BOUNDS FOR ENTROPY NUMBERS FOR SOME CRITICAL OPERATORS

M. A. LIFSHITS

Abstract. We provide upper bounds for entropy numbers for two types of operators: summation operators on binary trees and integral operators of Volterra type. Our efforts are concentrated on the critical cases where none of the known methods work. Therefore, we develop a method which seems to be completely new and probably merits further applications.

INTRODUCTION

We will investigate the entropy numbers of certain linear operators. Recall that for a set \( A \) in a metric space, its \( n \)-th diadic entropy number is the infimum of \( \varepsilon > 0 \) such that \( A \) admits a covering by \( 2^{n-1} \) balls of radius \( \varepsilon \). Moreover, given a compact linear operator \( V : X \to Y \) acting from one normed space to another, its entropy numbers \( e_n(V) \) are defined as those of the \( V \)-image of the unit ball of \( X \). The entropy numbers along with other measures of compactness such as approximation numbers and Kolmogorov numbers play an extremely important role in operator theory and its applications. We refer to classical monographs [6] and [8] for further details and references.

This work originates from a question of M. Lacey and W. Linde. They investigated entropy numbers for linear Volterra operators with relatively bad compactness properties and discovered that two types of the behavior of entropy numbers are possible [10], [12] (see more details in Section 3). On a certain boundary separating the two cases their methods did not apply and the problem remained open. Further hard efforts convinced us that the remaining case cannot be settled by a rich variety of traditional methods. Therefore, a new technique is required. It turned out that this new technique could be cleanly elaborated and better explained if we replace the Volterra operator by an analogous summation operator on the binary tree. This class of operators is quite simple and natural, but it is absolutely not investigated (its properties will be a subject of a separate work). Therefore, we start with consideration of summation operators and first prove our estimate in this case. Notice that the trees appear naturally in the study of functional spaces because the Haar base and other similar wavelet bases indeed have a structure close to that of a binary tree.

In the last section, we reproduce the same approach for the integral operator considered by Lacey and Linde.
Here is our main point: in most parts of classical methods, in order to evaluate entropy numbers $e_n(V)$ of an operator $V$, one approximates $V$ with a finite rank operator depending on $n$. Contrary to this, we approximate $V$ with a family of finite rank operators indexed by some finite set of “essential trees”, a notion introduced in this article.

1. Introduction to tree summation operators

We consider a tree $T$ and its levels $\{T_l\}$, $l = 0, 1, \ldots$, such that the level $T_0$ consists of the single node (the tree root) and the level $T_{l+1}$ is the set of all direct offsprings of nodes that belong to $T_l$.

We denote $O_T(t) = O(t)$ as the set of all direct and indirect offsprings of a node $t \in T$ including $t$ itself and let $O_l(t) = O(t) \cap T_l$. If $u \in O(t)$, we write $u \succeq t$ and $t \preceq u$. The strict inequalities have the same meaning with the additional assumption $u \neq t$.

For any element $\mu \in \ell_1(T)$ and any $t \in T$ we denote the mass and variation at $t$ as

$$s_\mu(t) = \sum_{u \preceq t} \mu(u), \quad ||\mu||(t) = \sum_{u \preceq t} |\mu(u)|.$$ 

Clearly, for any $t \in T$,

$$s_\mu(t) \leq ||\mu||(t),$$

and for any $t \in T_l$ and any $m \geq l$ we have

$$\sum_{u \in O_m(t)} ||\mu||(u) \leq ||\mu||(t).$$

Now assume that $T$ is equipped with a non-negative weight $W = \{w(t)\}_{t \in T}$.

The weight $W$ gives rise to the simple weighted summation operator, $\tilde{V} : \ell_2(T) \rightarrow \ell_\infty(T)$, given by

$$\tilde{V}_\mu(t) = \sum_{u \preceq t} \sqrt{w(u)} f(u), \quad t \in T,$$

where the summation is actually taken over the branch leading from the root to the node $t$. For technical reasons, we will investigate a slightly different form of this operator. Namely, let us introduce a pair of dual tree-summation operators, $V : \ell_2(T, W) \rightarrow \ell_\infty(T)$ and $V^* : \ell_1(T) \rightarrow \ell_2(T, W)$ defined by

$$Vf(t) = \sum_{u \preceq t} w(u) f(u), \quad t \in T,$$

and

$$(V^*\mu)(t) = s_\mu(t) = \mu(O(t)), \quad t \in T,$$

respectively. It is easy to see that

$$||V||^2 = ||V^*||^2 = \sup_{t \in T} \sum_{u \preceq t} w(u).$$

It is also clear that the operators $V$ and $\tilde{V}$ are isomorphic. We have chosen the representation (1.3) because of the simple form of the operator $V^*$, the one we will really handle.
2. The entropy of a summation operator on the binary tree

In this section we consider a binary tree $T$ with levels $\{T_l\}$, $l = 0, 1, \ldots$, such that the level $T_0$ consists of the single node (the tree root) and every node of level $T_l$ generates $2$ offsprings in $T_{l+1}$. Note that $|T_l| = 2^l$.

The weight $W = \{w(t)\}_{t \in T}$ is defined by

$$w(t) = (1 + |t|)^{-\beta}, \quad t \in T, \quad \beta > 1.$$  \hspace{1cm} (2.1)

2.1. Regular case.

**Theorem 2.1.** Let $\beta > 1$ and let the weight $W$ be given by (2.1). Consider the linear operator $V^*: \ell_1(T) \to \ell_2(T, W)$ defined by (1.4). There exist numeric constants $C_1, C_2$ depending on $\beta$ such that for all positive integers $n$ we have the following bounds for its entropy numbers:

$$\frac{C_1}{n^{\beta/2}} \leq \epsilon_n(V^*) \leq \frac{C_2}{n^{\beta/2}}, \quad 1 < \beta < 2,$$

$$\frac{C_1(\ln n)^{1-\beta/2}}{n^{1/2}} \leq \epsilon_n(V^*) \leq \frac{C_2(\ln n)^{1-\beta/2}}{n^{1/2}}, \quad \beta > 2,$$

$$\frac{C_1}{n^{1/2}} \leq \epsilon_n(V^*) \leq \frac{C_2 \ln n}{n^{1/2}}, \quad \beta = 2.$$  

**Proof. Upper bound.** Consider the set $D = \{V^* \delta_t, t \in T\} \subset \ell_2(W(T))$, where as usual $\delta_t$ denotes the delta-function at point $t$, i.e. $\delta_t(u) = 1_{u=t}$. Recall that

$$(V^* \delta_t)(u) = 1_{u=t}.$$  

It is rather easy to establish an upper bound for diadic entropy of $D$. Indeed, take a net $D_n = \{V^* \delta_t, t \in T, |t| \leq n\}$. Then $|D_n| \leq 2^{n+1}$ and, trivially, for any $t \in T$ we have

$$\text{dist}(V^* \delta_t, D_n)^2 = \begin{cases} 0, & |t| \leq n, \\ \sum_{|t|=n+1} w_t, & |t| > n. \end{cases}$$

