Entropy numbers of embeddings of Schatten classes

Aicke Hinrichs[*], Joscha Prochno[†], Jan Vybíral[‡]

December 23, 2016

Abstract

Let $0 < p, q \le \infty$ and denote by $\mathcal{S}_p^N$ and $\mathcal{S}_q^N$ the corresponding finite-dimensional Schatten classes. We prove optimal bounds, up to constants only depending on p and q, for the entropy numbers of natural embeddings between $\mathcal{S}_p^N$ and $\mathcal{S}_q^N$. This complements the known results in the classical setting of natural embeddings between finite-dimensional ℓ_p spaces due to Schütt, Edmunds-Triebel, Triebel and Guédon-Litvak/Kühn. We present a rather short proof that uses all the known techniques as well as a constructive proof of the upper bound in the range $N \le n \le N^2$ that allows deeper structural insight and is therefore interesting in its own right. Our main result can also be used to provide an alternative proof of recent lower bounds in the area of low-rank matrix recovery.

Keywords: Entropy numbers, Schatten classes, Grassmann manifolds

1 Introduction and Main Results

The probably most important class of unitary operator ideals are the Schatten p-classes $\mathcal{S}_p$ ($0 < p \le \infty$) consisting of all compact operators between Hilbert spaces for which the sequence of their singular values belongs to ℓ_p. In many ways, the structure of these spaces resembles that of the classical sequence spaces ℓ_p, and the Schatten classes $\mathcal{S}_p$ form, in the sense of multiplication, their non-commutative counterpart. Some classical results on the topological structure of $\mathcal{S}_p$ spaces can be found, for instance, in [16] and [2]. Recent years have seen increased interest in those non-commutative L_p spaces. One such case, in part motivating our work, lies in the area of compressed sensing in the form of low-rank matrix recovery (see, e.g., [4, 5, 8] and the references therein.)

The study of bounded linear operators between Banach spaces is a classical part of functional analysis. Of particular interest are compact operators, that is, those operators for which the image of the closed unit ball under the operator is relatively compact in the codomain space. In general, the extent to which a bounded linear operator is compact is measured by its sequence of entropy numbers. The n-th (dyadic) entropy number of

---

[*]Institute of Analysis, University Linz, Altenberger Str. 69, 4040 Linz, Austria, `aicke.hinrichs@jku.at`

[†]School of Mathematics & Physical Sciences, University of Hull, Cottingham Road, HU6 7RX Hull, United Kingdom, `j.prochno@hull.ac.uk`

[‡]Department of Mathematical Analysis, Charles University, Sokolovská 83, 186 00, Prague 8, Czech Republic, `vybiral@karlin.mff.cuni.cz`; this author was supported by the ERC CZ grant LL1203 of the Czech Ministry of Education and by the Neuron Fund for Support of Science.

a bounded linear operator T between Banach spaces X and Y (denoted by $e_n(T)$ in the sequel) is the infimum over all $\varepsilon > 0$ such that there are $y_1, \ldots, y_{2^{n-1}} \in Y$ with

$$T(B_X) \subseteq \bigcup_{j=1}^{2^{n-1}} (y_j + \varepsilon B_Y),$$

where B_X, B_Y denote the closed unit balls of X and Y, respectively. In fact, it is true that an operator between Banach spaces is compact if and only if the sequence of its entropy numbers converges to 0. The theory of entropy numbers and their systematic study was introduced by Pietsch in [19].

In the case of symmetric Banach sequence spaces, entropy numbers of diagonal operators were studied by Schütt in 1984 [23], where he obtained bounds of exact order up to a logarithmic factor. In the particular case of the natural identity between ℓ_p^N and ℓ_q^N in the Banach space regime $1 \le p < q \le \infty$, we write $\ell_p^N \hookrightarrow \ell_q^N$, he obtained the following asymptotically sharp estimates for the n-th dyadic entropy number e_n [23, Theorem 1]:

If $1 \le p < q \le \infty$ and $n, N \in \mathbb{N}$, then

$$e_n\big(\ell_p^N \hookrightarrow \ell_q^N\big) \asymp_{p,q} \begin{cases} 1, & 1 \le n \le \log N; \\ \left(\frac{\log(N/n+1)}{n}\right)^{1/p-1/q}, & \log N \le n \le N; \\ 2^{-n/N} \cdot N^{1/q-1/p}, & N \le n. \end{cases} \tag{1.1}$$

It was in 1996 that Edmunds and Triebel proved corresponding upper bounds also for the quasi-Banach space regime $0 < p \le q \le \infty$ [10, Proposition 3.2.2]. Shortly after, complementing and corresponding lower estimates for $1 \le n \le \log N$ and $N \le n$ were obtained by Triebel in [26, Theorem 7.3]. Finally, the gap on the medium scale, $\log N \le n \le N$, was closed by Guédon and Litvak [12] and, independently, by Kühn [15].

The main subject of this paper is to prove the non-commutative counterpart of (1.1), that is, the computation of entropy numbers of the natural identities between finite-dimensional Schatten classes in the full range $0 < p, q \le \infty$. For $0 < p \le \infty$, let $\mathcal{S}_p^N$ denote the N^2-dimensional space of all real $N \times N$ matrices A equipped with the Schatten p-norm

$$\|A\|_{\mathcal{S}_p^N} = \left(\sum_{j=1}^{N} \sigma_j(A)^p\right)^{1/p},$$

where $\sigma_j(A)$, $j = 1, \ldots, N$ are the singular values of A. For $0 < p, q \le \infty$, we denote the natural identity map from $\mathcal{S}_p^N$ to $\mathcal{S}_q^N$ by $\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N$.

Our main result is the following theorem.

Theorem 1.1. *Let $n, N \in \mathbb{N}$ and $0 < p, q \le \infty$. Then,*

$$e_n\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \asymp_{p,q} \begin{cases} 2^{-\frac{n}{N^2}} \cdot N^{1/q-1/p}, & 0 < q \le p \le \infty; \\ 1, & 0 < p \le q \le \infty, 1 \le n \le N; \\ \left(\frac{N}{n}\right)^{1/p-1/q}, & 0 < p \le q \le \infty, N \le n \le N^2; \\ 2^{-\frac{n}{N^2}} \cdot N^{1/q-1/p}, & 0 < p \le q \le \infty, N^2 \le n, \end{cases}$$

where $\asymp_{p,q}$ denotes equivalence up to constants depending only on p and q.

The following remark briefly explains the connection of Theorem 1.1 to low-rank matrix recovery.

Remark 1.2. Our main result can be used to provide an alternative proof of lower bounds in the area of low-rank matrix recovery. Let us assume that we want to recover a matrix X from the unit ball of $\mathcal{S}_p^N$ (which we denote by B_p^N in the sequel) from m linear measurements of X represented by the information map

$$\mathcal{A}(X) = (\langle A_1, X \rangle, \ldots, \langle A_m, X \rangle) \in \mathbb{R}^m.$$

The search for the optimal pair of information map $\mathcal{A} : \mathbb{R}^{N \times N} \to \mathbb{R}^m$ and recovery map $\Delta : \mathbb{R}^m \to \mathbb{R}^{N \times N}$ is expressed in terms of the quantities

$$E_m\big(B_p^N, \mathcal{S}_q^N\big) = \inf_{\substack{\mathcal{A} : \mathbb{R}^{N \times N} \to \mathbb{R}^m \\ \Delta : \mathbb{R}^m \to \mathbb{R}^{N \times N}}} \sup_{X \in B_p^N} \|X - \Delta(\mathcal{A}(X))\|_{\mathcal{S}_q^N}.$$

Using their relation to Gelfand numbers and Carl's inequality (see Section 2 and [11, Section 10] for details), Theorem 1.1 immediately implies that, for $0 < p \leq q \leq \infty$ and $1 \leq m \leq N^2$,

$$E_m\big(B_p^N, \mathcal{S}_q^N\big) \gtrsim_{p,q} \min\Big\{1, \frac{N}{m}\Big\}^{1/p - 1/q}.$$

This provides another proof for the lower bound recently given in [8, Theorem 5.5], where the same result was obtained in the regime $0 < p \leq 1$ and $p < q \leq 2$.

As typical in the theory of Banach spaces, our proofs reflect a variety of ideas and tools of analytical and geometric flavor. These include

- Carl's inequality for the case of quasi-Banach spaces [13]

- two-sided asymptotic estimates for the Gelfand numbers of identities of Schatten classes [8]

- volume estimates for metric balls in the Grassmannian going back to the work of Szarek [24] (see also [17]), and

- interpolation estimates of entropy numbers in the domain and target spaces (see [18] and [20]).

