSITION 4.8. *There is a constant $c = c(q)$ such that*

$$e_n(W^1_{\mathcal{L}}) \le c n^{-(1-1/q)}.$$

Proof. For $s \in T$ let L be the unique domain in $\mathcal{L}$ with $s \in L$. Clearly, $r^\bullet_L \preceq s$, hence if $s \not\equiv r^\bullet_L$, then $W^1_{\mathcal{L}}\delta_s = 0$. Thus it suffices to treat the case $s \equiv r^\bullet_L$ and then

$$(W^1_{\mathcal{L}}\delta_s)(t) = \alpha(t)\sigma(r^\bullet_L)\,\mathbf{1}_{\{t \preceq r^\bullet_L,\, t \equiv r^\bullet_L\}}. \tag{4.31}$$

Let $\Upsilon^\bullet_{\mathcal{L}} := \{r^\bullet_L : L \in \mathcal{L}\}$ and define an operator $V_{\mathcal{L}}$ from $\ell_1(\Upsilon^\bullet_{\mathcal{L}})$ into $\ell_q(T)$ by

$$(V_{\mathcal{L}}\delta_{r^\bullet_L})(t) := \alpha(t)\sigma(r^\bullet_L)\,\mathbf{1}_{\{t \preceq r^\bullet_L,\, t \equiv r^\bullet_L\}}. \tag{4.32}$$

Then, if U_1 is the unit ball in $\ell_1(T)$, by (4.31) and (4.32) it follows that $W^1_{\mathcal{L}}(U_1) = V_{\mathcal{L}}(U_1)$, hence $e_n(W^1_{\mathcal{L}}) = e_n(V_{\mathcal{L}})$. In order to estimate the latter entropy numbers we will use the following convenient result from [Ca, Proposition 1]. It provides a control of the entropy numbers for operators from ℓ_1-spaces into those of type q, based on the dimension of the first space. We refer to [MP] or [Pi] for the definition of type q.

PROPOSITION 4.9. *Let V be an operator from ℓ_1^N into a Banach space X of type q. Then for all $n = 1, 2, \ldots$,*

$$e_n(V) \le c(X) f(n, N, q)\|V\|$$

where the constant $c(X)$ depends only on the type q constant of the space X and

$$f(n, N, q) := 2^{-\max(n/N;1)} \min\left\{1; \left[\max\left(\frac{\log(N/n+1)}{n}; \frac{1}{N}\right)\right]^{1-1/q}\right\}.$$

Suppose now $n \ge N$. Then

$$f(n, N, q) = 2^{-n/N} N^{-(1-1/q)} \le c(q) n^{-(1-1/q)}$$

and we arrive at

$$e_n(V) \le c(X, q) n^{-(1-1/q)} \tag{4.33}$$

whenever $n \ge N$. Here $c(X,q)$ only depends on q and the type q constant of X.

In our case the operator $V := V_{\mathcal{L}}$ is defined on $\ell_1^{N_{\mathcal{L}}} = \ell_1(\Upsilon_{\mathcal{L}}^{\bullet})$ with $N_{\mathcal{L}} := \#\Upsilon_{\mathcal{L}}^{\bullet}$ and it acts into the space $\ell_q(T)$, which (cf. [MP]) for $1 < q \le 2$ is of type q with type q constant bounded by $\sqrt{q}$. Therefore, the important dimension parameter is $N_{\mathcal{L}}$. Here by (4.17) we have

$$N_{\mathcal{L}} = \#\Upsilon_{\mathcal{L}}^{\bullet} \le \#\mathcal{B}_{\mu}^{\bullet} \le n \tag{4.34}$$

where $\mu \in \ell_1(T)$ and $\mathcal{L} \in \mathbb{L}_n$ are related via $\mathcal{L} = \mathcal{L}_\mu$. Thus (4.33) applies and leads to

$$e_n(W_{\mathcal{L}}^1) = e_n(V_{\mathcal{L}}) \le c\|V_{\mathcal{L}}\|\; n^{-(1-1/q)} \le c\|W\| n^{-(1-1/q)}$$

with c only depending on q. In view of Proposition 3.1, this completes the proof of Proposition 4.8 since $\|W\| \le 2\|V_{\alpha,\sigma}\| \le 2\kappa$. ∎

Our next objective is to estimate $e_n(W_{\mathcal{L}}^2)$. Here Proposition 4.3 will play an important role.