We see that

$$\epsilon_{n+2}(D) \leq \sum_{|t| \leq n} w_t = \left[ \sum_{|t|=n+1} (1 + |t|)^{-\beta} \right]^{1/2} \leq C n^{-\beta/2}.$$  

Now recall that a polynomial upper bound $\epsilon_n(D) \leq C n^{-\alpha}$ for any set $D$ in a Hilbert space yields a bound on $\epsilon_n(aco D)$, where $aco D$ denotes the absolutely convex hull of $D$. Namely, as established in [4] for $\alpha \neq 1/2$ and in [5] for $\alpha = 1/2$, under this assumption we have

$$\epsilon_n(aco D) \leq C n^{-\alpha}, \quad \alpha < 1/2,$$

$$\epsilon_n(aco D) \leq C n^{-1/2} \ln n, \quad \alpha = 1/2,$$

$$\epsilon_n(aco D) \leq C n^{-1/2} (\ln n)^{1/2-\alpha}, \quad \alpha > 1/2.$$  

By letting $\alpha = \frac{\beta-1}{2}$ here, we obtain the desired upper bounds in Theorem 2.1 because by the property of the unit ball in the $\ell_1$-space we have $\epsilon_n(V^*) = \epsilon_n(aco D)$.

**Lower bound.** For any $n \in \mathbb{N}$ let $m = 2^n$ and denote

$$\{t : |t| = n\} := (t_j)_{1 \leq j \leq m}.$$  

Take any \((s_j)_{1 \leq j \leq m}\) such that \(|s_j| = 2n\) and \(s_j\) is an (indirect) offspring of \(t_j\). Let \(\mu_j = \delta_{s_j} - \delta_{t_j}\). Then
\[
(V^* \mu_j)(u) = 1_{\{t_j \prec u \preceq s_j\}}.
\]
These image vectors are orthogonal, since they have disjoint supports, and for appropriate \(C_1\)
\[
||V^* \mu_j||^2_{2,W} = \sum_{l=n+1}^{2n} w_l \geq C_1^2 n^{-(\beta - 1)}.
\]
We notice that we found \(m = 2^n\) elements \(\mu_j\) such that \(||\mu_j||_1 = 2\), and for \(i \neq j\) we have \(||V^*(\mu_j - \mu_i)||_{2,W} \geq C_1 n^{-\frac{\beta - 1}{2}}\). It follows that
\[
e_{n+1}(V^*) \geq \frac{C_1}{2} n^{-\frac{\beta - 1}{2}}.
\]
This is true for any \(\beta > 1\) but it is optimal only for \(1 < \beta \leq 2\), while for \(\beta > 2\) we need a refined argument.

By using the same vectors, we see that the restriction of \(V^*\) on the span of vectors \((\mu_j)\) is isometric to the embedding \(I_m : \ell^m_1 \to \ell^m_\infty\) up to the coefficient
\[
\frac{||V^* \mu_j||_{2,W}}{||\mu_j||_1} \geq \frac{C_1}{2} n^{-\frac{\beta - 1}{2}}.
\]
Recall that with an appropriate numerical \(c > 0\) we have a (sharp) estimate
\[
e_k(I_m) \geq \left[\frac{c \ln(1+m/k)}{k}\right]^{1/2}, \quad \log_2 m \leq k \leq m;
\]
see [13]. Choose \(n = n(k)\) such that \(2^{n/2} \leq k \leq 2^{(n+1)/2}\). Since \(m = m(n) = 2^n\), the parameter \(k\) fits in the range and we obtain
\[
e_k(V^*) \geq \frac{C_1}{2} n^{-\frac{\beta - 1}{2}} e_k(I_m)
\geq \frac{C_1}{2} n^{-\frac{\beta - 1}{2}} \left[\frac{c \ln(1 + 2^{(n-1)/2})}{k}\right]^{1/2} \geq \frac{\tilde{C}_1}{k^{1/2}} (\ln k)^{1-\beta/2},
\]
which is sharp for \(\beta \geq 2\).

\[\square\]

**Remark 2.2.** There are many available proofs for upper bounds in Theorem 2.1. The one presented here is probably the shortest one. It is due to W. Linde.

**Remark 2.3.** We refer to [3] for the studies of other summation operators in the language of small deviation probabilities.

### 2.2. Critical case.
We see that in Theorem 2.1 the upper and lower estimates for \(\beta \neq 2\) are of the same order and are “easy” to obtain modulo known results, although at the point \(\beta = 2\) the behavior of entropy undergoes a striking change. Moreover, in the case \(\beta = 2\) the estimates of Theorem 2.1 do not fit together and leave a logarithmic gap.

Apparently this gap is impossible to close just by combining the existing results.

Therefore, we call the case \(\beta = 2\) a critical one. We will show that the lower bound of Theorem 2.1 is in fact sharp but the proof of the corresponding upper bound is by far more complicated and requires a new method.
Theorem 2.4. Let the weight \( W \) be given by (2.1) with \( \beta = 2 \). Consider the linear operator \( V : \ell_2(T, W) \to \ell_\infty(T) \) defined by (1.3) and its dual \( V^* : \ell_1(T) \to \ell_2(T, W) \) defined by (1.4). There exists a numeric constant \( C \) such that for all positive integers \( n \) we have the following upper bound for its entropy numbers:

\[
\max \{ e_n(V), e_n(V^*) \} \leq C \sqrt{n}.
\]

Proof. The proof of Theorem 2.4 will be split into a few steps, with each step having its own clear meaning. We keep the notation \( O(t), O_l(t), |t|, s_\mu(t), ||\mu||(t) \) from the previous section.

Step 1: Essential subtrees. A subset \( \Upsilon \subset T \) is called a subtree if \( t \prec u \in \Upsilon \) yields \( t \in \Upsilon \). In particular, the root is contained in any subtree.

The evaluation of entropy numbers will be based on the construction of some family of subtrees \( \Upsilon^\mu \) based on a stopping rule. Namely, let \( \sigma_1 = \frac{1}{n} \). For \( \mu \in \ell_1(T) \) satisfying \( ||\mu||_1 \leq 1 \) we define the \( n \)-essential subtree \( \Upsilon^\mu \) by starting from the root and including all nodes in \( \Upsilon^\mu \) while \( ||\mu||(t) > \sigma_{|t|} \) and stopping the construction while \( ||\mu||(t) \leq \sigma_{|t|} \). We denote by \( B^\mu \) the set of nodes where construction was stopped.