The rest of the paper is organized as follows. In the next section, we give the necessary definitions and notation and recall some results needed for the proof of Theorem 1.1. In Section 3 a proof of Theorem 1.1, which is intended to be as short as possible using all known machinery to estimate entropy numbers, is presented. Most of the cases are rather standard, but some require deeper tools from the structure theory of Schatten classes. For some purposes it is much more instructive to find explicit covers achieving optimal upper bounds for the entropy numbers. This construction will be given in Section 4.

2   Notation and Preliminaries

In this section we present the notation and preliminaries we use throughout this text. For the sake of clarity, we subdivide the background material by topic. Since we have a rather broad audience in mind, we try to keep this work as self contained as possible.

2.1 General notation

The cardinality of a set A will be denoted by $|A|$. We denote by $\mathrm{vol}_n(A)$ the n-dimensional Lebesgue measure of a Borel measurable set $A \subseteq \mathbb{R}^n$ and will usually suppress the index n for brevity. We supply the n-dimensional Euclidean space $\mathbb{R}^n$ with its standard inner product $\langle \cdot, \cdot \rangle$ and corresponding Euclidean structure $\| \cdot \|_2$. For $0 < p \leq \infty$, we denote by ℓ_p^n the space $\mathbb{R}^n$ equipped with the (quasi-)norm

$$\|(x_i)_{i=1}^n\|_p = \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}, & 0 < p < \infty; \\ \max_{1 \leq i \leq n} |x_i|, & p = \infty. \end{cases}$$

Given two (quasi-)Banach spaces X, Y, we denote by B_X the closed unit ball of X and by $\mathcal{L}(X, Y)$ the space of bounded linear operators between X and Y equipped with the standard operator (quasi-)norm.

Finally, for two sequences $(a(n))_{n\in\mathbb{N}}$ and $(b(n))_{n\in\mathbb{N}}$ of non-negative real numbers, we write $a(n) \asymp b(n)$ provided that there exist constants $c, C \in (0, \infty)$ such that $cb(n) \leq a(n) \leq Cb(n)$ for all $n \in \mathbb{N}$. If the constants depend on some parameter p, we shall write $a(n) \lesssim_p b(n)$, $a(n) \gtrsim_p b(n)$ or, if both hold, $a(n) \asymp_p b(n)$.

2.2 Quasi-Banach spaces

We briefly present some background on quasi-Banach spaces. If X is a (real) vector space, then we call $\| \cdot \|_X : X \to [0, \infty)$ a quasi-norm if

1. $\|x\|_X = 0$ if and only if $x = 0$

2. $\|\alpha x\|_X = |\alpha| \cdot \|x\|_X$ for all $\alpha \in \mathbb{R}$ and $x \in X$

3. There exists a constant $C \in [1, \infty)$ such that $\|x + y\|_X \leq C(\|x\|_X + \|y\|_X)$ for all $x, y \in X$.

By the Aoki-Rolewicz theorem (see [1, 21]), every quasi-norm is equivalent to some p-norm, that is, there exists a mapping $\|| \cdot \||_X : X \to [0, \infty)$, which is equivalent to $\| \cdot \|_X$ on X, and some $0 < p \leq 1$ such that $\|| \cdot \||_X$ satisfies 1. and 2. above, while 3. is replaced by $\||x + y\||_X^p \leq \||x\||_X^p + \||y\||_X^p$. The mapping $\|| \cdot \||_X$ is then called a p-norm and if X is complete with respect to the metric induced by $\|| \cdot \||_X^p$, X is called a p-Banach space and the original space X with the quasi-norm $\| \cdot \|_X$ is called a quasi-Banach space. Therefore, as far as estimates up to a multiplicative constant are concerned, we may assume that a quasi-Banach space is equipped with an equivalent p-norm instead.

2.3 Singular values and Schatten p-classes

The singular values $\sigma_1, \ldots, \sigma_N$ of a real $N \times N$ matrix A are defined to be the square roots of the eigenvalues of the positive self-adjoint operator A^*A, which are simply the eigenvalues of $|A| := \sqrt{A^*A}$. The singular values are arranged in non-increasing order, that is, $\sigma_1(A) \geq \cdots \geq \sigma_N(A) \geq 0$. The singular value decomposition shall be used in the form $A = U\Sigma V^T$, where $U, V \in \mathbb{R}^{N \times N}$ are orthogonal matrices, and $\Sigma \in \mathbb{R}^{N \times N}$ is a diagonal matrix with $\sigma_1(A), \ldots, \sigma_N(A)$ on the diagonal.

For $0 < p \leq \infty$, the Schatten p-class $\mathcal{S}_p^N$ is the N^2-dimensional space of all $N \times N$ real matrices acting from ℓ_2^N to ℓ_2^N equipped with the Schatten p-norm

$$\|A\|_{\mathcal{S}_p^N} = \left(\sum_{j=1}^N \sigma_j(A)^p\right)^{1/p}.$$

4

Let us remark that $\| \cdot \|_{\mathcal{S}_1^N}$ is the nuclear norm, $\| \cdot \|_{\mathcal{S}_2^N}$ the Hilbert-Schmidt norm, and $\| \cdot \|_{\mathcal{S}_\infty^N}$ the operator norm. We denote the unit ball of $\mathcal{S}_p^N$ by

$$B_p^N := \left\{ A \in \mathbb{R}^{N \times N} \, : \, \|A\|_{\mathcal{S}_p^N} \leq 1 \right\}.$$

The following volume estimate for the unit balls of the Schatten p-classes can be found in [22, Corollary 8] for the case $p \geq 1$. Following along the lines of the proof in [22] it is not too hard to see that it carries over to the case $0 < p \leq 1$.

Lemma 2.1. *Let $0 < p \leq \infty$ and $N \in \mathbb{N}$. Then*

$$\mathrm{vol}(B_p^N)^{1/N^2} \asymp_p N^{-1/2-1/p}. \tag{2.1}$$

Please note that if $p \geq 1$, then $\| \cdot \|_{\mathcal{S}_p^N}$ is a norm, and if $0 < p < 1$, then $\| \cdot \|_{\mathcal{S}_p^N}$ is a p-norm, i.e., a quasi-norm satisfying the p-triangle inequality

$$\|A + B\|_{\mathcal{S}_p^N}^p \leq \|A\|_{\mathcal{S}_p^N}^p + \|B\|_{\mathcal{S}_p^N}^p, \qquad A, B \in \mathcal{S}_p^N.$$

For $p, q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$, the Schatten classes satisfy the following Hölder type inequality,

$$\|AB\|_{\mathcal{S}_1^N} \leq \|A\|_{\mathcal{S}_p^N} \|B\|_{\mathcal{S}_q^N}, \qquad A \in \mathcal{S}_p^N, B \in \mathcal{S}_q^N.$$

2.4 Unitarily invariant and symmetric norms

The Schatten p-norms (for $p \geq 1$) are examples of unitarily invariant norms. In particular, unitarily invariant norms correspond to 1-symmetric norms, similarly as the Schatten p-norm corresponds to the ℓ_p-norm.

A basis $\{e_i\}_{i=1}^N$ of an N-dimensional Banach space X is called 1-symmetric if and only if for all permutations π of $\{1, \ldots, N\}$, all signs $\varepsilon_i \in \{-1, 1\}$, $i = 1, \ldots, N$, and all $a_1, \ldots, a_N \in \mathbb{R}$,

$$\left\| \sum_{i=1}^n a_i e_i \right\|_X = \left\| \sum_{i=1}^N \varepsilon_i a_{\pi(i)} e_i \right\|_X.$$

A norm on $\mathcal{L}(\ell_2^N, \ell_2^N)$ is called a unitarily invariant norm, we write $\| \cdot \|_{inv}$, if it satisfies

$$\|A\|_{inv} = \|UAV\|_{inv}$$

for any $A \in \mathcal{L}(\ell_2^N, \ell_2^N)$ and all isometries U, V on ℓ_2^N. Since we are only considering the real case, U, V are simply orthogonal matrices. It is well-known and easy to see that to any such norm $\| \cdot \|_{inv}$ corresponds a 1-symmetric norm $\| \cdot \|$ on $\mathbb{R}^N$ so that, for all $A \in \mathcal{L}(\ell_2^N, \ell_2^N)$,

$$\|A\|_{inv} = \|(\sigma_i(A))_{i=1}^N\|,$$

where $(\sigma_i(A))_{i=1}^N$ is the vector of singular values of A, and, on the other hand, $\|A\|_{inv} = \|(x_1, \ldots, x_N)\|$ for any diagonal matrix $A = \mathrm{diag}(x_1, \ldots, x_N)$. For some background on unitarily invariant norms, we refer the reader to [14].

