Entropy Estimates on Tensor Products of Banach Spaces and Applications to Low-Rank Recovery

Kiryung Lee*, Rakshith Sharma Srinivasa†, Marius Junge‡, and Justin Romberg†
*ECE, The Ohio State University, †ECE, Georgia Tech, ‡Mathematics, Univ. Illinois at Urbana-Champaign
*lee.8763@osu.edu, †{rakshith.sharma.s,jrom}@gatech.edu, ‡mjunge@illinois.edu

***Abstract*—Low-rank matrix models have been universally useful for numerous applications starting from classical system identification to more modern matrix completion in signal processing and statistics. The Schatten-1 norm, also known as the nuclear norm, has been used as a convex surrogate of the low-rankness since it induces a low-rank solution to inverse problems. While the Schatten-1 norm for low-rankness has a nice analogy with the ℓ_1 norm for sparsity through the singular value decomposition, other matrix norms also induce low-rankness. Particularly as one interprets a matrix as a linear operator between Banach spaces, various tensor product norms generalize the role of the Schatten-1 norm. Inspired by a recent work on the max-norm-based matrix completion, we provide a unified view on a class of tensor product norms and their interlacing relations on low-rank operators. Furthermore we derive entropy estimates between the injective and projective tensor products of a family of Banach space pairs and demonstrate their applications to matrix completion and decentralized subspace sketching.**

***Index Terms*—Low-rank matrix, tensor product, Banach spaces, entropy number, matrix completion, sketching**

I. Introduction

Low-rank matrix models play a key role in solving inverse problems in signal processing, statistics, and data science [1]. The effectiveness of these models is often characterized through the geometry of a convex relaxation of the set of rank-constrained matrices. These convex relaxations allow us to reformulate inverse problems as convex optimization programs and derive theoretical guarantees on the accuracy of their solution under certain randomness assumptions on the observation model. The now standard relaxation for the set of rank-constrained matrices specified through the Schatten-1 norm (sum of singular values), also known as the *nuclear norm*. The set of rank-r matrices is included in a nonconvex cone whose members satisfy

$$\|M\|_{\mathrm{HS}} \leqslant \|M\|_{S_1} \leqslant \sqrt{r}\|M\|_{\mathrm{HS}}, \tag{1}$$

where $\|\cdot\|_{S_1}$ is the Schatten-1 norm and $\|\cdot\|_{\mathrm{HS}}$ is the Hilbert-Schmidt or Frobenius norm. The interlacing inequalities in (1) implies that the Schatten-1 norm of a rank-r matrix is equivalent to its spectral norm up to factor $\sqrt{r}$. The dependence on the rank here is sharp.

For inverse problems that involve random projections of low-rank matrices [2] or random quadratic forms [3], the Schatten-1 norm relaxation allows recovery with near-optimal sample complexity without any additional constraints on the matrix M being sensed.

The Schatten-1 norm is not the only relaxation for the set of rank-constrained matrices. For instance, the "max norm" [4] of a matrix M is given by

$$\|M\|_{\max} := \inf\left\{\|U\|_{2\to\infty}\|V\|_{2\to\infty} \mid M = UV^{\mathsf{T}}\right\},$$

where $\|U\|_{2\to\infty}$ is the induced norm $\|U\|_{2\to\infty} = \max_{\|x\|_2 \leqslant 1}\|Ux\|_\infty$. If M is rank-r, then its max norm obeys an analog of (1),

$$\|M\|_\infty \leqslant \|M\|_{\max} \leqslant \sqrt{r}\|M\|_\infty,$$

where $\|M\|_\infty$ is the largest magnitude entry in M. In [5], it is shown that randomly selecting samples provides a stable embedding (see Proposition IV.2 below) of all matrices in the set

$$\left\{M : \max\left(\|M\|_{\max}/\sqrt{r},\, \|M\|_\infty\right) \leqslant \alpha\right\}, \tag{2}$$

and that this stable embedding leads to near-optimal sample complexity guarantees for the matrix completion problem when the underlying matrix does indeed have rank at most r and maximum entry at most α.