Since \( \sigma_{n+1} > 1 \), we have \( \Upsilon^\mu \cap T_{n+1} = \emptyset \), that is, we stop the construction no later than at the level \( n \). In particular, \( \Upsilon^\mu \) is finite. Notice that we have a partition

\[
(2.2) \quad T = \Upsilon^\mu \cup \left( \bigcup_{t \in B^\mu} O(t) \right).
\]

Now we evaluate the size of \( \Upsilon^\mu \) and will see that:

The size of \( \Upsilon^\mu \) is dramatically small.

This is a decisive step towards our goal. Let \( N_l = |\Upsilon^\mu \cap T_l| \). We have

Lemma 2.5. Let \( Q \) be the set of terminal nodes of \( \Upsilon^\mu \). It is true that

\[
(2.3) \quad \sum_{t \in Q} |t| \leq n.
\]

Moreover,

\[
(2.4) \quad \sum_{l=1}^{\infty} N_l \leq n + 1.
\]

Proof of Lemma 2.5. By the definition of an \( n \)-essential tree we have

\[
1 \geq \sum_{t \in Q} ||\mu||(t) \geq \sum_{t \in Q} \frac{|t|}{n}.
\]

It follows that \( \sum_{t \in Q} |t| \leq n \), as required in (2.3). On the other hand, for any tree \( \Upsilon \) and its terminal set \( Q \) it is true that

\[
\sum_{l} |\Upsilon_l| = 1 + \sum_{u \in \Upsilon, |u| > 0} 1 \leq 1 + \sum_{u \in \Upsilon, |u| > 0} |Q \cap O(u)|
\]

\[
= 1 + \sum_{t \in Q, u: |u| > 0, u \preceq t} 1 = 1 + \sum_{t \in Q} |t|;
\]

thus (2.4) follows. \( \square \)
Remark 2.6. Actually, we proved something stronger than (2.4); namely,
\[ \sum_{u \in \mathcal{T}_n, |u| > 0} |Q \cap \mathcal{O}(u)| \leq n. \]

We can also easily evaluate the number of possible $n$-essential trees.

Lemma 2.7. The number of subtrees of a binary tree whose terminal set $Q$ satisfies (2.3) does not exceed $(4e)^n$.

Proof of Lemma 2.7. Since a subtree is entirely defined by its terminal set, we have to find out how many sets $Q$ satisfy (2.3). Denote $q_l = |Q \cap T_l|$. Then (2.3) writes as
\[ \sum_l q_l \leq n. \]

Since $q_l \leq n$, the number of non-negative integer solutions of this inequality does not exceed
\[ \prod_{l=1}^n (1 + q_l) = \prod_{l=1}^n \left(1 + \frac{2n}{l}\right) \leq \left(\frac{2n}{n/e}\right)^n = (2e)^n. \]

Moreover, for a given sequence $q_l$, while constructing a set $Q$, on each level $l$ of a binary tree we have to choose $q_l$ elements from at most $2^{q_l}$ elements of this level. Therefore, the number of possible sets does not exceed
\[ \prod_{l=1}^n \left(2^{q_l} \right) \leq \prod_{l=1}^n (2^{q_l}) = 2^{\sum_{l=1}^n q_l} \leq 2^n. \]

We finish the discussion of essential subtrees by proving their useful approximation property. It follows from (1.1) and (1.2) that for any $t \in \mathcal{T}$ it is true that
\[ \sum_{|t| = l} \sum_{u \in \mathcal{O}_l(t)} w(u)s_\mu(u)^2 \leq \sum_{|t| = l} \max_{u \in \mathcal{O}_l(t)} |s_\mu(u)| \sum_{u \in \mathcal{O}_l(t)} w(u)|s_\mu(u)| \\
\leq \sum_{|t| = l} \max_{u \in \mathcal{O}_l(t)} ||\mu||(u) \sum_{u \in \mathcal{O}_l(t)} w(u)||\mu(u)|| \\
\leq ||\mu|||t|^2 \sum_{|t| = l} (1 + |l|)^{-2} \leq ||\mu|||t|^2|t|^{-1}. \]

Moreover, (2.5) and the definition of $B^\mu$ yield
\[ \sum_{t \in \mathcal{T}_n} s_\mu(u)^2w(u) = \sum_{t \in B^\mu} \sum_{u \in \mathcal{O}_l(t)} s_\mu(u)^2w(u) \\
\leq \sum_{t \in B^\mu} ||\mu|||t|^2|t|^{-1} \leq \sum_{t \in B^\mu} \sigma_t ||\mu|||t|^2|t|^{-1} \\
= \sum_{t \in B^\mu} \frac{|t|}{n} ||\mu|||t| |t|^{-1} \\
\leq \sum_{t \in B^\mu} ||\mu|||t| n^{-1} \leq n^{-1}. \]
Therefore, as we will soon see, the part of operator $V^\ast$ related to the complement of $\Upsilon^\mu$ is not essential at the precision level $n^{-1/2}$ which explains the name “essential” we gave to this family.

**Step 2: Approximating operators.** We are now going to construct a family of finite rank operators approximating the operator $V^\ast$. Each operator will correspond to an $n$-essential subtree. However the construction is valid for any subtree of $T$. Given a subtree $\Upsilon \subset T$ we define three operators related to $\Upsilon$. The operator $V^\ast_\Upsilon : \ell_1(\Upsilon) \to \ell_2(\Upsilon, W)$ is defined by
\[
(V^\ast_\Upsilon \mu)(t) = \sum_{u \in O(t) \cap \Upsilon} \mu(u), \quad t \in \Upsilon.
\]
This is essentially the same operator as $V^\ast$ restricted on elements supported by $\Upsilon$.

Now define a mapping $z$ from the complement of $\Upsilon$ to the boundary of $\Upsilon$ by letting $z(s)$ be the last node in $\Upsilon$ on the way from the root to $s$. We denote $Z(t) = z^{-1}(t)$. This set will be non-empty only if $t$ belongs to the boundary of $\Upsilon$.

Now the flush-projection operator $P_\Upsilon : \ell_1(\Upsilon) \to \ell_1(\Upsilon)$ is defined by
\[
(P_\Upsilon \mu)(t) = \mu(t) + \sum_{u \in Z(t)} \mu(u), \quad t \in \Upsilon.
\]
This operator projects measures supported by $T$ onto the measures supported by $\Upsilon$. It is clear that $\|P_\Upsilon\| \leq 1$, i.e. $P_\Upsilon$ is a contraction.