5

2.5 Geometry of Grassmannians

Given $N \in \mathbb{N}$ and $k \in \{0, 1, \ldots, N\}$, we denote by $G_{N,k}$ the Grassmannian of k-dimensional linear subspaces of $\mathbb{R}^N$. Recall that its dimension is $k(N-k)$. We equip $G_{N,k}$ with a metric $\mathrm{d}(\cdot, \cdot)$ induced by some unitarily invariant norm $\| \cdot \|_{inv}$ on $\mathcal{L}(\ell_2^N, \ell_2^N)$ by defining

$$\mathrm{d}(E, F) := \|P_E - P_F\|_{inv}, \qquad E, F \in G_{N,k}, \tag{2.2}$$

under the embedding $F \mapsto P_F$ of the Grassmann manifold into $\mathcal{L}(\ell_2^N, \ell_2^N)$, where P_E, P_F stand for the orthogonal projections onto E and F, respectively. Hence, we can and will think of the Grassmann manifold $G_{N,k}$ as the set of orthogonal projections on ℓ_2^N of given rank k. A typical example of such an induced metric on the Grassmannian $G_{N,k}$ is the one induced by the operator norm. Note that, since we identify $G_{N,k}$ with the orthogonal projections on ℓ_2^N with given rank k, the singular value decomposition shows that $k^{-1/p} P_E \in B_p^N$ for each $E \in G_{N,k}$. Therefore,

$$k^{-1/p} G_{N,k} \subseteq B_p^N \subseteq B_\infty^N.$$

The following result shows that the volume of metric balls in the Grassmannian is essentially independent of the chosen unitarily invariant norm. It is an immediate consequence of Szarek's result from [24] on nets in Grassmann manifolds.

Lemma 2.2 ([17], Corollary 2.2). *Let $1 \leq k \leq N-1$, $\mathrm{d}(\cdot, \cdot)$ be a metric on $G_{N,k}$ induced by some unitarily invariant norm on $\mathcal{L}(\ell_2^N, \ell_2^N)$ normalized such that $\mathrm{diam}(G_{N,k}, \mathrm{d}) = 1$. Let $0 < \delta < 1$. Then, for any $F \in G_{N,k}$,*

$$\left(\frac{\delta}{c_1}\right)^{k(N-k)} \leq \mu_{N,k}\big(B_{\mathrm{d}}(F, \delta)\big) \leq \left(\frac{\delta}{c_2}\right)^{k(N-k)},$$

where $c_1, c_2 \in (0, \infty)$ are absolute constants and $B_{\mathrm{d}}(F, \delta) := \{E \in G_{N,k} : \mathrm{d}(F, E) < \delta\}$.

We will need this lemma in the particular case that $\mathrm{d}(\cdot, \cdot)$ is induced by the Schatten q-norm. If $1 \leq k \leq N/2$ and $E, F \in G_{N,k}$ are orthogonal, we have

$$\mathrm{d}(E, F) := \|P_E - P_F\|_{\mathcal{S}_q^N} = (2k)^{1/q}.$$

On the other hand, the triangle inequality shows that $\mathrm{d}(E, F) \leq 2k^{1/q}$. Hence, the diameter d of $G_{N,k}$ with respect to the Schatten q-norm satisfies $(2k)^{1/q} \leq d \leq 2k^{1/q}$ and Lemma 2.2 implies

Lemma 2.3. *Let $1 \leq k \leq N/2$, $q \geq 1$ and $0 < \delta < k^{1/q}$. Then, for any $F \in G_{N,k}$,*

$$\left(\frac{\delta}{c_1 k^{1/q}}\right)^{k(N-k)} \leq \mu_{N,k}\big(B(F, \delta)\big) \leq \left(\frac{\delta}{c_2 k^{1/q}}\right)^{k(N-k)},$$

with absolute constants $c_1, c_2 \in (0, \infty)$ and $B(F, \delta) := \{E \in G_{N,k} : \|P_E - P_F\|_{\mathcal{S}_q^N} < \delta\}$.

2.6 Entropy numbers and other s-numbers

Let X be a quasi-Banach space and let $n \in \mathbb{N}$. The n-th dyadic entropy number of $A \subseteq X$, $e_n(A, X)$, is defined as

$$e_n(A) := e_n(A, X) := \inf \left\{ \varepsilon > 0 : \exists x_1 \ldots, x_{2^{n-1}} \in X : A \subseteq \bigcup_{j=1}^{2^{n-1}} \big(x_j + \varepsilon B_X\big) \right\}.$$

For a bounded linear operator T between quasi-Banach spaces X and Y, the n-th dyadic entropy number, $e_n(T)$, is defined as

$$e_n(T) := e_n\big(T(B_X), Y\big) = \inf\Big\{\varepsilon > 0 \ : \ \exists\, y_1, \ldots, y_{2^{n-1}} \in Y \ : \ T(B_X) \subseteq \bigcup_{j=1}^{2^{n-1}} \big(y_j + \varepsilon B_Y\big)\Big\}.$$

The dyadic entropy numbers are almost s-numbers and hence satisfy the following properties that we will use frequently in this text without always referring to them:

1. $\|T\| \geq e_1(T) \geq e_2(T) \geq \cdots \geq 0$ for $T \in \mathcal{L}(X, Y)$.

2. $e_n(S + T) \leq e_n(S) + \|T\|$ for all $n \in \mathbb{N}$ and $S, T \in \mathcal{L}(X, Y)$.

3. $e_n(RTS) \leq \|R\|\, e_n(T)\, \|S\|$ for all $n \in \mathbb{N}$, $S \in \mathcal{L}(W, X)$, $T \in \mathcal{L}(X, Y)$, and $R \in \mathcal{L}(Y, Z)$.

4. $e_{n+m-1}(S + T) \leq C_Y\big(e_n(S) + e_m(T)\big)$ for all $n, m \in \mathbb{N}$ and $S, T \in \mathcal{L}(X, Y)$.

5. $e_{n+m-1}(ST) \leq e_n(S) \cdot e_m(T)$ for all $n, m \in \mathbb{N}$ and $T \in \mathcal{L}(X, Y)$, $S \in \mathcal{L}(Y, Z)$.

The reader may find the proofs of these facts, for instance, in [7] or [10].

The following result is an analogue of [19, Proposition 12.1.13] for entropy numbers in the quasi-Banach space case.

Lemma 2.4 ([13], Lemma 2.1). *Let $0 < p \leq 1$, $m \in \mathbb{N}$, and let X be a real m-dimensional p-Banach space. Then, for all $n \in \mathbb{N}$,*

$$e_n(X \hookrightarrow X) \leq 4^{1/p} \cdot 2^{-\frac{n-1}{m}}. \tag{2.3}$$

We will need Carl's inequality for the case of quasi-Banach spaces, which was recently proved in [13]. Before we state the result recall that if X, Y are quasi-Banach spaces and $T \in \mathcal{L}(X, Y)$, then the Gelfand numbers $c_n(T)$, $n \in \mathbb{N}$ are defined as

$$c_n(T) := \inf_{\substack{M \subseteq X \\ \mathrm{codim}(M) < n}} \ \sup_{\substack{x \in M \\ \|x\|_X \leq 1}} \ \|Tx\|_Y.$$

Lemma 2.5 ([13], Theorem 1.1). *Let $p, \alpha > 0$. Then there exists a constant $\gamma_{\alpha,p} \in (0, \infty)$ such that, for all p-Banach spaces X, Y and any $T \in \mathcal{L}(X, Y)$,*

$$\sup_{1 \leq k \leq n} k^\alpha e_k(T) \leq \gamma_{\alpha,p} \sup_{1 \leq k \leq n} k^\alpha c_k(T).$$

We will also essentially need the following recent result on Gelfand numbers of identities between Schatten classes (cf. also [6]).

Lemma 2.6 ([8], Lemma 2.1). *Let $0 < p \leq 1$, $p \leq q \leq 2$. Then, for all $n, N \in \mathbb{N}$ with $1 \leq n \leq N^2$,*

$$c_n\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \asymp_{p,q} \min\Big\{1, \frac{N}{n}\Big\}^{1/p - 1/q}. \tag{2.4}$$

Another type of s-numbers we will use are the Kolmogorov numbers. Given two quasi-Banach spaces X, Y and $T \in \mathcal{L}(X, Y)$, the Kolmogorov numbers d_n, $n \in \mathbb{N}$ are defined as

$$d_n(T) := \inf_{\substack{N \subseteq Y \\ \dim(N) < n}} \ \sup_{\|x\|_X \leq 1} \ \inf_{z \in N} \|Tx - z\|_Y.$$

2.7 Interpolation properties of entropy numbers

To prove the estimates in Theorem 1.1, we need certain interpolation properties of the entropy numbers of a bounded linear map acting between quasi-Banach spaces, where one end point is fixed (see, for instance, [10, Chapter 1.3]). Recall that a couple of quasi-Banach spaces (X_0, X_1) is called an interpolation couple if both X_0 and X_1 are continuously embedded in a suitable quasi-Banach space.