Proposition 4.10. *There is a constant $c = c(q)$ such that*

$$e_n(W_{\mathcal{L}}^2) \le c n^{-(1-1/q)}. \tag{4.35}$$

Proof. Take $s \in T$ and choose as before the corresponding $L \in \mathcal{L}$ with $s \in L$. In the case where L is degenerate we have $W_{\mathcal{L}}^2 \delta_s = 0$, thus it suffices to investigate those $s \in T$ for which the corresponding L is generic, i.e., we have $r_L^{\bullet} \preceq r_L^{-} \prec r_L^{\circ}$. Furthermore, whenever $s \not\equiv r_L^-$, then $W_{\mathcal{L}}^2 \delta_s = 0$ as well. On the other hand, if $s \equiv r_L^-$, then $\sigma(s) = \sigma(r_L^-) = \sigma(r_L^{\circ})$ and

$$\mathbf{1}_{\{r_L^{\bullet} \prec t \preceq r_L^-,\, t \equiv s\}} = \mathbf{1}_{\{r_L^{\bullet} \prec t \preceq r_L^-,\, t \equiv r_L^{\circ}\}}.$$

For generic $L \in \mathcal{L}$ we define elements $x_L \in \ell_q(T)$ by

$$x_L(t) := \alpha(t)\sigma(r_L^{\circ})\, \mathbf{1}_{\{r_L^{\bullet} \prec t \preceq r_L^-,\, t \equiv r_L^{\circ}\}}$$

and a set $C_{\mathcal{L}} \subseteq \ell_q(T)$ by

$$C_{\mathcal{L}} := \{x_L : L \in \mathcal{L} \text{ is generic}\}.$$

Then

$$e_n(W_{\mathcal{L}}^2) = e_n(\mathrm{aco}(C_{\mathcal{L}}))$$

where $\mathrm{aco}(C_{\mathcal{L}})$ denotes the absolutely convex hull of $C_{\mathcal{L}} \subseteq \ell_q(T)$.

Take a generic $L \in \mathcal{L}$. Then there is an $m \ge 1$ such that $L = B_{\tau,m}$ with $\tau := r_L^{\circ} \in R_m$. Set $r := r_L^{\bullet}$. Then $r \in R_{m-1}$, $r \prec \tau$, $\tau \in R_m \setminus R_{m-1}$ and $\tau \in B_{r,m-1}$. Thus we are exactly in the situation of Proposition 4.3 with $\tau = r_L^{\circ}$ and $r = r_L^{\bullet}$, which implies

$$\|x_L\|_q = \sigma(r_L^{\circ}) \Big(\sum_{\substack{r_L^{\bullet} \prec v \preceq r_L^- \\ v \equiv r_L^{\circ}}} \alpha(v)^q \Big)^{1/q} \le c m^{-1} = c|L|^{-1}. \tag{4.36}$$

Hence, for any $h > 0$,

$$\#\{L \in \mathcal{L} : \|x_L\|_q \geq h\} \leq \#\{L \in \mathcal{L} : |L| \leq c/h\} \leq 2^{c/h+2} \tag{4.37}$$

where we have used $\#\{L \in \mathcal{L} : |L| = m\} \leq \#R_m \leq 2^{m+1}$ in the last estimate.

By [CKP, Proposition 6.2], which handles the entropy of convex hulls in type q spaces in the non-critical case, estimate (4.37) yields

$$e_k(\mathrm{aco}(C_{\mathcal{L}})) \leq ck^{-(1-1/q)}(\log k)^{-1/q}, \quad k \geq 1.$$

For $k = n$ we have

$$e_n(W^2_{\mathcal{L}}) = e_n(\mathrm{aco}(C_{\mathcal{L}})) \leq cn^{-(1-1/q)}(\log n)^{-1/q}.$$

This is even slightly better than required in (4.35) and completes the proof. ∎

Our final objective is to estimate $e_n(W^3_{\mathcal{L}})$ suitably.

PROPOSITION 4.11. *There is a constant $c = c(q)$ such that*

$$e_n(W^3_{\mathcal{L}}) \leq cn^{-(1-1/q)}.$$

Proof. Take $s \in T$ and $L \in \mathcal{L}$ with $s \in L$. If L is degenerate, then $W^3_{\mathcal{L}}\delta_s = 0$. This is so too if $s \not\equiv r^\circ_L$. Consequently,

$$\begin{aligned}&\{W^3_{\mathcal{L}}\delta_s : s \in T\}\\ &\quad = \{\sigma(s)\alpha(r^\circ_L)\delta_{r^\circ_L} : s \equiv r^\circ_L,\ L \text{ with } s \in L \text{ is generic}, s \in T\} \cup \{0\}\\ &\quad = \{\sigma(r^\circ_L)\alpha(r^\circ_L)\delta_{r^\circ_L} : L \text{ is generic}\} \cup \{0\}.\end{aligned}$$

Set

$$G_{\mathcal{L}} := \{r^\circ_L : L \text{ is generic}\}$$

and define a diagonal operator $D^3_{\mathcal{L}} : \ell_1(G_{\mathcal{L}}) \to \ell_q(G_{\mathcal{L}})$ by

$$D^3_{\mathcal{L}}(\delta_{r^\circ_L}) := \gamma_L\, \delta_{r^\circ_L}, \quad L \text{ generic},$$

where $\gamma_L := \sigma(r^\circ_L)\alpha(r^\circ_L)$. Then $e_n(W^3_L) = e_n(D^3_{\mathcal{L}})$ and it suffices to estimate the γ_L suitably.

