RESTRICTED INVERTIBILITY AND THE BANACH-MAZUR DISTANCE TO THE CUBE

PIERRE YOUSSEF

ABSTRACT. We prove a normalized version of the restricted invertibility principle obtained by Spielman-Srivastava in [15]. Applying this result, we get a new proof of the proportional Dvoretzky-Rogers factorization theorem recovering the best current estimate in the symmetric setting while we improve the best known result in the nonsymmetric case. As a consequence, we slightly improve the estimate for the Banach-Mazur distance to the cube: the distance of every n-dimensional normed space from ℓ_∞^n is at most $(2n)^{\frac{5}{6}}$. Finally, using tools from the work of Batson-Spielman-Srivastava in [2], we give a new proof for a theorem of Kashin-Tzafriri [11] on the norm of restricted matrices.

1. INTRODUCTION

Given an $n \times m$ matrix U, viewed as an operator from ℓ_2^m to ℓ_2^n, the restricted invertibility problem asks if we can extract a large number of linearly independent columns of U and provide an estimate for the norm of the restricted inverse. If we write U_σ for the restriction of U to the columns Ue_i, $i \in \sigma \subset \{1, \dots, m\}$, we want to find a subset σ, of cardinality k as large as possible, such that $\|U_\sigma x\|_2 \geqslant c\|x\|_2$ for all $x \in \mathbb{R}^\sigma$ and to estimate the constant c (which will depend on the operator U). This question was studied by Bourgain-Tzafriri [4] who obtained a result for square matrices:

Given an $n \times n$ matrix T (viewed as an operator on ℓ_2^n) whose columns are of norm one, there exists $\sigma \subset \{1, \dots, n\}$ with $|\sigma| \geqslant d\frac{n}{\|T\|^2}$ such that $\|T_\sigma x\|_2 \geqslant c\|x\|_2$ for all $x \in \mathbb{R}^\sigma$, where $d, c > 0$ are absolute constants.

Here and in the rest of the paper, $\|\cdot\|_2$ denotes the Euclidean norm. For any matrix A, $\|A\|$ denotes its operator norm seen as an operator on l_2 and $\|A\|_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm, i.e.

$$\|A\|_{\mathrm{HS}} = \sqrt{Tr(A \cdot A^*)} = \left(\sum_i \|C_i\|_2^2\right)^{1/2}$$

where C_i are the columns of A.

Since the identity operator can be decomposed in the form $Id = \sum_{j \leqslant n} e_j e_j^t$ where $(e_j)_{j \leqslant n}$ is the canonical basis of $\mathbb{R}^n$, the previous result states that one can find a large part of this basis (of cardinality greater than $d\frac{n}{\|T\|^2}$) on the span of which the operator T is invertible and the norm of its inverse is controlled by an absolute constant.

In [21], Vershynin generalized this result for any decomposition of the identity and improved the estimate for the size of the subset. Using a technical iteration scheme based on the previous result of Bourgain-Tzafriri, combined with a theorem of Kashin-Tzafriri which we will discuss in the last section, he obtained the following :

Let $Id = \sum_{j \leqslant m} x_j x_j^t$ and let T be a linear operator on ℓ_2^n. For any $\varepsilon \in (0,1)$ one can find $\sigma \subset \{1, \dots, m\}$ with

$$|\sigma| \geqslant (1-\varepsilon)\frac{\|T\|_{\mathrm{HS}}^2}{\|T\|^2}$$

such that

$$\left\| \sum_{j\in\sigma} a_j \frac{Tx_j}{\|Tx_j\|_2} \right\|_2 \geqslant c(\varepsilon) \left(\sum_{j\in\sigma} a_j^2 \right)^{\frac{1}{2}}$$

for all scalars (a_j).

One can easily check that, in the case of the canonical decomposition, this is a generalization of the Bourgain-Tzafriri theorem, which was previously only proved for a fixed value of ε. The constant $c(\varepsilon)$ plays a crucial role in applications and finding the right dependence is an important problem. Let us mention that Veshynin obtained also a non trivial upper bound which we don't state here; we refer to [21] for a full statement.

Back to the original restricted invertibility problem, a recent work of Spielman-Srivastava [15] provides the best known estimate for the norm of the inverse matrix. Their proof uses a new deterministic method based on linear algebra, while the previous works on the subject employed probabilistic, combinatorial and functional-analytic arguments.