Another relaxation is given in [6], where the problem of decentralized subspace sketching is considered. Decentralized subspace sketching estimates a low-dimensional subspace from their compressive linear measurements of sampled vectors. Mathematically this problem is identical to multivariate regression [7]. In [6], the "mixed norm" is introduced,

$$\|M\|_{\mathrm{mixed}} := \inf\left\{\|U\|_{\mathrm{HS}}\|V^{\mathsf{T}}\|_{1\to 2} \mid M = UV^{\mathsf{T}}\right\},$$

where $\|V^{\mathsf{T}}\|_{1\to 2}$ is the induced norm $\|V^{\mathsf{T}}\|_{1\to 2} = \max_{\|x\|_1 = 1}\|V^{\mathsf{T}}x\|_2$. The authors show that sketching each column of a matrix independently provides a stable embedding (see Proposition IV.4 below) of all matrices in the set

$$\left\{M : \max\left(\|M^{\mathsf{T}}\|_{\mathrm{mixed}}/\sqrt{r},\, \|M\|_{1\to 2}\right) \leqslant \alpha\right\}, \tag{3}$$

and that this stable embedding leads to near-optimal sample complexity guarantees for the decentralized sketching problem when the underlying matrix does indeed have rank at most r and α is chosen appropriately.

In this paper, we show how the Schatten-1, max, and mixed norms are all examples of tensor product norms on linear operators that map between pairs of Banach spaces. Then following the properties and relations among tensor product norms in the functional analysis literature (e.g., [8]), we generalize the notions of "operator norm" and "nuclear norm", and arrive at a family of convex relaxations for low-rank matrices that can be chosen to match the structure of the inverse problems for which they are being used. We give bounds on the complexity of the analogs of the sets (2),(3) by providing entropy estimates of the associated operators.

Considering both matrix completion and decentralized subspace sketching as our main examples, we are particularly interested in the tensor product of ℓ_∞^m and ℓ_p^n with $2 \leqslant p \leqslant \infty$. We show the rank-driven interlacing property between tensor norms and derive entropy estimates using Maurey's empirical method [9] through embedding of a Hilbert space to a higher dimensional Banach space with ℓ_1 norm.

The remainder of this paper is organized as follows: We recall relevant definitions and properties of Banach spaces and norms in Section II. The interlacing properties between the injective and projective tensor norms of rank-r operators are derived in Section III. Entropy estimates on the identity map on tensor products with different norms are shown in Section IV, followed by its applications to matrix completion and decentralized subspace sketching in Section V. Finally we finish the paper with concluding remarks and the summary of our extension to more general settings.

II. Preliminaries

We first recall definitions and properties of Banach spaces and norms in the context of the tensor product.

A. Tensor product of Banach spaces

We interpret a matrix $M \in \mathbb{R}^{m\times n}$ as the matrix representation of a linear operator from a vector space of dimension n to another vector space of dimension m. Let X^* denote the Banach space dual of X. The vector space of all linear operators from X^* to another Banach space will be denoted by $L(X^*, Y)$. Then $B(X^*, Y)$ will denote the subset of $L(X^*, Y)$ consisting of linear operators with a finite operator norm. The tensor product of two Banach spaces X and Y, denoted by $X \otimes Y$, is identified to $L(X^*, Y)$ provided that either X or Y is finite dimensional.

B. Tensor product norms

There are many ways to define a norm on the tensor product of two Banach spaces. We are mostly interested in the following two tensor norms.

The injective tensor norm on $X \otimes Y$ is defined by

$$\|T\|_\vee := \sup_{x^*\in B_{X^*}, y^* \in B_{Y^*}} |\langle x^* \otimes y^*, T\rangle|,$$

where B_{X^*} and B_{Y^*} are the unit balls in X^* and Y^* respectively. Then $X \check{\otimes} Y$ denotes the corresponding Banach space. The injective tensor norm of T coincides with the operator norm of $T \in L(X^*, Y)$.