Recall that r°_L and $r^\bullet_L$ belong to the same element of the partition $\mathcal{B}_{|L|-1}$, hence, if L is generic, i.e., if $r^\bullet_L \prec r^\circ_L$, by (4.10) we obtain

$$\gamma_L := \sigma(r^\circ_L)\alpha(r^\circ_L) \leq d_{\mathcal{I}}(r^\bullet_L, r^\circ_L) \leq \varepsilon_{|L|-1}.$$

It follows that

$$\#\{L : \gamma_L \geq \varepsilon_m\} \leq \#\{L : |L| \leq m+1\} \leq 2^{m+3}.$$

Again we have used $\#\{L \in \mathcal{L} : |L| = m\} \leq \#R_m \leq 2^{m+1}$ in the last step. If $\{\gamma^*_k\}_{k\geq 1}$ is the non-increasing rearrangement of $\{\gamma_L\}_{L\in\mathcal{L}}$, we have

$$\gamma^*_k \leq c(\log k)^{-(1-1/q)}.$$

By using [Ku, Proposition 3.1] where the entropy of critical diagonal operators with logarithmic diagonal is handled, we obtain

$$e_k(W^3_{\mathcal{L}}) = e_k(D^3_{\mathcal{L}}) \le ck^{-(1-1/q)}, \quad k \ge 1.$$

For $k = n$ we have

$$e_n(W^3_{\mathcal{L}}) = e_n(D^3_{\mathcal{L}}) \le cn^{-(1-1/q)}$$

as asserted. ∎

5. Final remarks. We must acknowledge that the proof of Theorem 1.1, or Theorem 3.2, is quite complicated. One of the reasons for this is that so many operators $W^i_{\mathcal{L}}$ are involved. Thus a natural question is why two or three operators do not suffice. Indeed, once a very natural bound for $W^4_{\mathcal{L}}$ (Proposition 4.7) is obtained, it is tempting to use (for the generic case) a splitting into two pieces instead of four as in (4.24), i.e., to split $[\mathbf{0}, s]$ only as

$$[\mathbf{0}, s] = [\mathbf{0}, r^\circ_L] \cup (r^\circ_L, r].$$

In other words, why cannot we add up $W^1_{\mathcal{L}}, W^2_{\mathcal{L}}, W^3_{\mathcal{L}}$ into one operator and deal with it as we did with $W^1_{\mathcal{L}}$? In fact, the corresponding bound is dimension-based. Therefore, proceeding in this way, we must replace the dimension bound (4.34) with some bound for $\#\Upsilon^\circ_{\mathcal{L}}$ where $\Upsilon^\circ_{\mathcal{L}} := \{r^\circ_L : L \in \mathcal{L}\}$. Unfortunately, $\#\Upsilon^\circ_{\mathcal{L}} = \#\mathcal{L}$, the number of *extremal light* domains, does not admit any uniform estimate, unlike the number of *heavy* domains we used in the proof. The only chance to estimate $\#\Upsilon^\circ_{\mathcal{L}}$ is to make further assumptions about the structure of the underlying tree T. Therefore, the splitting into two pieces does not work for general trees.

Once this difficulty is understood, the next natural idea is to use a splitting into three pieces,

$$[\mathbf{0}, s] = [\mathbf{0}, r^\bullet_L] \cup (r^\bullet_L, r^\circ_L] \cup (r^\circ_L, r].$$