The main property of the operators introduced so far reads as
\[
(V^\ast_\Upsilon P_\Upsilon \mu)(t) = \sum_{u \in O(t)} \mu(u) = (V^\ast \mu)(t), \quad \forall t \in \Upsilon, \mu \in \ell_1(\Upsilon).
\]

Finally, we will use the natural embedding $\iota_\Upsilon : \ell_2(\Upsilon, W) \to \ell_2(T, W)$ defined by
\[
\iota_\Upsilon \mu(t) = \begin{cases} 
\mu(t), & t \in \Upsilon, \\
0, & t \notin \Upsilon.
\end{cases}
\]
Combining these all together we define the approximating operator $A_\Upsilon : \ell_1(\Upsilon) \to \ell_2(T, W)$ by $A_\Upsilon = \iota_\Upsilon V^\ast_\Upsilon P_\Upsilon$. It follows from (2.7) that for any subtree $\Upsilon$ and any $\mu \in \ell_1(\Upsilon)$ we have
\[
(V^\ast \mu)(t) - (A_\Upsilon \mu)(t) = \begin{cases} 
0, & t \in \Upsilon, \\
s_\mu(t), & t \notin \Upsilon.
\end{cases}
\]
Hence,
\[
||(V^\ast - A_\Upsilon)\mu||^2_{2,W} = \sum_{t \notin \Upsilon} |s_\mu(t)|^2 w(t).
\]

Finally notice that since $\|P_\Upsilon\| \leq 1$ and $\|\iota_\Upsilon\| \leq 1$, we have for any $m \in \mathbb{N}$ that
\[
\epsilon_m(A_\Upsilon) \leq \epsilon_m(V^\ast_\Upsilon).
\]

So far we have not specified our subtree. Now we will use the $n$-essential subtrees constructed above. For any given $n$ let $\Gamma = \{\Upsilon\}$ be the set of all subtrees $\Upsilon \subset T$ satisfying (2.3), hence (2.4). Recall that by Lemma 2.7 we have $|\Gamma| \leq (4e)^n$, and for any $\mu \in \ell_1(\Upsilon)$ its $n$-essential subtree $\Upsilon^\mu$ belongs to $\Gamma$.

By comparing inequality (2.6) with (2.8) we see that for any $\mu \in \ell_1(\Upsilon)$ with $||\mu||_1 \leq 1$ we have
\[
||(V^\ast - A_{\Upsilon^\mu})\mu||^2_{2,W} \leq n^{-1}.
\]
In other words,

\[(2.10) \quad \sup_{\mu: ||\mu||_1 \leq 1} \inf_{\gamma \in \Gamma} ||(V^* - A_{\gamma})\mu||_{2,W} \leq n^{-1/2}.\]

Recall that for every \(\mu\) its own approximating operator is used. We will now show how the properties like this one can be applied. This simple idea seems to be of independent interest, thus we state it as a separate statement.

**Step 3: Approximation lemma.** The following lemma shows how a linear operator \(V\) can be approximated by a family of operators \((V_\gamma)_{\gamma \in \Gamma}\) in the sense that for every element \(x\) its image \(Vx\) is approximated by \(V_\gamma x\) with an appropriate \(\gamma\) depending on \(x\).

**Lemma 2.8.** Let \(X, Y\) be the normed spaces and \(V, (V_\gamma)_{\gamma \in \Gamma}\) be the linear operators acting from \(X\) to \(Y\). Then for any \(n \in \mathbb{N}\) it is true that

\[(2.11) \quad e_{n+\lfloor \log_2 |\Gamma| \rfloor +1}(V) \leq \sup_{\gamma \in \Gamma} e_n(V_\gamma) + \sup_{x \in B_X} \inf_{\gamma \in \Gamma} ||Vx - V_\gamma x||_Y,
\]

where \(B_X = \{x \in X : ||x||_X \leq 1\}\).

**Proof.** Denote by \(S_1\) and \(S_2\) the expressions in (2.11) and fix a small \(\delta > 0\). For every \(\gamma\) we choose an \(\{e_n(V_\gamma) + \delta\}\)-net \(N_\gamma\) of size \(2^{n-1}\) for the set \(V_\gamma(B_X)\) in the space \(Y\). Let

\[N = \bigcup_{\gamma \in \Gamma} N_\gamma\]

be a global net. Clearly,

\[#\{N\} = |\Gamma| \cdot 2^{n-1} \leq 2^{\lfloor \log_2 |\Gamma| \rfloor + n}.
\]

For any \(x \in B_X\) we first find a \(\gamma\) such that

\[||Vx - V_\gamma x||_Y \leq S_2 + \delta.
\]

Then we find an element \(y \in N_\gamma \subset N\) such that

\[||V_\gamma x - y||_Y \leq e_n(V_\gamma) + \delta.
\]

By a triangle inequality, we have

\[||Vx - y||_Y \leq S_2 + \delta + e_n(V_\gamma) + \delta \leq S_1 + S_2 + 2\delta.
\]

Therefore, \(N\) is an \((S_1 + S_2 + 2\delta)\)-net for the set \(V(B_X)\) and its size does not exceed \(2^{\lfloor \log_2 |\Gamma| \rfloor + n}\). The assertion of the lemma follows. \(\square\)

We will apply Lemma 2.8 to \(X = \ell_1(T), B_X = \{\mu : ||\mu||_1 \leq 1\}, Y = \ell_2(T,W), V = V^*,\) and the approximating family of operators \((A_\gamma)_{\gamma \in \Gamma}\). Now (2.11), (2.10), and (2.9), together with the known estimate of \(|\Gamma|\) given in Lemma 2.7, yield

\[(2.12) \quad e_{(1+\lfloor \log_2 (4e) \rfloor)n+1}(V^*) \leq \sup_{\gamma \in \Gamma} e_n(V^*_\gamma) + n^{-1/2}.
\]
Step 4: Evaluation of operators on short trees. With \([2.12]\) at hand, it remains to evaluate \(e_n(\mathcal{V}_\mathcal{T}^+)\) for fixed \(\mathcal{T} \in \Gamma\). In other words, we have to evaluate the entropy of the operator restricted to a tree of a very small size (due to the bound \([2.4]\) for the size of \(\mathcal{T}\)).

Towards this aim, recall an important entropy bound from \([5, Corollary 2.4 (i)]\). There exists a constant \(c > 0\) such that for any operator \(W\) acting from \(\ell^m_1\) to a Hilbert space and any \(k \in \mathbb{N}\) it is true that

\[
e_k(W) \leq c \ln (m+1) \|W\| k^{-1/2},
\]

We will apply this estimate to a particular situation of tree operators. Let \(\Delta\) be the tree that consists of the first \(\lfloor \ln n/4 \rfloor\) levels of a binary tree.