The projective tensor norm on $X \otimes Y$ is defined by

$$\|T\|_\wedge := \inf\left\{ \sum_n \|x_n\|_X \|y_n\|_Y \;\middle|\; T = \sum_n x_n \otimes y_n \right\}$$

and the corresponding Banach space is written as $X \hat{\otimes} Y$. When X and Y are finite dimensional, the projective tensor norm of T coincides with the 1-nuclear norm [8].

The 2-summing norm of $T \in X \otimes Y$, denoted by $\pi_2(T)$, is defined as the smallest constant c that satisfies

$$\left(\sum_i \|Tx_i^*\|_Y^2\right)^{1/2} \leqslant c \sup\left\{ \left(\sum_i |\langle x, x_i^*\rangle|^2\right)^{1/2} \;\middle|\; x \in B_X \right\}$$

for any sequence $(x_i^*) \subset X^*$.

We also consider the γ_2 norm, which is defined though the optimal factorization through a finite dimensional Hilbert space. The γ_2 norm of $T \in X \otimes Y$ is defined by

$$\begin{aligned}\gamma_2(T) := \inf\{&\|T_1\|_\vee \|T_2\|_\vee \mid d \in \mathbb{N}, T_1 \in B(X^*, \ell_2^d), \\ &T_2 \in B(\ell_2^d, Y),\, T = T_2 T_1\}.\end{aligned} \tag{4}$$

Particularly, if $\operatorname{rank}(T) \leqslant r$, the decomposition $T = T_2 T_1$ in (4) can be made via a Hilbert space of dimension up to r.

III. Relation between tensor product norms for rank-r operators

Next we demonstrate the relation between the injective and projective tensor norms of rank-r operators from ℓ_1^n to a class of Banach spaces.

A. Linear operators in $\ell_\infty^n \otimes \ell_\infty^m$

Linial et al. [10] showed that the γ_2 norm on $\ell_\infty^n \otimes \ell_\infty^m$ is upper-bounded by the operator norm multiplied by the square root of the rank. In fact, their result is derived from the fact that the Banach-Mazur distance between a finite dimensional Banach space and a Hilbert space is no larger than the square root of the dimension. It indeed applies to any pair of Banach spaces.

Lemma III.1. *Suppose that $T \in X \otimes Y$ with Banach spaces X, Y satisfies $\operatorname{rank}(T) \leqslant r$. Then $\|T\|_\vee \leqslant \gamma_2(T) \leqslant \sqrt{r}\|T\|_\vee$.*

Furthermore for a linear operator from ℓ_1^n to ℓ_∞^m, it is well known that its γ_2 norm and the 1-nuclear norm are equivalent up to the Grothendieck constant.

Lemma III.2 (little Grothendieck (e.g., [8])). *Let $T \in \ell_\infty^n \otimes \ell_\infty^m$. Then $\gamma_2(T) \leqslant \|T\|_\wedge \leqslant K_G \gamma_2(T)$, where K_G denotes the Grothendieck constant that satisfies $1.67 \leqslant K_G \leqslant 1.79$.*

The following corollary that shows the relation between the injective and projective tensor norms on $\ell_\infty^n \otimes \ell_\infty^m$ is a direct consequence of Lemmas III.1 and III.2.

Corollary III.3. *Let $T \in \ell_\infty^n \otimes \ell_\infty^m$ satisfy $\operatorname{rank}(T) \leqslant r$. Then $\|T\|_\vee \leqslant \|T\|_\wedge \leqslant K_G\sqrt{r}\|T\|_\vee$.*

B. Linear operators in $\ell_\infty^n \otimes \ell_p^m$ with $2 \leqslant p < \infty$

When $T \in \ell_\infty^n \otimes \ell_p^m$ with $2 \leqslant p < \infty$, we derive the relation between the injective and projective tensor norms in several steps given by the following lemmas. We first show that the 2-summing norm of the adjoint does not exceed the operator norm multiplied by the square root of the rank.