Lemma 2.7 ([10], Theorem 1.3.2 (i)). *Let X be a quasi-Banach space and (Y_0, Y_1) be an interpolation couple of p-Banach spaces. Let $0 < \theta < 1$ and Y_θ be a quasi-Banach space such that $Y_0 \cap Y_1 \subseteq Y_\theta \subseteq Y_0 + Y_1$ and*

$$\|y\|_{Y_\theta} \le \|y\|_{Y_0}^{1-\theta} \cdot \|y\|_{Y_1}^\theta, \qquad \text{for all } y \in Y_0 \cap Y_1.$$

Let $T \in \mathcal{L}(X, Y_0 \cap Y_1)$, $Y_0 \cap Y_1$ being equipped with the quasi-norm $\max\{\|y\|_{Y_0}, \|y\|_{Y_1}\}$. Then, for all $k_0, k_1 \in \mathbb{N}$,

$$e_{k_0+k_1-1}(T : X \to Y_\theta) \le 2^{1/p} \cdot e_{k_0}^{1-\theta}(T : X \to Y_0) \cdot e_{k_1}^\theta(T : X \to Y_1).$$

Since Schatten p-norms satisfy Hölder's inequality, we immediately obtain the following corollary.

Corollary 2.8. *Let $N \in \mathbb{N}$, $0 < q_0, q_1, q \le \infty$ and $0 < \theta < 1$ be such that $\frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}$. Then, for all $k_0, k_1 \in \mathbb{N}$,*

$$e_{k_0+k_1-1}(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N) \le 2^{1/\min\{q_0,q_1,1\}} \cdot e_{k_0}^{1-\theta}(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_{q_0}^N) \cdot e_{k_1}^\theta(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_{q_1}^N).$$

To formulate the analogue result for interpolation in the domain space, we recall that the K-functional for an interpolation couple (X_0, X_1) is given by

$$K(t, x) = \inf\{\|x_0\|_{X_0} + t\|x_1\|_{X_1} : x = x_0 + x_1, \ x_0 \in X_0, \ x_1 \in X_1\}.$$

Lemma 2.9 ([10], Theorem 1.3.2 (ii)). *Let (X_0, X_1) be an interpolation couple of quasi-Banach spaces and let Y be a p-Banach space. Let $0 < \theta < 1$ and let X be a quasi-Banach space such that $X \subseteq X_0 + X_1$ and*

$$t^{-\theta} K(t, x) \le \|x\|_X, \qquad \text{for all } x \in X, \ t \in (0, \infty).$$

Let $T : X_0 + X_1 \to Y$ be linear and such that its restrictions to X_0 and X_1 are continuous. Then its restriction to X is also continuous and, for all $k_0, k_1 \in \mathbb{N}$,

$$e_{k_0+k_1-1}(T : X \to Y) \le 2^{1/p} \cdot e_{k_0}^{1-\theta}(T : X_0 \to Y) \cdot e_{k_1}^\theta(T : X_1 \to Y).$$

Using the singular value decomposition and the corresponding K-functional estimate for ℓ_p^N-norms it is easy to see that, for $X_i = \mathcal{S}_{p_i}^N$ with $i \in \{0,1\}$ and $X = \mathcal{S}_p^N$ with $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$, the assumption of Lemma 2.9 is satisfied up to a constant only depending on the parameters p_0, p_1, θ.

Corollary 2.10. *Let $N \in \mathbb{N}$, $0 < p_0, p_1, p \le \infty$ and $0 < \theta < 1$ be such that $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$. Then, for all $k_0, k_1 \in \mathbb{N}$,*

$$e_{k_0+k_1-1}(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N) \le c_{p_0,p_1,\theta,q} \cdot e_{k_0}^{1-\theta}(\mathcal{S}_{p_0}^N \hookrightarrow \mathcal{S}_q^N) \cdot e_{k_1}^\theta(\mathcal{S}_{p_1}^N \hookrightarrow \mathcal{S}_q^N).$$

Lemma 2.7 and Lemma 2.9 show that entropy numbers behave well with respect to interpolation in the domain as well as in the target space. On the other hand, by the famous result of Edmunds and Netrusov [9], this is not the case when interpolating simultaneously in the domain and in the target space. We mention also that interpolation properties for Schatten classes in the Banach space case were already established in [18]. For generalizations to non-commutative L_p-spaces, see also [20].

3 Proof of Theorem 1.1

In this section, we present the proof of Theorem 1.1. We subdivide it into several cases, which are presented in the subsections below. These cases cover all the upper and lower bounds in Theorem 1.1.

3.1 Upper bound in the case $0 < q \leq p \leq \infty$

Here we use factorization and the entropy estimate for finite dimensional spaces given in Lemma 2.4. First, Lemma 2.4 shows that for all $n \in N$ and all $p > 0$

$$e_n\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_p^N\big) \leq C_p 2^{-\frac{n-1}{N^2}} = C_p 2^{-\frac{n}{N^2}} 2^{\frac{1}{N^2}} \leq 2C_p 2^{-\frac{n}{N^2}},$$

with $C_p = 4^{1/\min\{1,p\}}$. This is already the required bound for $p = q$. Now we use the factorization $\mathcal{S}_p^N \hookrightarrow \mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N$ and obtain

$$e_n\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \leq e_n\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_p^N\big) \cdot \|\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\| \leq 2C_p 2^{-\frac{n}{N^2}} N^{1/q-1/p}.$$

This shows, up to constant, the optimal upper bound in the full range.

3.2 Lower bound in the cases $0 < q \leq p \leq \infty$ and $0 < p \leq q \leq \infty, N^2 \leq n$

These bounds follow from a volume comparison argument. Let us assume that B_p^N is covered by 2^{n-1} translates of the ball $\varepsilon \cdot B_q^N$, $\varepsilon > 0$. Then

$$\mathrm{vol}(B_p^N) \leq 2^{n-1}\mathrm{vol}(\varepsilon \cdot B_q^N) = 2^{n-1}\varepsilon^{N^2}\mathrm{vol}(B_q^N)$$

and the volume estimates for the unit balls of Schatten classes in Lemma 2.1 imply

$$c_{p,q} N^{1/q-1/p} \leq \left(\frac{\mathrm{vol}(B_p^N)}{\mathrm{vol}(B_q^N)}\right)^{1/N^2} \leq 2^{(n-1)/N^2}\varepsilon.$$

Hence, by the very definition of the entropy numbers,

$$e_n\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \geq c_{p,q} 2^{-\frac{n-1}{N^2}} N^{1/q-1/p} \geq c_{p,q} 2^{-\frac{n}{N^2}} N^{1/q-1/p}.$$

This establishes the lower bound in the case $0 < q \leq p \leq \infty$ for the full range of $n, N \in \mathbb{N}$ and the estimate is also of the right order when $0 < p \leq q \leq \infty$ and $N^2 \leq n$.

3.3 Upper bound in the case $0 < p \leq q \leq \infty, N^2 \leq n$

The proof for this bound is again based on a standard volume argument. In the following, let $\bar{r} := \min\{1, r\}$ for some $0 < r \leq \infty$. We define

$$\varepsilon = c_{p,q} 2^{-\frac{n}{N^2}} N^{1/q-1/p},$$

where $c_{p,q} > 0$ is some constant to be specified later. We choose a maximal ε-net $\mathcal{N} \subseteq B_p^N$ in the $\|\cdot\|_{\mathcal{S}_q^N}$ sense, i.e., for all $A, B \in \mathcal{N}$ with $A \neq B$, we have $\|A - B\|_{\mathcal{S}_q^N} > \varepsilon$. Then,

$$B_p^N \subseteq \bigcup_{A \in \mathcal{N}} \big(A + \varepsilon B_q^N\big).$$

9

For $n \geq \lambda_{p,q} N^2$, where $\lambda_{p,q} \in (0, \infty)$ is some constant depending only on p and q, we will show that

$$|\mathcal{N}| \leq 2^{n-1}. \tag{3.1}$$

This will imply the desired bound for e_n. First, we observe that the $\bar{q}$-triangle inequality implies that the balls

$$A + \frac{\varepsilon}{2^{1/\bar{q}}} B_q^N, \quad A \in \mathcal{N}$$

are disjoint. Since $B_q^N \subseteq N^{1/p-1/q} B_p^N$, it holds for all $A \in \mathcal{N}$ that

$$A + \frac{\varepsilon}{2^{1/\bar{q}}} B_q^N \subseteq B_p^N + \frac{\varepsilon}{2^{1/\bar{q}}} B_q^N \subseteq B_p^N + \frac{\varepsilon N^{1/p-1/q}}{2^{1/\bar{q}}} B_p^N$$

$$= B_p^N + \frac{c_{p,q} 2^{-\frac{n}{N^2}}}{2^{1/\bar{q}}} B_p^N \subseteq \gamma \cdot B_p^N,$$

where, in the last inclusion, we used the $\bar{p}$-triangle inequality, and defined

$$\gamma := \gamma(p, q, n, N) := \left(1 + c_{p,q}^{\bar{p}} 2^{(-\frac{n}{N^2} - \frac{1}{\bar{q}})\bar{p}}\right)^{1/\bar{p}}.$$