More precisely, Spielman-Srivastava proved the following:

Theorem A (Spielman-Srivastava)**.** *Let* $x_1, \dots x_m \in \mathbb{R}^n$ *such that* $Id = \sum_i x_i x_i^t$ *and let* $0 < \varepsilon < 1$. *For every linear operator* $T : \ell_2^n \to \ell_2^n$ *there exists a subset* $\sigma \subset \{1, \dots, m\}$ *of size* $|\sigma| \geqslant \left\lfloor (1-\varepsilon)^2 \frac{\|T\|_{\mathrm{HS}}^2}{\|T\|^2} \right\rfloor$ *for which* $\{Tx_i\}_{i\in\sigma}$ *is linearly independent and*

$$\lambda_{\min}\left(\sum_{i\in\sigma} (Tx_i)(Tx_i)^t \right) > \frac{\varepsilon^2 \|T\|_{\mathrm{HS}}^2}{m},$$

where $\lambda_{\min}$ *is computed on* $\mathrm{span}\{Tx_i\}_{i\in\sigma}$ *or simply here* $\lambda_{\min}$ *denotes the smallest nonzero eigenvalue of the corresponding operator. .*

One can view the previous result as an invertibility theorem for rectangular matrices. Given, as above, a decomposition of the identity and a linear operator T, we can associate to these an $n \times m$ matrix U whose columns are the vectors $(Tx_j)_{j\leqslant m}$. Since $Id = \sum_j x_j x_j^t$, one can easily check that

$$U \cdot U^t = T \cdot T^t = \sum_j (Tx_j) \cdot (Tx_j)^t.$$

Hence,

$$\|U\|_{\mathrm{HS}} = \|T\|_{\mathrm{HS}} \quad \text{and} \quad \|U\| = \|T\|,$$

and thus the previous result can be written in terms of the rectangular matrix U.

In the applications, one might need to extract multiples of the columns of the matrix. Adapting the proof of Spielman-Srivastava, we will generalize the restricted invertibility theorem for any rectangular matrix and, under some conditions, for any choice of multiples.

If D is an $m \times m$ diagonal matrix with diagonal entries $(\alpha_j)_{j\leqslant m}$, we set $I_D := \{j \leqslant m \mid \alpha_j \neq 0\}$ and write D_σ^{-1} for the restricted inverse of D i.e the diagonal matrix whose diagonal entries are the inverses of the respective entries of D for indices in σ and zero elsewhere. The main result of this paper is the following:

Theorem 1.1. *Given an* $n \times m$ *matrix* U *and a diagonal* $m \times m$ *matrix* D *with* $(\alpha_j)_{j\leqslant m}$ *on its diagonal, with the property that* $\mathrm{Ker}(D) \subset \mathrm{Ker}(U)$, *then for any* $\varepsilon \in (0,1)$ *there exists* $\sigma \subset I_D$ *with*

$$|\sigma| \geqslant \left\lceil (1-\varepsilon)^2 \frac{\|U\|_{\mathrm{HS}}^2}{\|U\|^2} \right\rceil$$

such that

$$s_{\min}\left(U_\sigma D_\sigma^{-1} \right) > \frac{\varepsilon \|U\|_{\mathrm{HS}}}{\|D\|_{\mathrm{HS}}},$$

where $s_{\min}$ *denotes the smallest singular value.*

Note that if we apply this fact to the matrix U which we associated with a linear operator T and a decomposition of the identity, and we take D to be the identity operator, we recover the restricted invertibility theorem of Spielman-Srivastava. Taking D the diagonal matrix with diagonal entries $\|Tx_j\|_2$, it is easy to see that we recover the "normalized" restricted invertibility principle stated by Vershynin with $c(\varepsilon) = \varepsilon$.

In Section 2, we give the proof of the main result. In section 3, we use Theorem 1.1 to give an alternative proof for the proportional Dvoretzky-Rogers factorization; in the symmetric case, we recover the best known dependence and improve the constants involved which allows us to improve the estimate of the Banach-Mazur distance to the cube, while in the nonsymmetric case, we improve the best known dependence for the proportional Dvoretzky-Rogers factorization. Finally, in Section 4 we give a new proof of a theorem due to Kashin-Tzafriri [11] which deals with the norm of coordinate projections of a matrix; our proof slightly improves the result of Kashin-Tzafriri and has the advantage of producing a deterministic algorithm.