Lemma III.4. *Let $T \in \ell_\infty^n \otimes \ell_p^m$ satisfy that its adjoint T^* is 2-summing and* $\operatorname{rank}(T) \leqslant r$. *Then* $\|T\|_\vee \leqslant \pi_2(T^*) \leqslant \sqrt{r}\|T\|_\vee$.

Next we show that the 1-nuclear norm of $T \in L(\ell_1^n, \ell_p^m)$ with $2 \leqslant p < \infty$ is equivalent to the 2-summing norm of the adjoint T^* up to constant $\sqrt{2p}$.

Lemma III.5. *Let $T \in \ell_\infty^n \otimes \ell_p^m$ with $2 \leqslant p < \infty$ satisfy that T^* is 2-summing and* $\operatorname{rank}(T) \leqslant r$. *Then* $\pi_2(T^*) \leqslant \|T\|_\wedge \leqslant \sqrt{2p}\,\pi_2(T^*)$.

By combining Lemmas III.4 and III.5, we obtain the following corollary.

Corollary III.6. *Let $T \in \ell_\infty^n \otimes \ell_p^m$ with $2 \leqslant p < \infty$ satisfy* $\operatorname{rank}(T) \leqslant r$. *Then* $\|T\|_\vee \leqslant \|T\|_\wedge \leqslant \sqrt{2p}\sqrt{r}\|T\|_\vee$.

While the relation between the injective and projective tensor norms of rank-r operators is useful for the entropy estimate, it is often easier to compute the 2-summing norm of the adjoint than the projective norm. Below we show that $\pi_2(T^*)$ is written through an optimal factorization similarly to the γ_2 norm for the tensor product $\ell_\infty^n \otimes \ell_\infty^m$.

Lemma III.7. *Let $T \in \ell_\infty^n \otimes \ell_p^m$. Then*

$$\begin{aligned}\pi_2(T^*) := \inf\{&\pi_2(T_1^*)\|T_2^*\|_\vee \,|\, d \in \mathbb{N},\, T_1^* \in L(\ell_{p'}^m, \ell_2^d), \\ &T_2^* \in L(\ell_2^d, \ell_\infty^n),\, T^* = T_2^* T_1^*\},\end{aligned} \tag{5}$$

where $1/p + 1/p' = 1$.

In a special case when $T \in \ell_\infty^n \otimes \ell_2^m$, the 2-summing norm of T^* is computed by the following optimization problem:

$$\begin{aligned}\pi_2(T^*) = \inf\{&\|T_1\|_{\mathrm{HS}}\|T_2\|_\vee \,|\, d \in \mathbb{N},\, T_1 \in L(\ell_2^d, \ell_2^m), \\ &T_2 \in L(\ell_1^n, \ell_2^d),\, T = T_1 T_2\}.\end{aligned} \tag{6}$$

Indeed, it trivially holds that $\|T_1^*\|_\vee = \|T_1\|_\vee$. Furthermore, for every $T_1^* : \ell_2^m \to \ell_2^d$, we have

$$\pi_2(T_1^*) = \|T_1^*\|_{\mathrm{HS}} = \|T_1\|_{\mathrm{HS}} = \pi_2(T_1).$$

Therefore, (6) follows from (5). In a companion paper, we show that (6) is written as a semidefinite program.