By volume comparison, we conclude

$$|\mathcal{N}| \cdot \mathrm{vol}\left(\frac{\varepsilon}{2^{1/\bar{q}}} B_q^N\right) \leq \mathrm{vol}\left(\gamma B_p^N\right).$$

Taking the N^2-root, using Lemma 2.1 and the definitions of ε and γ, we obtain

$$|\mathcal{N}|^{1/N^2} \leq \frac{\alpha_{p,q}}{c_{p,q}} \left(2^{(1/\bar{q}+n/N^2)\bar{p}} + c_{p,q}^{\bar{p}}\right)^{1/\bar{p}},$$

where $\alpha_{p,q} \in (0, \infty)$ depends on the constants in the volume estimates in Lemma 2.1. To show (3.1), we have to choose $c_{p,q}, \lambda_{p,q} \in (0, \infty)$ such that

$$\frac{\alpha_{p,q}}{c_{p,q}} \left(2^{(1/\bar{q}+n/N^2)\bar{p}} + c_{p,q}^{\bar{p}}\right)^{1/\bar{p}} \leq 2^{\frac{n-1}{N^2}},$$

which is equivalent to

$$\alpha_{p,q}^{\bar{p}} 2^{\bar{p}/\bar{q}} \leq c_{p,q}^{\bar{p}} \left(2^{-\bar{p}/N^2} - \alpha_{p,q}^{\bar{p}} 2^{-\bar{p}n/N^2}\right). \tag{3.2}$$

Obviously, $2^{-\bar{p}/N^2} \geq 2^{-\bar{p}}$ and, for suitable $\lambda_{p,q} \in (0, \infty)$, we have $\alpha_{p,q}^{\bar{p}} 2^{-\bar{p}n/N^2} \leq 4^{-\bar{p}}$. Therefore, we can choose a suitably large constant $c_{p,q} \in (0, \infty)$ such that the inequality (3.2) is satisfied. This completes the proof for $n \geq \lambda_{p,q} N^2$. To close the gap left, observe that, for $N^2 \leq n \leq \lambda_{p,q} N^2$, the monotonicity of the entropy numbers and the upper estimate for $e_{N^2}(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N)$ (see next subsection), gives

$$e_n(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N) \leq e_{N^2}(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N) \leq c_{p,q}' N^{1/q-1/p} \leq c_{p,q}'' 2^{-n/N^2} N^{1/q-1/p},$$

where $c_{p,q}', c_{p,q}'' \in (0, \infty)$.

10

3.4 Upper bound in the case $0 < p \le q \le \infty, 1 \le n \le N^2$

For $1 \le n \le 2N$, this bound follows trivially from

$$e_n\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \le \|\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\| = 1.$$

For $2N \le n \le N^2$ we resort to the Gelfand number estimate in Lemma 2.6, the Carl inequality in Lemma 2.5 to transfer this into an entropy estimate and, finally, to interpolation to cover the remaining cases for p and q.

We will show that, for all $N \le n \le N^2$,

$$e_{2n}\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \le c_{p,q}\Big(\frac{N}{n}\Big)^{1/p-1/q}.$$

The gaps for odd n can be filled as in the last subsection, by suitably increasing the constant. To prove the desired estimate in the range $N \le n \le N^2$, we consider four different cases.

Case $0 < p \le 1$ and $p \le q \le 2$: Recall that, by Lemma 2.6,

$$c_n(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N) \le c_{p,q}\Big(\frac{N}{n}\Big)^{1/p-1/q}.$$

Therefore, by Carl's inequality (see Lemma 2.5) used with $\alpha = 1/p-1/q$ there, we obtain

$$n^{1/p-1/q}e_n\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \le c_{p,q} \sup_{k \le n} \min\{k, N\}^{1/p-1/q} = c_{p,q}\,N^{1/p-1/q}.$$

Hence,

$$e_n\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \le c_{p,q}\Big(\frac{N}{n}\Big)^{1/p-1/q}. \tag{3.3}$$

Case $0 < p \le 1$ and $q = \infty$: Factorizing $\mathcal{S}_p^N \hookrightarrow \mathcal{S}_\infty^N$ through $\mathcal{S}_2^N$, we obtain

$$e_{2n}\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_\infty^N\big) \le e_n\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_2^N\big) \cdot e_n\big(\mathcal{S}_2^N \hookrightarrow \mathcal{S}_\infty^N\big).$$

From (3.3), we get

$$e_{2n}\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_\infty^N\big) \le c_p\Big(\frac{N}{n}\Big)^{1/p-1/2} \cdot e_n\big(\mathcal{S}_2^N \hookrightarrow \mathcal{S}_\infty^N\big). \tag{3.4}$$

To estimate the second factor in (3.4), we use the duality of Gelfand and Kolmogorov numbers and again Lemma 2.6. This gives

$$d_n\big(\mathcal{S}_2^N \hookrightarrow \mathcal{S}_\infty^N\big) = c_n\big(\mathcal{S}_1^N \hookrightarrow \mathcal{S}_2^N\big) \le c\Big(\frac{N}{n}\Big)^{1/2}.$$

Again, using Carl's inequality with Kolmogorov numbers, we obtain

$$e_n\big(\mathcal{S}_2^N \hookrightarrow \mathcal{S}_\infty^N\big) \le c\Big(\frac{N}{n}\Big)^{1/2}.$$

Together with the estimate (3.4), this yields

$$e_{2n}\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_\infty^N\big) \le c_p\Big(\frac{N}{n}\Big)^{1/p}. \tag{3.5}$$

11

This completes the proof in this case.

Case $0 < p \le 1$ and $p < q \le \infty$: We now use interpolation in the target space (see Corollary 2.8) and obtain the estimate

$$e_{2n}\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \le 2^{1/\bar{q}}\|\mathcal{S}_p^N \hookrightarrow \mathcal{S}_p^N\|^{1-\theta} \cdot e_{2n}\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_\infty^N\big)^\theta \le 2^{1/\bar{q}}e_{2n}\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_\infty^N\big)^\theta,$$

with $\frac{1-\theta}{p} + \frac{\theta}{\infty} = \frac{1}{q}$. Using the upper bound in (3.5), we have

$$e_{2n}\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \le c_{p,q}\Big(\frac{N}{n}\Big)^{\theta/p} = c_{p,q}\Big(\frac{N}{n}\Big)^{1/p-1/q}. \tag{3.6}$$

Case $1 < p < q \le \infty$: Here we use interpolation in the domain space (see Corollary 2.10. Then,

$$\begin{aligned}
e_{2n}\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) &\le c_{p,q} \cdot e_{2n}\big(\mathcal{S}_1^N \hookrightarrow \mathcal{S}_q^N\big)^{1-\theta}\|\mathcal{S}_q^N \hookrightarrow \mathcal{S}_q^N\|^\theta \\
&\le c_{p,q} \cdot e_{2n}\big(\mathcal{S}_1^N \hookrightarrow \mathcal{S}_q^N\big)^{1-\theta},
\end{aligned} \tag{3.7}$$

where $\frac{1-\theta}{1} + \frac{\theta}{q} = \frac{1}{p}$ and with $c_{p,q} \in (0,\infty)$ being the constant from Corollary 2.10. Using the estimate obtained in the previous case (see inequality (3.6)) with $p = 1$ there, we get

$$e_{2n}\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \le c_{p,q}\Big(\frac{N}{n}\Big)^{(1-1/q)(1-\theta)} = c_{p,q}\Big(\frac{N}{n}\Big)^{1/p-1/q}.$$

This completes the proof for all $0 < p \le q \le \infty$.

3.5 Lower bound in the case $0 < p \le q \le \infty, 1 \le n \le N^2$

Note that the lower bound for $1 \le n \le N$ trivially follows if the lower bound for $N \le n \le N^2$ is proved. So we assume $N \le n \le N^2$.