In other words, why cannot we add up $W^2_{\mathcal{L}}$ and $W^3_{\mathcal{L}}$ into one operator and deal with it as we did with $W^2_{\mathcal{L}}$? Recall that the corresponding bound from Proposition 4.10 is based on the size evaluation $\|x_L\|_q \le c\,|L|^{-1}$ from (4.36). That one in turn was built upon the tricky property (4) from Proposition 4.2. Once we use the three-piece splitting, we can only use (4.10) for the evaluation of $\|x_L\|$, as we did when working with $W^3_{\mathcal{L}}$. In this way we only obtain $\|x_L\| \le \varepsilon_{|L|-1} = c|L|^{-(1-1/q)}$. Unfortunately, we *do not know* whether or not this weaker bound provides the necessary bound $cn^{-(1-1/q)}$ for the entropy numbers of the convex hull of a sequence. To the best of our knowledge, the required result is missing in the literature for subsets of spaces of type q (or even for subsets of ℓ_q-spaces). Thus it is this gap that forced us to struggle with partition constructions having property (4), and then extract the well studied diagonal operators $W^3_{\mathcal{L}}$ or $D^3_{\mathcal{L}}$ by the further splitting $(r^\bullet_L, r^\circ_L] = (r^\bullet_L, r^-_L] \cup \{r^\circ_L\}$, eventually coming to the proof presented here. Let us mention that for $q = 2$, i.e., for sets in Hilbert spaces, such entropy estimates for convex hulls of sequences are known (cf. Proposition 4 in [CE]). Hence, if $q = 2$, we may add up $W^2_{\mathcal{L}}$ and $W^3_{\mathcal{L}}$ into one operator, which slightly simplifies the proof in that case because here we need neither property (4) of Proposition 4.2 nor Proposition 4.3.

Another difficulty comes from the partition $\mathcal{I}$ of T generated by the weight σ. This forced us to replace the metric d on T by the localized "distance" $d_{\mathcal{I}}$. Of course, this additional difficulty does not appear for one-weight summation operators $V_{\alpha,\sigma}$ with $\sigma(t) = 1$, $t \in T$. Hence, also in that case the proof of Theorem 1.1 becomes slightly less involved.

Acknowledgements. We are grateful to M. Lacey for his interest and useful discussions that led us to the present research. Furthermore, the research was supported by the RFBR-DFG grant 09-01-91331 "Geometry and asymptotics of random structures". The work of the first named author was also supported by RFBR grants 11-01-12104-ofi-m and 10-01-00154a, as well as by Federal Focused Programme 2010-1.1-111-128-033.

References

[Ca] B. Carl, *Inequalities of Bernstein–Jackson-type and the degree of compactness of operators in Banach spaces*, Ann. Inst. Fourier (Grenoble) 35 (1985), no. 3, 79–118.

[CE] B. Carl and D. E. Edmunds, *Gelfand numbers and metric entropy of convex hulls in Hilbert spaces*, Studia Math. 159 (2003), 391–402.

[CKP] B. Carl, I. Kyrezi and A. Pajor, *Metric entropy of convex hulls in Banach spaces*, J. London Math. Soc. 60 (1999), 871–896.

[CS] B. Carl and I. Stephani, *Entropy, Compactness and Approximation of Operators*, Cambridge Univ. Press, Cambridge, 1990.

[CSt] J. Creutzig and I. Steinwart, *Metric entropy of convex hulls in type p spaces—the critical case*, Proc. Amer. Math. Soc. 130 (2002), 733–743.

[ET] D. E. Edmunds and H. Triebel, *Function Spaces, Entropy Numbers and Differential Operators*, Cambridge Univ. Press, Cambridge, 1996.

[Ga] F. Gao, *Metric entropy of convex hulls*, Israel J. Math. 123 (2001), 359–364.

[Ku] Th. Kühn, *Entropy numbers of general diagonal operators*, Rev. Mat. Complut. 18 (2005), 479–491.

[Lif] M. A. Lifshits, *Bounds for entropy numbers for some critical operators*, Trans. Amer. Math. Soc., to appear; www.arxiv.org/abs/1002.1377.

[LL1] M. A. Lifshits and W. Linde, *Compactness properties of weighted summation operators on trees*, Studia Math. 202 (2011), 17–47.

[LL2] —, —, *Random Gaussian sums on trees*, Electron. J. Probab. 16 (2011), 739–763.

[Lin] W. Linde, *Nondeterminism of linear operators and lower entropy estimates*, J. Fourier Anal. Appl. 14 (2008), 568–587.

[MP] B. Maurey et G. Pisier, *Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach*, Studia Math. 58 (1976), 45–90.
[Pi] G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry*, Cambridge Univ. Press, Cambridge, 1989.

Mikhail Lifshits
Department of Mathematics and Mechanics
St. Petersburg State University
Bibliotechnaya pl. 2
198504 Stary Peterhof, Russia
E-mail: lifts@mail.rcom.ru

Werner Linde
Institut für Stochastik
Friedrich-Schiller-Universität Jena
Ernst-Abbe-Platz 2
07743 Jena, Germany
E-mail: werner.linde@uni-jena.de

Received August 2, 2011 (7265)