IV. Entropy Estimates of Tensor Products

For symmetric convex bodies D and E, the *covering number* $N(D, E)$ and the *packing number* $M(D, E)$ are respectively defined by

$$N(D, E) := \min\Big\{l \,\Big|\, \exists y_1, \ldots, y_l \in D,\, D \subset \bigcup_{1 \leqslant j \leqslant l} (y_j + E)\Big\},$$

$$M(D, E) := \max\Big\{l \,\Big|\, \exists y_1, \ldots, y_l \in D,\, y_j - y_k \notin E,\, \forall j \neq k\Big\}.$$

Let X, Y be Banach spaces. For $T \in L(X, Y)$, the *dyadic entropy number* [11] is defined by

$$e_k(T) := \inf\{\epsilon > 0 \,|\, M(T(B_X), \epsilon B_Y) \leqslant 2^{k-1}\}.$$

We will use the following shorthand notation for the weighted summation of the dyadic entropy numbers:

$$\mathcal{E}_{2,1}(T) := \sum_{k=1}^\infty k^{-1/2} e_k(T),$$

which is up to a constant equivalent to the entropy integral $\int_0^\infty \sqrt{\ln N(T(B_X), \epsilon B_Y)}\,d\epsilon$ [12], which plays a key role in analyzing properties on random linear operators on low-rank matrices.

In this section, we derive the $\mathcal{E}_{2,1}$ of the identity operator from the injective tensor product to the projective tensor product of a set of Banach space pairs. Note that these tensor product spaces are valid Banach spaces too. The main machinery in deriving these estimates is Maurey's empirical method [9], summarized in the following lemma.

Lemma IV.1. *Let $T \in L(\ell_1^n, \ell_\infty^m)$. Then*

$$\mathcal{E}_{2,1}(T) \leqslant C\sqrt{1 + \ln(m \vee n)}\,(1 + \ln(m \wedge n))^{3/2}\|T\|_\vee.$$

In order to apply Lemma IV.1 to $\ell_\infty^n \otimes \ell_\infty^m$, we use the fact that $\ell_\infty^m \check{\otimes} \ell_\infty^n$ is isometrically isomorphic to ℓ_∞^{mn}. In fact,

$$\|M\|_{\ell_\infty^m \check{\otimes} \ell_\infty^n} = \max_{1 \leqslant j \leqslant n} \|Me_j\|_\infty = \|\mathrm{vec}(M)\|_\infty,$$

where e_j denotes the jth column of the n-by-n identity matrix and $\mathrm{vec}(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}^{mn}$ vectorizes $M \in \mathbb{R}^{m\times n}$ by stacking its columns vertically. On the other hand, the trace dual and Banach space dual of $\ell_1^n \hat{\otimes} \ell_1^m$ are $\ell_1^m \check{\otimes} \ell_1^n$ and $\ell_\infty^n \check{\otimes} \ell_\infty^m$, respectively. Therefore, we also have that $\ell_1^m \hat{\otimes} \ell_1^n$ is isotropically isomorphic to ℓ_1^{mn}. With these isometric isomorphisms, Lemma IV.1 provides the following estimate.

Proposition IV.2. *There exists a numerical constant C such that*

$$\mathcal{E}_{2,1}(\mathrm{id} : \ell_\infty^m \hat{\otimes} \ell_\infty^n \to \ell_\infty^m \check{\otimes} \ell_\infty^n) \leqslant C\sqrt{m+n}\,(1 + \ln(mn))^{3/2}.$$

Next, to apply Lemma IV.1 to $\ell_\infty^n \otimes \ell_p^m$ with $2 \leqslant p < \infty$, we need the following result that shows embedding of finite dimensional ℓ_p space to ℓ_1 up to a small Banach-Mazur distance.

Lemma IV.3 ([9, Lemma 5])**.** *Let $1 < p \leqslant 2$ and $\epsilon > 0$. There is a constant $c(p, \epsilon) > 0$ for which the following property holds: For each m, there exists $k \geqslant c(p, \epsilon)m$ so that ℓ_1^m contains a subspace $(1+\epsilon)$-isomorphic to ℓ_p^k, i.e., the Banach-Mazur distance is upper-bounded by $(1 + \epsilon)$.*

Then we obtain the following entropy estimate for $\ell_\infty^n \otimes \ell_p^m$ with $2 \leqslant p < \infty$ by combining Lemmas IV.1 and IV.3.