2. PROOF OF THEOREM 1.1

Since the rank and the eigenvalues of $(U_\sigma D_\sigma^{-1})^t \cdot (U_\sigma D_\sigma^{-1})$ and $(U_\sigma D_\sigma^{-1}) \cdot (U_\sigma D_\sigma^{-1})^t$ are the same, it suffices to prove that $(U_\sigma D_\sigma^{-1}) \cdot (U_\sigma D_\sigma^{-1})^t$ has rank equal to $k = |\sigma|$ and its smallest positive eigenvalue is greater than $\varepsilon^2 \frac{\|U\|_{\text{HS}}^2}{\|D\|_{\text{HS}}^2}$. Note that

$$(U_\sigma D_\sigma^{-1}) \cdot (U_\sigma D_\sigma^{-1})^t = \sum_{j\in\sigma} \left(UD_\sigma^{-1}e_j\right) \cdot \left(UD_\sigma^{-1}e_j\right)^t = \sum_{j\in\sigma} \left(\frac{Ue_j}{\alpha_j}\right) \cdot \left(\frac{Ue_j}{\alpha_j}\right)^t$$

We are going to construct the matrix $A_k = \sum_{j\in\sigma} \left(UD_\sigma^{-1}e_j\right) \cdot \left(UD_\sigma^{-1}e_j\right)^t$ by iteration. We begin by setting $A_0 = 0$ and at each step we will be adding a rank one matrix $\left(\frac{Ue_j}{\alpha_j}\right) \cdot \left(\frac{Ue_j}{\alpha_j}\right)^t$ for a suitable j, which will give a new positive eigenvalue. This will guarantee that the vector $UD_\sigma^{-1}e_j$ chosen in each step is linearly independent from the previous ones.

If A and B are symmetric matrices, we write $A \preceq B$ if $B - A$ is a positive semidefinite matrix. Recall the Sherman-Morrison Formula which will be needed in the proof. For any invertible matrix A and any vector v we have

$$(A + v\cdot v^t)^{-1} = A^{-1} - \frac{A^{-1}v \cdot v^t A^{-1}}{1 + v^t A^{-1} v}.$$

We will also apply the following lemma which appears as Lemma 6.3 in [16]:

Lemma 2.1. *Suppose that $A \succeq 0$ has q nonzero eigenvalues, all greater than $b' > 0$. If $v \neq 0$ and*

$$v^t(A - b'I)^{-1}v < -1, \tag{1}$$

then $A + vv^t$ has $q+1$ nonzero eigenvalues, all greater than b'.

Proof. The proof of the lemma is simple and makes use of the Sherman-Morrison formula. Denote $\lambda_1 \geqslant \lambda_2 \geqslant .. \geqslant \lambda_q > b'$ the eigenvalues of A and $\lambda_1' \geqslant .. \geqslant \lambda_q' \geqslant \lambda_{q+1}' \geqslant 0$ the eigenvalues of $A + vv^t$. By the Cauchy interlacing Theorem we have

$$\lambda_1' \geqslant \lambda_1 \geqslant \lambda_2' \geqslant .. \geqslant \lambda_q' \geqslant \lambda_q \geqslant \lambda_{q+1}'$$

Since $\lambda_q > b'$ then $\lambda_q' > b'$ and it remains to prove that $\lambda_{q+1}' > b'$.

$$\text{Tr}(A + vv^t - b'I)^{-1} = \sum_{j\leqslant q+1} \frac{1}{\lambda_j' - b'} - \sum_{j>q+1} \frac{1}{b'}$$

$$\text{Tr}(A - b'I)^{-1} = \sum_{j\leqslant q} \frac{1}{\lambda_j - b'} - \sum_{j>q} \frac{1}{b'}$$

$$\begin{aligned}\mathrm{Tr}(A+vv^t-b'I)^{-1}-\mathrm{Tr}(A-b'I)^{-1}&=\sum_{j\leqslant q}\left[\frac{1}{\lambda'_j-b'}-\frac{1}{\lambda_j-b'}\right]+\frac{1}{\lambda'_{q+1}-b'}-\frac{1}{b'}\\&\leqslant\frac{1}{\lambda'_{q+1}-b'}-\frac{1}{b'}\end{aligned}$$

By the Sherman-Morrison's formula we have:

$$(A+vv^t-b'I)^{-1}=(A-b'I)^{-1}-\frac{(A-b'I)^{-1}vv^t(A-b'I)^{-1}}{1+v^t(A-b'I)^{-1}v}$$

Now taking the Trace we get:

$$\mathrm{Tr}(A+vv^t-b'I)^{-1}-\mathrm{Tr}(A-b'I)^{-1}=-\frac{v^t(A-b'I)^{-2}v}{1+v^t(A-b'I)^{-1}v}$$

Since $(A-b'I)^{-2}$ is positive definite and using the hypothesis, one can see that the right hand side in the previous equality is positive. Therefore we get:

$$\frac{1}{\lambda'_{q+1}-b'}-\frac{1}{b'}>0$$

wich means that $\lambda'_{q+1}>b'$.