Let $1 \le k \le N/2$, $q \ge 1$ and $0 < \delta < k^{1/q}$. It follows from Lemma 2.3 that, for all $F \in G_{N,k}$,

$$\mu_{N,k}\big(B(F,\delta)\big) \le \Big(\frac{\delta}{ck^{1/q}}\Big)^{k(N-k)}$$

for some absolute constant $c \in (0,\infty)$, where $B(F,\delta) := \{E \in G_{N,k} : \|P_E - P_F\|_{\mathcal{S}_q^N} < \delta\}$. We may assume that $c < 4$ and choose $\delta = \frac{ck^{1/q}}{4} < k^{1/q}$. We conclude that

$$\mu_{N,k}\big(B(F,\delta)\big) \le 4^{-k(N-k)} = 2^{-2k(N-k)} \le 2^{-kN}.$$

It follows that there exist a set $\mathcal{N} \subseteq G_{N,k}$ with cardinality 2^{kN} such that $\|P_E - P_F\|_{\mathcal{S}_q^N} \ge \delta$ or, equivalently,

$$\|k^{-1/p}P_E - k^{-1/p}P_F\|_{\mathcal{S}_q^N} \ge k^{-1/p}\delta = \frac{ck^{1/q-1/p}}{4},$$

for distinct $E, F \in \mathcal{N}$. Since $k^{-1/p}P_E \in B_p^N$ for $E \in G_{N,k}$, we obtain

$$e_{kN}\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \ge \frac{ck^{1/q-1/p}}{8}.$$

This shows the required lower bound in the case $n = kN$ and $0 < p \le q \le \infty$ and $q \ge 1$. For $(k-1)N \le n \le kN$, the lower bound follows from the monotonicity of entropy

12

numbers. The same is true for $N^2/2 \leq n \leq N^2$ taking into account the already proved lower bound for $n = N^2$.

The case $0 < p \leq q < 1$ follows by factorization,

$$e_{2n}\left(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_1^N\right) \leq e_n\left(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\right) \cdot e_n\left(\mathcal{S}_q^N \hookrightarrow \mathcal{S}_1^N\right).$$

Using the bounds we have already obtained (upper and lower), we derive the lower bound for $e_n\left(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\right)$.

4   Construction of Nets

Let $K, N \in \mathbb{N}$ with $K \leq N$. We denote by V_K^N the Stiefel manifold of $N \times K$ real matrices with orthonormal columns, i.e., if $\mathrm{id}_{\mathbb{R}^K}$ is the identity matrix in $\mathbb{R}^{K \times K}$, then

$$V_K^N = \left\{U \in \mathbb{R}^{N \times K} : U^T U = \mathrm{id}_{\mathbb{R}^K}\right\}. \tag{4.1}$$

We will denote the dimension of the Stiefel manifold V_K^N by $d_{N,K} = K(N - (K+1)/2)$. Furthermore, if $M \subseteq X$ is a subset of a quasi-Banach space $(X, \|\cdot\|_X)$ and $\varepsilon > 0$ is a real number, we denote by $N(M, X, \varepsilon)$ (or $N(M, X, \varepsilon)$ if we want to specify the norm) the covering number of M in X, i.e., the minimal number of (closed) ε-balls in X covering M. Moreover, $\|U\|_{op}$ stands for the operator norm of a rectangular matrix $U \in \mathbb{R}^{N \times K}$. If $K = N$, then $\|U\|_{op} = \|U\|_{\mathcal{S}_\infty^N}$. The following lemma provides an upper bound on the covering number of the Stiefel manifold in $\mathbb{R}^{N \times K}$ with respect to the operator norm.

Lemma 4.1. *Let $0 < \varepsilon < 1$ and let $K, N \in \mathbb{N}$ with $K \leq N$. Then*

$$N\left(V_K^N, \|\cdot\|_{op}, \varepsilon\right) \leq \left(\frac{c}{\varepsilon}\right)^{d_{N,K}},$$

where $c \in (0, \infty)$ is a universal constant.

Proof. Let $\mathcal{N}_G \subseteq G_{N,K}$ be an $\varepsilon > 0$ net of $G_{N,K}$ in the operator norm. Then, to every $F \in G_{N,K}$ there is $E \in \mathcal{N}_G$ with $\|P_E - P_F\|_{op} < \varepsilon$. For every $E \in \mathcal{N}_G$ let us fix an orthonormal basis $(u_1^E, \ldots, u_K^E)$ of E and denote by U_E the $N \times K$ matrix with columns $u_1^E, \ldots, u_K^E$. In that notation, we can write $P_E = U_E U_E^T$.

We denote by $\mathcal{O}(K)$ the group of $K \times K$ orthogonal matrices and let $\mathcal{N}_\mathcal{O} \subseteq \mathcal{O}(K)$ be an $\varepsilon > 0$ net of $\mathcal{O}(K)$ in the operator norm. This means that to every $W \in \mathcal{O}(K)$ there is $Z \in \mathcal{N}_\mathcal{O}$ with $\|W - Z\|_{op} < \varepsilon$. We put

$$\mathcal{N} = \{U_E Z : E \in \mathcal{N}_G, Z \in \mathcal{N}_\mathcal{O}\} \subseteq \mathbb{R}^{N \times K}. \tag{4.2}$$

First, let us observe that for any $U_E Z \in \mathcal{N}$,

$$(U_E Z)^T (U_E Z) = Z^T U_E^T U_E Z = Z^T \, \mathrm{id}_{\mathbb{R}^K} \, Z = Z^T Z = \mathrm{id}_{\mathbb{R}^K},$$

which implies that $\mathcal{N} \subseteq V_K^N$. We will now show that $\mathcal{N}$ is a 3ε-net of V_K^N in the operator norm.

Indeed, let $U \in V_K^N$ be arbitrary and let us denote by F its column space. Then there is an $E \in \mathcal{N}_G$ with $\|P_E - P_F\|_{op} < \varepsilon$. Also, we can write

$$U = P_F U = (P_F - P_E)U + P_E U = (P_F - P_E)U + U_E U_E^T U.$$

13

It is now left to show that there is a $Z \in \mathcal{N}_\mathcal{O}$ with $\|U_E^T U - Z\|_{op} < 2\varepsilon$, because then we obtain

$$\begin{aligned}
\|U - U_E Z\|_{op} &= \|(P_F - P_E)U + U_E U_E^T U - U_E Z\|_{op} \\
&\leq \|(P_F - P_E)U\|_{op} + \|U_E(U_E^T U - Z)\|_{op} \\
&\leq \|P_F - P_E\|_{op} \cdot \|U\|_{op} + \|U_E\|_{op} \cdot \|U_E^T U - Z\|_{op} < 3\varepsilon.
\end{aligned}$$

We show the existence of $Z \in \mathcal{N}_\mathcal{O}$ with $\|U_E^T U - Z\|_{op} < 2\varepsilon$. For that sake, it is sufficient to prove that $\mathrm{d}(U_E^T U, \mathcal{O}(K), \|\cdot\|_{op}) < \varepsilon$. By considering the singular value decomposition of $U_E^T U$, it is even enough to show that the singular values of $U_E^T U$ lie in the interval $(1 - \varepsilon, 1 + \varepsilon)$. Since we trivially have $\|U_E^T U\|_{op} \leq \|U_E^T\|_{op} \cdot \|U\|_{op} \leq 1$, we only need to show that the singular values are larger than $1 - \varepsilon$. Observe that

$$\begin{aligned}
\{\|U_E^T U x\|_2 \ : \ x \in \mathbb{R}^K, \ \|x\|_2 = 1\} &= \{\|U_E^T y\|_2 \ : \ y \in F, \ \|y\|_2 = 1\} \\
&= \{\|P_E y\|_2 \ : \ y \in F, \ \|y\|_2 = 1\},
\end{aligned}$$

and that for $y \in F$ with $\|y\|_2 = 1$, we have

$$\begin{aligned}
\|P_E y\|_2 = \|P_F y + (P_E - P_F)y\|_2 &\geq \|P_F y\|_2 - \|(P_E - P_F)y\|_2 \\
&= \|y\|_2 - \|(P_E - P_F)y\|_2 > 1 - \varepsilon.
\end{aligned}$$

This shows that $\mathrm{d}(U_E^T U, \mathcal{O}(K), \|\cdot\|_{op}) < \varepsilon$.