Proposition IV.4. *Let $2 \leqslant p < \infty$. Then*

$$\mathcal{E}_{2,1}(\mathrm{id} : \ell_\infty^n \hat{\otimes} \ell_p^m \to \ell_\infty^n \check{\otimes} \ell_p^m) \leqslant C\sqrt{n+m}\,(1 + \ln(nm))^{3/2}.$$

V. Applications

We illustrate the utility of the entropy estimates in Section IV to derive performance guarantees of low-rank recovery problems. We consider the linear inverse problem that reconstructs the unknown matrix M_0 of rank up to r from the measurements given by

$$b = \mathcal{A}(M_0) + z$$

with a linear operator $\mathcal{A}$ and random noise z. Particularly we illustrate on two applications where each entry of $\mathcal{A}(M_0)$ is written as the inner product of M_0 and a given rank-1 matrix.

We interpret $M_0 \in \mathbb{C}^{m\times n}$ as a rank-r linear operator or equivalently a tensor in $\ell_p^n \otimes \ell_q^m$, where p and q are chosen according to the structure in $\mathcal{A}$. Then we consider the estimate given by the following optimization problem:

$$\begin{array}{ll} \underset{M}{\text{minimize}} & \|b - \mathcal{A}(M)\|_2^2 \\ \text{subject to} & \|\| M \|\|_{p,q,r} \leqslant \alpha, \end{array} \tag{7}$$

where

$$\|\| M \|\|_{p,q,r} := \max\left(\|M\|_\wedge/\sqrt{r},\ \|M\|_\vee\right).$$

The projective tensor norm can be replaced by an equivalent norm to enable efficient implementation. This does not break performance guarantees we present below.

A. Matrix completion

The optimization in (7) reduces to the max-norm constrained matrix completion [5] when $p = \infty$ and $q = \infty$. For brevity, we consider the square matrix case ($m = n$). By applying the entropy estimate in Proposition IV.2 to a version of the Rudelson-Vershynin lemma [13, Proposition 2.6], which generalizes upon previous works [14], [15], [16], we obtain the following concentration result on the quadratic form with random entriwise sampling operator.

Proposition V.1. *Let (i_l, j_l) for $l = 1, \ldots, L$ be independent copies of a uniform random variable on $\{1, \ldots, n\} \times \{1, \ldots, n\}$. Then*

$$\sup_{\|\|M\|\|_{\infty,\infty,r}\leqslant\alpha} \left| \frac{1}{L}\sum_{l=1}^{L} |\langle e_{i_l} \otimes e_{j_l}, M\rangle|^2 - \frac{\|M\|_{\mathrm{HS}}^2}{n^2} \right| \leqslant \frac{C\alpha\sqrt{rn\log^3 n}}{\sqrt{L}} \left(\frac{\|M\|_{\mathrm{HS}}}{n} + \frac{\alpha\sqrt{rn\log^3 n}}{\sqrt{L}} \right)$$

holds except with probability $O(e^{-n})$.

This concentration inequality provides an alternative derivation of the performance guarantee on the max-norm constrained matrix completion [5].

B. Decentralized subspace sketching

Decentralized subspace sketching is formulated as the optimization in (7) with $\mathcal{A}$ that randomly samples entries of each column of the unknown matrix M_0. We choose Banach spaces by $p = \infty$ and $q = 2$ for this problem. By plugging in the entropy estimate in Proposition IV.4 to the suprema of chaos processes by Krahmer et al. [17], we obtain the following concentration inequality.