□

For any symmetric matrix A and any $b>0$, we define

$$\phi(A,b)=\mathrm{Tr}\left(U^t(A-bI)^{-1}U\right)$$

as the potential corresponding to the barrier b.
At each step l, the matrix already constructed is denoted by A_l and the barrier by b_l. Suppose that A_l has l nonzero eigenvalues all greater than b_l. As mentioned before, we will try to construct A_{l+1} by adding a rank one matrix $v\cdot v^t$ to A_l so that A_{l+1} has $l+1$ nonzero eigenvalues all greater than $b_{l+1}=b_l-\delta$ and $\phi(A_{l+1},b_{l+1})\leqslant\phi(A_l,b_l)$. Note that

$$\begin{aligned}\phi(A_{l+1},b_{l+1})&=\mathrm{Tr}\left(U^t(A_l+vv^t-b_{l+1}I)^{-1}U\right)\\&=\mathrm{Tr}\left(U^t(A_l-b_{l+1}I)^{-1}U\right)-\mathrm{Tr}\left(\frac{U^t(A_l-b_{l+1}I)^{-1}vv^t(A_l-b_{l+1}I)^{-1}U}{1+v^t(A_l-b_{l+1}I)^{-1}v}\right)\\&=\phi(A_l,b_{l+1})-\frac{v^t(A_l-b_{l+1}I)^{-1}UU^t(A_l-b_{l+1}I)^{-1}v}{1+v^t(A_l-b_{l+1}I)^{-1}v}.\end{aligned}$$

So, in order to have $\phi(A_{l+1},b_{l+1})\leqslant\phi(A_l,b_l)$, we must choose a vector v verifying

$$-\frac{v^t(A_l-b_{l+1}I)^{-1}UU^t(A_l-b_{l+1}I)^{-1}v}{1+v^t(A_l-b_{l+1}I)^{-1}v}\leqslant\phi(A_l,b_l)-\phi(A_l,b_{l+1}).\tag{2}$$

Since $v^t(A_l-b_{l+1}I)^{-1}UU^t(A_l-b_{l+1}I)^{-1}v$ and $\phi(A_l,b_l)-\phi(A_l,b_{l+1}))$ are positive, choosing v verifying conditions (1) and (2) is equivalent to choosing v which satisfies the following:

$$v^t(A_l-b_{l+1}I)^{-1}UU^t(A_l-b_{l+1}I)^{-1}v\leqslant\left(\phi(A_l,b_l)-\phi(A_l,b_{l+1})\right)\left(-1-v^t(A_l-b_{l+1}I)^{-1}v\right)$$

Since $UU^t\preceq\|U\|^2Id$ and $(A_l-b_{l+1}I)^{-1}$ is symmetric, it is sufficient to choose v so that

$$v^t(A_l-b_{l+1}I)^{-2}v\leqslant\frac{1}{\|U\|^2}\left(\phi(A_l,b_l)-\phi(A_l,b_{l+1})\right)\left(-1-v^t(A_l-b_{l+1}I)^{-1}v\right)\tag{3}$$

Recall the notation $I_D:=\{j\leqslant m\mid\alpha_j\neq0\}$ where $(\alpha_j)_{j\leqslant m}$ are the diagonal entries of D. Since we have assumed that $\mathrm{Ker}(D)\subset\mathrm{Ker}(U)$, we have

$$\|U\|_{\mathrm{HS}}^2=\sum_{j\leqslant m}\|Ue_j\|_2^2=\sum_{j\in I_D}\|Ue_j\|_2^2\leqslant|I_D|\cdot\|U\|^2,$$

and thus $|I_D| \geqslant \frac{\|U\|_{\mathrm{HS}}^2}{\|U\|^2}$. At each step, we will select a vector v satisfying (3) among $(\frac{Ue_j}{\alpha_j})_{j\in I_D}$. Our task therefore is to find $j \in I_D$ such that

$$(4) \quad (Ue_j)^t(A_l - b_{l+1}I)^{-2}Ue_j \leqslant \frac{\phi(A_l, b_l) - \phi(A_l, b_{l+1})}{\|U\|^2}\Big(-\alpha_j^2 - (Ue_j)^t(A_l - b_{l+1}I)^{-1}Ue_j\Big)$$