We now use the estimates of Szarek (see [25, Prop. 6 and Prop. 8]) and the identity

$$\begin{aligned}
\dim G_{N,K} + \dim \mathcal{O}(K) &= K(N - K) + \frac{K(K-1)}{2} \\
&= KN - \frac{K(K+1)}{2} = \dim V_K^N = d_{N,K}
\end{aligned}$$

to show that

$$|\mathcal{N}| \leq |\mathcal{N}_G| \cdot |\mathcal{N}_\mathcal{O}| \leq \left(\frac{c}{\varepsilon}\right)^{\dim G_{N,K}} \cdot \left(\frac{c}{\varepsilon}\right)^{\dim \mathcal{O}(K)} = \left(\frac{c}{\varepsilon}\right)^{d_{N,K}}.$$

$\square$

The construction that we will present in the proof of Theorem 4.3 is crucially based on a decomposition of a matrix into a sum of low-rank matrices. Therefore, let us define

$$R_{K,q}^N = \{X \in \mathbb{R}^{N \times N} \ : \ \mathrm{rank}(X) \leq K, \ \|X\|_{\mathcal{S}_q^N} \leq 1\}. \tag{4.3}$$

The next lemma estimates the covering numbers of $R_{K,q}^N$ in the Schatten q-classes $\mathcal{S}_q^N$.

Lemma 4.2. *Let $0 < q \leq \infty$. Then there exists $c_q \in [1, \infty)$ such that for every $0 < \varepsilon \leq c_q$ and all $K, N \in \mathbb{N}$ with $K \leq N$, it holds*

$$N\left(R_{K,q}^N, \mathcal{S}_q^N, \varepsilon\right) \leq \left(\frac{c_q}{\varepsilon}\right)^{2d_{N,K}+K} = \left(\frac{c_q}{\varepsilon}\right)^{K(2N-K)}. \tag{4.4}$$

Proof. We write again $\bar{q} = \min(1, q)$. Moreover, let us assume that

$$\mathcal{N}_{N,K} \subseteq V_K^N$$

14

is an $\varepsilon/3^{1/\bar{q}}$-net of V_K^N in the operator norm and that

$$\mathcal{N}_q \subseteq B_{\ell_q^K}$$

be an $\varepsilon/3^{1/\bar{q}}$-net of the unit ball $B_{\ell_q^K}$ in the quasi-Banach space ℓ_q^K. We now define

$$\mathcal{N}_{K,q}^N = \left\{ \widetilde{U}\widetilde{\Sigma}\widetilde{V}^T \; : \; \widetilde{U}, \widetilde{V} \in \mathcal{N}_{N,K}, \; \widetilde{\Sigma} = \mathrm{diag}(\widetilde{\sigma}), \; \widetilde{\sigma} \in \mathcal{N}_q \right\}.$$

Let us consider $X \in R_{K,q}^N$ with $X = U\Sigma V^T$, where $U, V \in V_K^N$ and $\Sigma = \mathrm{diag}(\sigma)$ with $\|\sigma\|_{\ell_q^K} \leq 1$. Moreover, consider $Y = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^T$, where $\widetilde{U}, \widetilde{V} \in \mathcal{N}_{N,K}$ such that $\|U - \widetilde{U}\|_{op} < \varepsilon/3^{1/\bar{q}}$, $\|V - \widetilde{V}\|_{op} < \varepsilon/3^{1/\bar{q}}$ and where $\widetilde{\sigma} \in \mathcal{N}_q$ with $\|\sigma - \widetilde{\sigma}\|_{\ell_q^K} < \varepsilon/3^{1/\bar{q}}$. Then, we have $Y \in \mathcal{N}_{K,q}^N$ and obtain the estimate

$$
\begin{aligned}
\|X - Y\|_{\mathcal{S}_q^N}^{\bar{q}} &= \|U\Sigma V^T - \widetilde{U}\widetilde{\Sigma}\widetilde{V}^T\|_{\mathcal{S}_q^N}^{\bar{q}} \\
&\leq \|U\Sigma(V^T - \widetilde{V}^T)\|_{\mathcal{S}_q^N}^{\bar{q}} + \|U(\Sigma - \widetilde{\Sigma})\widetilde{V}^T\|_{\mathcal{S}_q^N}^{\bar{q}} + \|(U - \widetilde{U})\widetilde{\Sigma}\widetilde{V}^T\|_{\mathcal{S}_q^N}^{\bar{q}} \\
&= \|\Sigma(V^T - \widetilde{V}^T)\|_{\mathcal{S}_q^N}^{\bar{q}} + \|\Sigma - \widetilde{\Sigma}\|_{\mathcal{S}_q^N}^{\bar{q}} + \|(U - \widetilde{U})\widetilde{\Sigma}\|_{\mathcal{S}_q^N}^{\bar{q}} \\
&\leq \|\Sigma\|_{\mathcal{S}_q^N}^{\bar{q}} \cdot \|V^T - \widetilde{V}^T\|_{op}^{\bar{q}} + \|\sigma - \widetilde{\sigma}\|_{\ell_q^N}^{\bar{q}} + \|U - \widetilde{U}\|_{op}^{\bar{q}} \cdot \|\widetilde{\Sigma}\|_{\mathcal{S}_q^N}^{\bar{q}} \\
&< \varepsilon^{\bar{q}}/3 + \varepsilon^{\bar{q}}/3 + \varepsilon^{\bar{q}}/3 = \varepsilon^{\bar{q}}.
\end{aligned}
$$

Here, we have used the inequality $\|AB\|_{\mathcal{S}_q^N} \leq \|A\|_{op} \cdot \|B\|_{\mathcal{S}_q^N}$ (see, for instance, [3, Proposition IV.2.4]). Finally, we use Lemmas 4.2 and 2.4 to estimate the number of elements in $\mathcal{N}_{K,q}^N$. We obtain

$$|\mathcal{N}_{K,q}^N| \leq |\mathcal{N}_{N,K}| \cdot |\mathcal{N}_q| \cdot |\mathcal{N}_{N,K}| \leq \left(\frac{c_q}{\varepsilon}\right)^{2d_{N,K}+K} = \left(\frac{c_q}{\varepsilon}\right)^{K(2N-K)}.$$

$\square$

In the next theorem, we present a constructive proof for the upper bound of the entropy numbers of natural identities of Schatten classes in the range $N \leq n \leq N^2$. Our construction is crucially based on a low-rank matrix decomposition.

Theorem 4.3. *Let $0 < p \leq q \leq \infty$ and $N \leq n \leq N^2$. Then*

$$e_n\big(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N\big) \leq C_{p,q}\Big(\frac{N}{n}\Big)^{1/p-1/q}, \tag{4.5}$$

where $C_{p,q} \in (0, \infty)$ depends only on p and q.

Proof. We subdivide the proof into three steps.
Step 1: By the monotonicity of entropy numbers, it is enough to consider $N = 2^\nu$ and $n = 2^\ell \cdot 2^\nu$, where $1 \leq \ell \leq \nu$ are natural numbers. In that case, $N/n = 2^{-\ell}$.

Let now $A \in \mathcal{S}_p^N$ be an arbitrary matrix with $\|A\|_{\mathcal{S}_p^N} \leq 1$. We use the singular value decomposition of A in the form $A = U\Sigma V^T$ to decompose it into a sum of low-rank matrices. If $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_N)$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_N \geq 0$ is the diagonal matrix collecting the singular values of A on its diagonal, we define

$$A_1 := U\Sigma_1 V^T, \qquad \Sigma_1 := \mathrm{diag}(\sigma_1, 0, \ldots, 0)$$

15

and, for $j = 2, \ldots, \ell$,

$$A_j := U\Sigma_j V^T, \qquad \Sigma_j := \mathrm{diag}(0, \ldots, 0, \sigma_{2^{j-1}}, \ldots, \sigma_{2^j - 1}, 0, \ldots, 0).$$

That way, we obtain $\mathrm{rank}(A_j) \leq 2^{j-1}$ for $j = 1, \ldots, \ell$. Furthermore, $\|A_1\|_{\mathcal{S}_q^N} \leq 1$ and, for all $j = 2, \ldots, \ell$, we have

$$\|A_j\|_{\mathcal{S}_q^N} = \left(\sum_{u=2^{j-1}}^{2^j - 1} \sigma_u^q \right)^{1/q} \leq \left(2^{j-1} \sigma_{2^{j-1}}^q \right)^{1/q} = 2^{(j-1)/q} \left(\sigma_{2^{j-1}}^p \right)^{1/p}$$

$$\leq 2^{(j-1)/q} \left(\frac{1}{2^{j-1}} \sum_{u=1}^{2^{j-1}} \sigma_u^p \right)^{1/p} \leq 2^{(j-1)(1/q - 1/p)}.$$

We can therefore decompose as follows,

$$A = A_1 + A_2 + \cdots + A_\ell + A^c,$$

where $\mathrm{rank}(A_j) \leq 2^{j-1}$ and $\|A_j\|_{\mathcal{S}_q^N} \leq 2^{(j-1)(1/q - 1/p)}$ for all $j = 1, \ldots, \ell$. Finally, we obtain

$$\|A^c\|_{\mathcal{S}_q^N}^q = \sum_{u=2^\ell}^N \sigma_u^q = \sum_{u=2^\ell}^N \sigma_u^p \sigma_u^{q-p} \leq \sigma_{2^\ell}^{q-p} \cdot \sum_{u=2^\ell}^N \sigma_u^p$$

$$\leq \left(\frac{1}{2^\ell} \sum_{u=1}^{2^\ell} \sigma_u^p \right)^{\frac{q-p}{p}} \cdot \sum_{u=2^\ell}^N \sigma_u^p$$

$$\leq 2^{-\ell \cdot \frac{q-p}{p}} \|A\|_{\mathcal{S}_p^n}^{q-p} \cdot \|A\|_{\mathcal{S}_p^N}^p$$

and $\|A^c\|_{\mathcal{S}_q^N} \leq 2^{\ell(1/q - 1/p)}$.