Proposition V.2. *Let $b_{l,k}$ for $l = 1, \ldots, L$ and $k = 1, \ldots, n$ be independent copies of a Gaussian vector $\xi \in \mathbb{R}^m$ with $\mathbb{E}[\xi] = 0$ and $\mathbb{E}[\xi\xi^\top] = I_m$. Let $d = m + n$. Then*

$$\sup_{\|\|M\|\|_{\infty,2,r}\leqslant\alpha} \left| \|M\|_{\mathrm{HS}}^2 - \frac{1}{L}\sum_{l=1}^{L}\sum_{k=1}^{n} |\langle b_{l,k} \otimes e_k, M\rangle|^2 \right| \leqslant \frac{C\alpha^2\sqrt{rd\log^3 d}}{\sqrt{Ln}} \left(1 + \frac{\sqrt{rd\log^3 d}}{\sqrt{Ln}} \right)$$

holds except with probability $O(e^{-crd})$.

This result plays a key role in deriving a performance guarantee of decentralized subspace sketching by a convex program at a near optimal sample complexity, which will be presented in a companion paper.

VI. Discussion

We interpreted a low-rank matrix as a linear operator between two Banach spaces or equivalently as an element in the tensor product space. Various matrix norms that induced a low-rank solution to inverse problems are interpreted as tensor product norms extensively studied in the functional analysis literature. For a class of tensor products of ℓ_∞^n and ℓ_p^m for $2 \leqslant p \leqslant \infty$, we show the interlacing properties between the injective and projective tensor norms and derive entropy estimates between two corresponding tensor product spaces. We illustrated the use of obtained estimates on matrix completion and decentralized subspace sketching. Omitted proofs and further generalization on a broader class of tensor products including operators between infinite dimensional spaces will be available in the full version manuscript [18].

Acknowledgement

This work was supported in part by NSF grants CCF 17-18771 and DMS 18-00872.

References

[1] M. A. Davenport and J. Romberg, "An overview of low-rank matrix recovery from incomplete observations," *IEEE J. Sel. Topics Signal Process.*, 2016.

[2] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.*, 2010.

[3] E. Candès and X. Li, "Solving quadratic equations via PhaseLift when there are about as many equations as unknowns," *Found. of Comput. Math.*, 2014.

[4] N. Srebro and A. Shraibman, "Rank, trace-norm and max-norm," in *COLT 2005*. Springer.

[5] T. T. Cai and W.-X. Zhou, "Matrix completion via max-norm constrained optimization," *Electron. J. Stat.*, 2016.

[6] R. S. Srinivasa, K. Lee, M. Junge, and J. Romberg, "Decentralized subspace sketching," *submitted*.

[7] S. Negahban and M. J. Wainwright, "Estimation of (near) low-rank matrices with noise and high-dimensional scaling," *Ann. Stat.*, 2011.

[8] G. J. O. Jameson, *Summing and nuclear norms in Banach space theory*. Cambridge University Press, 1987.

[9] B. Carl, "Inequalities of Bernstein-Jackson-type and the degree of compactness of operators in Banach spaces," *Ann. Inst. Fourier*, 1985.

[10] N. Linial, S. Mendelson, G. Schechtman, and A. Shraibman, "Complexity measures of sign matrices," *Combinatorica*, 2007.

[11] B. Carl and I. Stephani, *Entropy, Compactness and the Approximation of Operators*. Cambridge University Press, 1990.

[12] G. Pisier, *The volume of convex bodies and Banach space geometry*. Cambridge University Press, 1999.

[13] M. Junge and K. Lee, "Generalized notions of sparsity and restricted isometry property. Part I: A unified framework," *Inf. Inference*, 2019.

[14] M. Rudelson and R. Vershynin, "On sparse reconstruction from fourier and gaussian measurements," *Comm. Pure Appl. Math*, 2008.

[15] H. Rauhut, "Compressive sensing and structured random matrices," 2010.

[16] S. Dirksen, "Tail bounds via generic chaining," *Electron. J. Probab.*, 2015.

[17] F. Krahmer, S. Mendelson, and H. Rauhut, "Suprema of chaos processes and the restricted isometry property," *Comm. Pure Appl. Math*, 2014.

[18] M. Junge and K. Lee, "Convex relexation of low-rank operators in tensor products of Banach spaces," *in preparation*.