Let us split our operator in a sum \(\mathcal{V}_\mathcal{T}^+ = \mathcal{V}_\mathcal{T}^+ + \mathcal{V}_\mathcal{T}^0\), where \(\mathcal{V}_\mathcal{T}^+\) corresponds to the layers distant from the root,

\[
\mathcal{V}_\mathcal{T}^+ \mu(t) = \begin{cases} 0, & t \in \Delta, \\ \mathcal{V}_\mathcal{T}^0 \mu(t), & t \not\in \Delta, \end{cases}
\]

while \(\mathcal{V}_\mathcal{T}^0\) corresponds to the first layers

\[
\mathcal{V}_\mathcal{T}^0 \mu(t) = \begin{cases} \mathcal{V}_\mathcal{T}^+ \mu(t), & t \in \Delta, \\ 0, & t \not\in \Delta. \end{cases}
\]

The idea behind this splitting is simple: the operator \(\mathcal{V}_\mathcal{T}^+\) has a small norm while \(\mathcal{V}_\mathcal{T}^0\) has a small image dimension. We first study the operator \(\mathcal{V}_\mathcal{T}^+\). Notice that

\[
\|\mathcal{V}_\mathcal{T}^+\| \leq \left( \sum_{l \geq \ln n/4} (1+l)^{-2} \right)^{1/2} \leq \left( \ln n/4 \right)^{-1/2} = 2(\ln n)^{-1/2}.
\]

For any tree \(\mathcal{T}\) of size bounded by \(m\), from \([2.13]\) we get

\[
e_k(\mathcal{V}_\mathcal{T}^+) \leq c \ln (m+1) \|\mathcal{V}_\mathcal{T}^+\| k^{-1/2}
\]

and applying this with \(k = n, m = n + 1\) we obtain

\[
e_n(\mathcal{V}_\mathcal{T}^+) \leq c n^{-1/2}.
\]

Now we have to consider the operator \(\mathcal{V}_\mathcal{T}^0\).

Notice that since weights on higher levels vanish, the operator \(\mathcal{V}_\mathcal{T}^0\) actually acts into \(\ell_{2,W}(\Delta)\). The size of \(\Delta\) is merely \(2^{1+\ln n/4} \leq 2n^{1/4}\); thus estimation can be rather crude.

Write \(\mathcal{V}_\mathcal{T}^0 = I \circ \mathcal{V}_\mathcal{T}^{00}\), where \(\mathcal{V}_\mathcal{T}^{00}\) is the same operator as \(\mathcal{V}_\mathcal{T}^0\) but acting into \(\ell_\infty(\Delta)\) and \(I\) is the embedding of \(\ell_\infty(\Delta)\) in \(\ell_{2,W}(\Delta)\). The operator \(\mathcal{V}_\mathcal{T}^{00}\) is a contraction, since

\[
\|\mathcal{V}_\mathcal{T}^{00} \mu\|_\infty = \max_{\mu \in \Delta} \left| \sum_{u \in \mathcal{O}(t) \cap \mathcal{T}} \mu(u) \right| \leq \|\mu\|_1.
\]

On the other hand, we can easily evaluate the entropy of \(I\). The net \(\mathcal{H}_\Delta \subset \ell_{2,W}(\Delta)\) will consist of all possible functions \(h\) of the form

\[
h(t) = j(t)n^{-1}, \quad t \in \Delta,
\]

where \(j(t)\) are odd integers satisfying \(|j(t)| \leq n\). Notice that there are no more than \(2n\) choices for each \(j(t)\).
Now we provide the estimates for approximation error and for the size of $\mathcal{H}_\Delta$. We start with evaluating approximation error. Let $x \in \ell_\infty(\Delta)$ be such that $\|x\|_\infty \leq 1$. Then for any $t \in \Delta$ we have $|x(t)| \leq 1$; hence, there exists a function $h \in \mathcal{H}_\Delta$ such that

$$|x(t) - h(t)| \leq n^{-1}, \quad \forall t \in \Delta.$$ 

Therefore,

$$\|Ix - h\|_2^2 = \sum_{t \in \Delta} w_t |x(t) - h(t)|^2 \leq |\Delta| n^{-2} \leq 2 n^{-7/4}.$$ 

The size of $\mathcal{H}_\Delta$ is bounded by

$$|\mathcal{H}_\Delta| \leq (2n)^{|\Delta|} \leq (2n)^{2n^{1/4}} = 2^{2n^{1/4}(1 + \log_2 n)} \leq 2^{2(n+1)}.$$ 

We conclude that

$$e_{2n+3}(\mathcal{V}_T^0) \leq e_{2n+3}(I) \leq 2n^{-7/8},$$ 

and we are done with the operator $\mathcal{V}_T^0$, too.

Having the bounds for both $e_n(\mathcal{V}_T^0)$ and $e_n(\mathcal{V}_T^+)$, by standard entropy estimates we get a bound for the sum of operators, i.e.

$$e_n(\mathcal{V}_T^+) \leq c n^{-1/2}.$$ 

Finally, it follows from (2.12) that

$$e_n(\mathcal{V}^*) \leq c n^{-1/2},$$ 

as required in the assertion of Theorem 2.4. Once the bound for $e_n(\mathcal{V}^*)$ is obtained, the bound for $e_n(\mathcal{V})$ follows from the famous duality connection for entropy numbers,

$$(2.15) \quad e_n(\mathcal{V}) \leq c_1 e_{c_2 n}(\mathcal{V}^*)$$

for some numerical constants $c_1, c_2$, which is still a conjecture for general Banach spaces but is a proved statement in our situation (one of the spaces is a Hilbert one); see [1], [2]. Actually, an older result [14] would suffice. $\square$

3. Entropy of an integral operator

Let $r < e^{-2}$ be a small number. In this section $(\cdot, \cdot)$ and $\|\cdot\|$ denote the scalar product and the norm in $L_2[0, r]$, respectively. We denote by $M[0, r]$ the space of signed measures of finite variation and $\|\cdot\|_1$ the respective variation norm. Moreover, $\|\mu\|_1(I)$ stands for the variation of $\mu \in M[0, r]$ on an interval $I$.

Our aim is to study the critical integral operator $\mathcal{V} : L_2[0, r] \to C[0, r]$ defined by

$$\mathcal{V} f(t) = \int_0^t f(s) K_t(s) ds = (f, K_t), \quad 0 \leq t \leq r,$$

and its adjoint $\mathcal{V}^* : M[0, r] \to L_2[0, r]$ defined by

$$\mathcal{V}^* \mu(s) = \int_0^r K_t(s) \mu(dt), \quad 0 \leq s \leq r,$$
where the critical kernel is
\[ K_t(s) = (t - s)^{-1/2} \ln(t - s)^{-1}. \]

Before we start the studies of \( K \), let us explain why it is critical in our context. Consider the family of kernels
\[ K_t^{(\beta)}(s) = (t - s)^{-1/2} \ln(t - s)^{-\beta}, \quad 1/2 < \beta < \infty, \]
and the corresponding operators \( V_\beta \). It is known from the works of Linde and Lacey \[10, 12\] that
\[
\begin{align*}
&c n^{1/2-\beta} \leq e_n(V_\beta) \leq C n^{1/2-\beta}, \quad 1/2 < \beta < 1, \\
&c n^{-1/2} \leq e_n(V_\beta) \leq C n^{-1/2} \ln n, \quad \beta = 1, \\
&c n^{-1/2} (\ln n)^{1-\beta} \leq e_n(V_\beta) \leq C n^{-1/2} (\ln n)^{1-\beta}, \quad \beta > 1.
\end{align*}
\]

Therefore, we see that the most interesting kernel \( K = K^{(1)} \) lays on the boundary between two different regimes and observe a logarithmic gap between the lower and upper bounds. The situation is exactly the same as in Theorem \[2.4\].