Step 2: We first observe that, for $j = 1, \ldots, \ell$, the low-rank matrix $2^{(j-1)(1/p - 1/q)} A_j$ belongs to $R_{2^{j-1}, q}^N$, where $R_{2^{j-1}, q}^N$ was defined in (4.3). Now let $\mathcal{N}_j \subseteq R_{2^{j-1}, q}^N$, $j = 1, \ldots, \ell$ be an ε_j-net. Then,

$$\mathcal{N} := \left\{ \sum_{j=1}^\ell 2^{-(j-1)(1/p - 1/q)} Z_j \; : \; Z_j \in \mathcal{N}_j \text{ for all } j = 1, \ldots, \ell \right\}$$

is an ε-net of the unit ball of $\mathcal{S}_p^N$ in $\mathcal{S}_q^N$, where

$$\varepsilon^{\bar{q}} = \sum_{j=1}^\ell \left[2^{-(j-1)(1/p - 1/q)} \varepsilon_j \right]^{\bar{q}} + 2^{\ell \bar{q}(1/q - 1/p)}. \tag{4.6}$$

Furthermore, $\mathcal{N}$ has $|\mathcal{N}| = \prod_{j=1}^\ell |\mathcal{N}_j|$ elements.

Step 3: For $j = 1, \ldots, \ell$, we choose $\varepsilon_j = c_q \cdot 2^{(j-\ell)(1/p - 1/q + \alpha)}$, where $c_q \in [1, \infty)$ is the constant from Lemma 4.2 and the value of the parameter $\alpha > 0$ will be specified later.

This gives

$$\varepsilon^{\bar{q}} = \sum_{j=1}^{\ell} \left[2^{-(j-1)(1/p-1/q)}\varepsilon_j\right]^{\bar{q}} + 2^{\ell\bar{q}(1/q-1/p)}$$

$$= \sum_{j=1}^{\ell} \left[c_q 2^{(1-\ell)(1/p-1/q)}2^{(j-\ell)\alpha}\right]^{\bar{q}} + 2^{\ell\bar{q}(1/q-1/p)}$$

$$= c_q^{\bar{q}} 2^{(1-\ell)\bar{q}(1/p-1/q)} \sum_{j=1}^{\ell} 2^{(j-l)\alpha\bar{q}} + 2^{\ell\bar{q}(1/q-1/p)} \qquad (4.7)$$

$$\leq c_{\alpha,q}^{\bar{q}} 2^{-\ell\bar{q}(1/p-1/q)} = c_{\alpha,q}^{\bar{q}}(N/n)^{\bar{q}(1/p-1/q)}.$$

Furthermore, by Lemma 4.2,

$$|\mathcal{N}| = \prod_{j=1}^{\ell} |\mathcal{N}_j| \leq \prod_{j=1}^{\ell} \left(\frac{c_q}{\varepsilon_j}\right)^{2^{j-1}(2N-2^{j-1})} \leq \prod_{j=1}^{\ell} \left(2^{(\ell-j)\cdot(1/p-1/q+\alpha)}\right)^{2^j N}$$

and

$$\log_2(|\mathcal{N}|) = \sum_{j=1}^{\ell} \log_2(|\mathcal{N}_j|) \leq \sum_{j=1}^{\ell} 2^j N \log_2(2^{(\ell-j)\cdot(1/p-1/q+\alpha)})$$

$$= (1/p-1/q+\alpha)N \sum_{j=1}^{\ell} 2^j(\ell-j) = (1/p-1/q+\alpha)N2^{\ell} \sum_{j=1}^{\ell} 2^{j-\ell}(\ell-j) \quad (4.8)$$

$$\leq 2(1/p-1/q+\alpha)2^{\ell} N = 2(1/p-1/q+\alpha)n.$$

We now choose $\alpha = 1$ and $\gamma \in \mathbb{N}$ with $\gamma \geq 1 + 2(1/p - 1/q + 1)$. Combining (4.7) with (4.8), we obtain the upper bound $e_{\gamma n}(\mathcal{S}_p^N \hookrightarrow \mathcal{S}_q^N) \leq c(N/n)^{1/p-1/q}$. $\qquad \square$

References

[1] T. Aoki, *Locally bounded linear topological spaces*, Proc. Imp. Acad. Tokyo 18 (10) (1942), 588–594

[2] J. Arazy, J. Lindenstrauss, *Some linear topological properties of the spaces C_p of operators on Hilbert space*, Composito Math. 30 (1) (1975), 81–111

[3] R. Bhatia, *Matrix analysis*, Graduate Texts in Mathematics, vol. 169. Springer, New York, 1997

[4] E. J. Candès, B. Recht, *Exact matrix completion via convex optimization*, Found. Comput. Math. 9 (6) (2009), 717–772

[5] E. J. Candès, T. Tao, *The power of convex relaxation: near-optimal matrix completion*, IEEE Trans. Inform. Theory 56 (5) (2010), 2053–2080

[6] B. Carl and A. Defant, *Asymptotic estimates for approximation quantities of tensor product identities*, J. Approx. Theory 88 (1997), 228–256

[7] B. Carl and I. Stephani, *Entropy, compactness and the approximation of operators*, Cambridge University Press, 1990

[8] J.A. Chávez-Domínguez and D. Kutzarova, *Stability of low-rank matrix recovery and its connections to Banach space geometry*, J. Math. Anal. Appl. 427 (1) (2015), 320–335

[9] D.E. Edmunds and Yu. Netrusov, *Entropy numbers and interpolation*, Math. Ann. 351 (4) (2011), 963–977

[10] D.E. Edmunds and H. Triebel, *Function spaces, entropy numbers and differential operators*, Cambridge University Press, 1996

[11] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing*, Birkhäuser, Boston, 2013

[12] O. Guédon and A.E. Litvak, *Euclidean projections of a p-convex body, Geometric aspects of functional analysis*, Israel Seminar (GAFA) 1996-2000, V.D. Milman and G. Schechtman, eds., Lecture Notes in Mathematics, 1745, 95–108, Springer, Berlin, 2000

[13] A. Hinrichs, A. Kolleck, and J. Vybíral, *Carl's inequality for quasi-Banach spaces*, J. Funct. Anal. 271 (8) (2016), 2293–2307

[14] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press (1986)

[15] T. Kühn, *A lower estimate for entropy numbers*, J. Approx. Theory 110 (1) (2001), 120–124

[16] C. A. McCarthy, C_p, Israel J. Math. 5 (4) (1967), 249–271

[17] G. Paouris, P. Valettas, *Neighborhoods on the Grassmannian of marginals with bounded isotropic constant*, J. Funct. Anal. 267 (9) (2014), 3427–3443

[18] A. Pietsch, *Interpolationsfunktoren, Folgenideale und Operatorenideale*, Czechoslovak Math. J. 21 (96) (1971), 644–652

[19] A. Pietsch, *Operator ideals*, North-Holland Publishing Co., Amsterdam-New York, 1980

[20] G. Pisier, Q. Xu, *Non-commutative L^p-spaces*, In: Handbook of the geometry of Banach spaces, Vol. 2, North-Holland, Amsterdam, 2003, 1459–1517

[21] S. Rolewicz, *On a certain class of linear metric spaces*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astrono. Phys. 5 (1957), 471–473

[22] J. Saint-Raymond, *Le volume des idéaux d'opérateurs classiques*, Studia Math. 80 (1) (1984), 63–75

[23] C. Schütt, *Entropy numbers of diagonal operators between symmetric Banach spaces*, J. Approx. Theory 40 (2) (1984), 121–128

[24] S. Szarek, *The finite dimensional basis problem with an appendix on nets on Grassmann manifolds*, Acta Math. 151 (1) (1983), 153–179

[25] S. Szarek, *Nets of Grassmann manifold and orthogonal group*, Proceedings of Banach Space Workshop, University of Iowa Press (1982), 169–185

[26] H. Triebel, *Fractals and Spectra*, Birkhäuser, Basel, 1997