The main property of the kernel \( K \) we need is its modulus of continuity\[1\] (3.1) For all positive integers \( n \) and \( r \) it is true that
\[ ||K_{t+u} - K_t||_2 \leq 2|\ln u|^{-1/2}. \]

**Proof of Lemma 3.1.** We have
\[
||K_{t+u} - K_t||_2^2 = \int_0^t \left( (t - s)^{-1/2} |\ln(t - s)|^{-1} - (t + u - s)^{-1/2} |\ln(t + u - s)|^{-1} \right)^2 ds \\
+ \int_t^{t+u} (t + u - s)^{-1} |\ln(t + u - s)|^{-1} ds
\]
\[
= \int_0^t \left( v^{-1/2} |\ln v|^{-1} - (v + u)^{-1/2} |\ln(v + u)|^{-1} \right)^2 dv + \int_0^u v^{-1} |\ln v|^{-2} dv
\]
\[
\leq \int_0^u \left( v^{-1/2} |\ln v|^{-1} - (v + u)^{-1/2} |\ln(v + u)|^{-1} \right)^2 dv + 2 \int_0^u v^{-1} |\ln v|^{-2} dv
\]
\[
\leq u^2 \int_u^0 v^{-3} |\ln v|^{-2} dv + 2 \int_0^u v^{-1} |\ln v|^{-2} dv \leq |\ln u|^{-2} + 2 |\ln u|^{-1},
\]
and the assertion of the lemma follows. \( \square \)

**Theorem 3.2.** For all positive integers \( n \) and for a numerical constant \( C \) we have
\[ \max \{ e_n(V), e_n(V^*) \} \leq \frac{C}{n^{1/2}}. \]

**Proof of Theorem 3.2.** We repeat the ideas applied earlier to the summation operator on a binary tree. We first find a family of good finite rank approximations to \( V^* \) by giving an interpretation for \( n \)-essential subtrees. We will construct an \( n \)-essential partition \( T_n^u \) of \([0, r]\) as follows. Given a positive integer \( n \) and an element

---

1 However, we will also use the fact that the kernel \( u^{-1/2} |\ln(u)|^{-1} \) is a decreasing convex function on \([0, r]\) by the choice of \( r \).
μ ∈ M[0, r] we start dividing the interval [0, r] in halves and continue dividing while a (binary) interval \( I = \left( \frac{i r}{2^l}, \frac{(i+1) r}{2^l} \right) \) subject to division satisfies

\[
\|\mu\|_1(I) \geq \frac{l}{n}.
\]

Once an interval does not satisfy (3.2) we do not divide it and include it in our partition \( I^\mu_n \). If \( \|\mu\|_1 \leq 1 \), the condition (3.2) fails for \( l > n \). Therefore, our construction provides a finite partition of [0, r] in binary intervals of variable length.

The partition \( I^\mu_n \) depends on \( \mu \), but we will now show that the number of possible partitions and their size are rather limited.

Let \( D \) be the set of all binary intervals we divided during the construction of \( I^\mu_n \).

Notice that \( D \) is a tree w.r.t. inclusion. Let \( Q \) be the set of all terminal intervals of \( D \). I n other words, \( I ∈ Q \) iff \( I \) satisfies (3.2), but neither of its halves satisfies it. It is important for us that \( Q \) uniquely determines both \( D \) and \( I^\mu_n \). Indeed, any subtree of the binary tree is determined by the set of its terminal nodes. Thus \( Q \) determines \( D \). Moreover, \( I^\mu_n \) consists of all direct offsprings of elements of \( D \) that do not belong to \( D \).

Let \( q_l = \#\{ I ∈ Q : |I| = 2^{-l} r \} \). Then by (3.2)

\[
1 \geq \|\mu\|_1 \geq \sum_{l ∈ Q} \|\mu\|_1(I) = \sum_{l=0}^{∞} \sum_{I ∈ Q, |I| = 2^{-l}} \|\mu\|_1(I) \geq \sum_{l=0}^{∞} q_l \frac{l}{n}.
\]

Hence,

\[
\sum_{l=0}^{∞} q_l \cdot 1 \leq n.
\]

By Lemma 2.7, the number of possible trees \( Q \), thus the number of possible n-essential partitions, does not exceed \((4e)^n\). It is also worthwhile to notice that the number of intervals in \( I^\mu_n \) satisfies

\[
|I^\mu_n| \leq 2|D| \leq 2(n + 1)
\]

by Lemma 2.8.

Consider a finite dimensional approximation for \( V^\ast \) generated by any partition \( \mathcal{I} \), the operator \( V^\ast_\mathcal{I} : M[0, r] → L_2[0, r] \) defined by

\[
(V^\ast_\mathcal{I} \mu) = \sum_{I ∈ \mathcal{I}} \mu(I) K_{t_I},
\]

where \( t_I \) is the left end of \( I \). We evaluate the approximation error \( Δ_\mathcal{I} = V^\ast - V^\ast_\mathcal{I} \).

By the definition,

\[
(Δ_\mathcal{I} \mu) = \sum_{I ∈ \mathcal{I}} \int_I (K_I - K_{t_I}) \mu(dt).
\]

We are going to show that the approximation error is particularly small when we use the n-essential partition.

**Proposition 3.3.** For any \( n ∈ \mathbb{N} \) and any \( \mu \) with \( \|\mu\|_1 \leq 1 \) we have

\[
\|Δ_\mathcal{I}^\mu\|_2 \leq \frac{C}{\sqrt{n}}.
\]
Proof of Proposition 3.3. Let $\mu = \mu_+ - \mu_-$ be the Hahn decomposition of $\mu$. It is enough to show that

$$
||\Delta T^n_\mu \mu_+||_2 \leq \frac{C}{\sqrt{n}}
$$

and to prove the similar inequality for $\mu_-$. We start with

$$
||\Delta T^n_\mu \mu_+||_2 = (\Delta T^n_\mu \mu_+, \Delta T^n_\mu \mu_+)
$$

$$
= \sum_{i_1, i_2 \in \mathcal{I}_n^+} \left( \int_{I_{i_1}} (K_{i_1} - K_{i_2}) \mu_+(dt), \int_{I_{i_2}} (K_{i_2} - K_{i_1}) \mu_+(dt) \right)
$$

$$
= \sum_{i_1, i_2 \in \mathcal{I}_n^+} \int_{I_{i_1}} \int_{I_{i_2}} (K_{i_1} - K_{i_2}, K_{i_2} - K_{i_1}) \mu_+(dt_1) \mu_+(dt_2).
$$

For the main (diagonal) terms of this sum we have

$$
\sum_{I \in \mathcal{I}_n^+} \int_{I} \int_{I} (K_i - K_j, K_j - K_i) \mu_+(dt_1) \mu_+(dt_2)
$$

$$
\leq \sum_{I \in \mathcal{I}_n^+} \int_{I} \int_{I} ||K_i - K_j||_2 ||K_j - K_i||_2 \mu_+(dt_1) \mu_+(dt_2)
$$

$$
\leq \sum_{I \in \mathcal{I}_n^+} \max_{I} ||K_i - K_j||_2^2 \mu_+(I)^2
$$

$$
\leq \sum_{I} \sum_{l \in \mathcal{I}_n^+, |l| = 2^{-r}} 4(\ln |l|)^{-1} \mu_+(l)^2 \quad \text{by Lemma 3.1}
$$

$$
\leq \sum_{l} \sum_{I \in \mathcal{I}_n^+} 4(\ln 2 \cdot l)^{-1} \cdot \frac{l}{n} \mu_+(I) \quad \text{by the definition of } \mathcal{I}_n^+
$$

$$
= \frac{4}{(\ln 2) n} \sum_{I \in \mathcal{I}_n^+} \mu_+(I) \leq \frac{4}{(\ln 2) n}.
$$

Unlike the tree case, the summands in the definition of $\Delta T$ are not orthogonal, therefore we cannot stop here. We will show that the non-diagonal terms do not give a positive contribution to the quantity we evaluate. From the probabilistic viewpoint, this is very much in the spirit of negative dependence of the increments for fractional Brownian motion with parameter $H < 1/2$. We will state the corresponding result as an independent lemma.

Let $g : \mathbb{R} \to \mathbb{R}$ be a function such that $g$ vanishes on $(-\infty, 0]$ and $g$ is a decreasing convex non-negative function on $(0, +\infty)$. Let $K_a(\cdot) = g(t - \cdot)$ for $t \geq 0$.

Lemma 3.4. For all $0 \leq a \leq b \leq c \leq d \leq r$ we have

$$
\int_0^r (K_d - K_c)(K_b - K_a) \leq 0.
$$

Proof of Lemma 3.4. First of all, let us notice that the function $s \to K_d(s) - K_c(s) = g(d-s) - g(c-s)$ is non-positive and non-increasing, while $s \in [0, c] \subseteq [0, b]$. Next, the function $s \to K_b(s) - K_a(s) = g(b-s) - g(a-s)$ is positive on $[a, b]$ and
negative on $[0, a]$. Therefore,

$$
\int_0^r (K_d - K_c)(K_b - K_a) = \int_0^b (K_d - K_c)(K_b - K_a) = \left( \int_0^a + \int_a^b \right) (K_d - K_c)(K_b - K_a) \leq (K_d(a) - K_c(a)) \left( \int_0^a (K_b - K_a) + \int_a^b K_b \right)
$$

$$
= (K_d(a) - K_c(a)) \left( \int_{b-a}^b g + \int_0^a g - \int_0^{b-a} g \right)
$$

$$
= (K_d(a) - K_c(a)) \int_a^b g \leq 0.
$$

By applying this result to our function $g(t) = t^{-1/2} \ln |t|^{-1}$ we obtain for any $t_1 \in I_1 \in \mathcal{I}_n, t_2 \in I_2 \in \mathcal{I}_n$,

$$(K_{t_1} - K_{t_1}, K_{t_2} - K_{t_2}) = \int_0^r (K_{t_1} - K_{t_1}, K_{t_2} - K_{t_2}) \leq 0,$$

provided $I_1 \neq I_2$. Hence,

$$
\sum_{I_1, I_2 \in \mathcal{I}_n, I_1 \neq I_2} \int_{I_1} \int_{I_2} (K_{t_1} - K_{t_1}, K_{t_2} - K_{t_2}) \mu_+(dt_1) \mu_+(dt_2) \leq 0.
$$

Therefore,

$$
\| \Delta_{\mathcal{I}_n} \mu_+ \|_2^2 \leq \sum_{I \in \mathcal{I}_n} \int_I \int_I (K_{t_1} - K_{t_1}, K_{t_2} - K_{t_1}) \mu_+(dt_1) \mu_+(dt_2) \leq \frac{4}{(\ln 2)^n},
$$

and (3.5) follows. The same inequality for $\mu_-$ is obtained by applying (3.5) to $-\mu$.

Now (3.4) is proved completely.

We continue the proof of Theorem 3.2. Let $\mathcal{J}_n = \{I \in \mathcal{I}_n : |\mu_1| \leq 1\}$ be the set of all possible $n$-essential partitions of $[0, r]$. Recall that

$$
|\mathcal{J}_n| \leq (4e)^n.
$$

We claim that

$$
\sup_{I \in \mathcal{J}_n} e_n(\mathcal{V}_I^+) \leq \frac{C}{n^{1/2}}.
$$

Assuming this is obtained, the application of Lemma 2.8 to the family of operators \{\mathcal{V}_I, I \in \mathcal{J}_n\}, along with the estimate of approximation error (3.4) and the estimate for the number of operators (3.6), lead to

$$
e_n(\mathcal{V}) \leq \frac{C}{n^{1/2}},
$$

as required by the assertion of Theorem 3.2.

The same estimate for $e_n(\mathcal{V})$ now follows by the duality argument (2.15).
Now it only remains to prove (3.7). Let us fix a partition \( \mathcal{I} \in \mathcal{J} \). From now on, we do not need any particular properties of \( n \)-essential partitions, except for the size bound (3.3).

Consider an auxiliary partition \( \mathcal{E} \) of \([0, r]\) constructed as follows. Take \( m \) such that \( 2^{-m} \leq n^{-1/4} \leq 2^{1-m} \). Divide \([0, r]\) in binary intervals of length \( r2^{-m} \). If a union of such intervals belongs to \( \mathcal{I} \), then replace them by this union. The result is a partition \( \mathcal{E} \). Notice that \( \mathcal{I} \) is a refinement of \( \mathcal{E} \) and \(|\mathcal{E}| \leq 2^m \leq 2n^{1/4} \). Write

\[
V^*_{\mathcal{I}} = V^*_{\mathcal{E}} + (V^*_{\mathcal{I}} - V^*_{\mathcal{E}})
\]

and evaluate the entropy of both operators.

First we handle the low rank operator \( V^*_{\mathcal{E}} \). Recall that \( V^*_{\mathcal{E}} \mu = \sum_{I \in \mathcal{E}} K_{t_I} \mu(I) \).

Consider the net

\[
\mathcal{N}_E = \left\{ \sum_{I \in \mathcal{E}} K_{t_I} \frac{j_I}{n}, j_I \in \{1 - n, \ldots, 0, \ldots, n - 1\} \right\}.
\]

Notice that

\[
|\mathcal{N}_E| \leq (2n - 1)^{|\mathcal{E}|} \leq (2n)^{2n^{1/4}} \leq 2^{2(n+1)}.
\]

On the other hand, for any \( \mu \) with \( ||\mu||_1 \leq 1 \) find an \( h = \sum_{I \in \mathcal{E}} K_{t_I} \frac{j_I}{n} \) such that \( \max_{I \in \mathcal{E}} |\mu(I) - \frac{j_I}{n}| \leq n^{-1} \). We have

\[
||V^*_{\mathcal{E}} \mu - h||_2 \leq \sum_{I \in \mathcal{E}} ||K_{t_I}||_2 \cdot |\frac{j_I}{n} - \mu(I)| \leq \max_{t \in [0,1]} ||K_t||_2 \cdot n^{-1} \cdot |\mathcal{E}| \leq C n^{-3/4}.
\]

It follows that

\[
\epsilon_{2n+3}(V^*_{\mathcal{E}}) \leq C n^{-3/4}.
\]

Now we handle the operator \( V^*_{\mathcal{I}} - V^*_{\mathcal{E}} \) which has a larger rank but smaller norm. By the definition,

\[
(V^*_{\mathcal{I}} - V^*_{\mathcal{E}}) \mu = \sum_{J \in \mathcal{I}} K_{t_J} \mu(J) - \sum_{I \in \mathcal{E}} K_{t_I} \mu(I)
= \sum_{J \in \mathcal{E} \setminus \mathcal{I}} K_{t_J} \mu(J) - \sum_{I \in \mathcal{E} \setminus \mathcal{I}} K_{t_I} \mu(I)
= \sum_{I \in \mathcal{E} \setminus \mathcal{I}} \sum_{J \in \mathcal{I}, J \subset I} (K_{t_J} - K_{t_I}) \mu(J).
\]

Notice that the conditions \( I \in \mathcal{E} \setminus \mathcal{I}, J \subset I \) imply

\[
|t_J - t_I| \leq |I| = 2^{-m} \leq n^{-1/4}.
\]

Hence by Lemma 3.1

\[
||(V^*_{\mathcal{I}} - V^*_{\mathcal{E}}) \mu||_2 \leq 2(\ln(n^{1/4}))^{-1/2} \sum_{J \in \mathcal{I}} |\mu(J)| \leq (\ln n)^{-1/2} ||\mu||_1,
\]

which simply means that \( ||V^*_{\mathcal{I}} - V^*_{\mathcal{E}}|| \leq (\ln n)^{-1/2} \).

On the other hand, recall that by (3.3)

\[
\text{rank } (V^*_{\mathcal{I}} - V^*_{\mathcal{E}}) \leq |\mathcal{I}| \leq 2(n + 1).
\]
We apply, as we did in the investigation of tree summation, the estimate (2.13) and obtain
\[ e_k(V_T^* - V_E^*) \leq c \ln^{1/2}(\text{rank}(V_T^* - V_E^*) + 1) \|V_T^* - V_E^*\| k^{-1/2} \]
\[ \leq c \ln^{1/2}(2n + 3) \cdot (\ln n)^{-1/2} \cdot k^{-1/2} \]
\[ \leq c k^{-1/2}. \]

By letting \( k = n \),
\[ e_n(V_T^* - V_E^*) \leq c n^{-1/2}. \]

We conclude that
\[ e_{3n+1}(V_T^*) \leq e_{2n+2}(V_E^*) + e_n(V_T^* - V_E^*) \leq c n^{-1/2}, \]
and (3.7) follows. \( \square \)

4. Relation to the entropy of convex hulls

Recall a well-known problem from the geometry of Banach spaces; see e.g. [4], [11]. Let \( A \) be a set in a Hilbert space and \( \text{aco}A \) its absolutely convex hull. If we know the behavior of entropy numbers \( e_n(A) \), what can we say about \( e_n(\text{aco}A) \)?

As we already mentioned in the proof of Theorem 2.1, it is known that for \( \beta < 1/2 \)
\[ e_n(A) \leq c n^{-\beta} \text{ implies } e_n(\text{aco}A) \leq C n^{-\beta}, \]
while
\[ e_n(A) \leq c n^{-1/2} \text{ only implies } e_n(\text{aco}A) \leq C n^{-1/2} \ln n. \]

F. Gao [9] was the first to construct a critical set \( A \) with properties (4.1)
\[ e_n(A) \leq c n^{-1/2} \text{ and } e_n(\text{aco}A) \geq C n^{-1/2} \ln n. \]

We can call a \textit{Gao set} any set satisfying (4.1). Later on, his arguments were streamlined and extended to the non-Hilbert case by J. Creutzig and I. Steinwart [7].

The relation to our problem is the following. For simplicity, let us take a setting of the summation operator on the binary tree (Section 2). Consider the critical summation operator \( V^* \) with the weight (2.1) and \( \beta = 2 \). Take a set
\[ A = \{ V^*(1_t), \ t \in T \} \]
in the Hilbert space \( \ell_2(T, W) \). It is plain that \( e_n(A) \leq n^{-1/2} \), and we see from the general fact above that \( e_n(\text{aco}A) \leq C n^{-1/2} \ln n \). Since it is quite difficult to get a better upper bound, one could think of \( A \) as a candidate to be a Gao set, although of a nature very different from the known ones.

On the other hand, \( \text{aco}A \) is the image of the unit ball w.r.t. the operator \( V^* \). In other words, \( e_n(V^*) = e_n(\text{aco}A) \). Therefore, Theorem 2.4 shows that \( A \) is \textit{not} a Gao set.

Curiously enough, one can construct another set \( B = \{ b_t, t \in T \} \) such that \( B \) is a Gao set, while the distance between any pair of points of \( B \) is smaller than that between their counterparts in \( A \). There is no contradiction, of course, but the reader can see from this observation that the problem is really delicate.

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
Acknowledgements

The author is grateful to M. Lacey and W. Linde for the problem statement, for interesting discussions around it and for informing him about their unpublished results.

Sincere thanks are also due to the American Institute of Mathematics (Palo-Alto) and its staff for organizing the conference “Small ball inequalities in analysis, probability, and irregularities of distribution” (December 2008), where many people mentioned here had a chance to discuss this and related problems.

References


Department of Mathematics and Mechanics, St. Petersburg State University, St. Petersburg, Russia

E-mail address: lifts@mail.rcom.ru