The existence of such a $j \in I_D$ is guaranteed by the fact that condition (4) holds true if we take the sum over all $(\frac{Ue_j}{\alpha_j})_{j\in D}$. The hypothesis $\mathrm{Ker}(D) \subset \mathrm{Ker}(U)$ implies that:

- $\sum\limits_{j\in I_D} (Ue_j)^t(A_l - b_{l+1}I)^{-2}Ue_j = \mathrm{Tr}\Big(U^t(A_l - b_{l+1}I)^{-2}U\Big)$,
- $\sum\limits_{j\in I_D} (Ue_j)^t(A_l - b_{l+1}I)^{-1}Ue_j = \mathrm{Tr}\Big(U^t(A_l - b_{l+1}I)^{-1}U\Big)$.

Therefore it is enough to prove that, at each step, one has

$$(5) \qquad \mathrm{Tr}(U^t(A_l - b_{l+1}I)^{-2}U) \leqslant \frac{\phi(A_l, b_l) - \phi(A_l, b_{l+1})}{\|U\|^2}\Big(-\|D\|_{\mathrm{HS}}^2 - \phi(A_l, b_{l+1})\Big)$$

The rest of the proof is similar to the one in [16]. One just needs to replace m by $\|D\|_{\mathrm{HS}}^2$. For the sake of completeness, we include the proof. The next lemma will determine the conditions required at each step in order to prove (5).

Lemma 2.2. *Suppose that A_l has l nonzero eigenvalues all greater than b_l, and write Z for the orthogonal projection onto the kernel of A_l. If*

$$(6) \qquad \phi(A_l, b_l) \leqslant -\|D\|_{\mathrm{HS}}^2 - \frac{\|U\|^2}{\delta}$$

and

$$(7) \qquad 0 < \delta < b_l \leqslant \delta \frac{\|ZU\|_{\mathrm{HS}}^2}{\|U\|^2},$$

then there exists $i \in I_D$ such that $A_{l+1} := A_l + \left(\frac{Ue_i}{\alpha_i}\right)\cdot\left(\frac{Ue_i}{\alpha_i}\right)^t$ has $l+1$ nonzero eigenvalues all greater than $b_{l+1} := b_l - \delta$ and $\phi(A_{l+1}, b_{l+1}) \leqslant \phi(A_l, b_l)$.

Proof. As mentioned before, it is enough to prove inequality (5). We set $\Delta_l := \phi(A_l, b_l) - \phi(A_{l+1}, b_{l+1})$. By (6), we get

$$\phi(A_l, b_{l+1}) \leqslant -\|D\|_{\mathrm{HS}}^2 - \frac{\|U\|^2}{\delta} - \Delta_l.$$

Inserting this in (5), we see that it is sufficient to prove the following inequality:

$$(8) \qquad \mathrm{Tr}\left(U^t(A_l - b_{l+1}I)^{-2}U\right) \leqslant \Delta_l\left(\frac{\Delta_l}{\|U\|^2} + \frac{1}{\delta}\right).$$

Now, denote by P the orthogonal projection onto the image of A_l. We set

$$\phi^P(A_l, b_l) := \mathrm{Tr}\left(U^tP(A_l - b_lI)^{-1}PU\right) \quad \text{and} \quad \Delta_l^P := \phi^P(A_l, b_l) - \phi^P(A_l, b_{l+1})$$

and use similar notation for Z. Since P, Z and A_l commute, one can write

$$\Delta_l = \Delta_l^P + \Delta_l^Z \quad \text{and} \quad \phi(A_l, b_l) = \phi^P(A_l, b_l) + \phi^Z(A_l, b_l).$$

Note that:

$$\begin{aligned}(A_l - b_lI)^{-1} - (A_l - b_{l+1}I)^{-1} &= (A_l - b_lI)^{-1}(b_lI - A_l + A_l - b_{l+1}I)(A_l - b_{l+1}I)^{-1} \\ &= \delta(A_l - b_lI)^{-1}(A_l - b_{l+1}I)^{-1}\end{aligned}$$

and since $P(A_l - b_lI)^{-1}P$ and $P(A_l - b_{l+1}I)^{-1}P$ are positive semidefinite, we have:

$$U^tP(A_l - b_lI)^{-1}PU - U^tP(A_l - b_{l+1}I)^{-1}PU \succeq \delta U^tP(A_l - b_{l+1}I)^{-2}PU.$$

Inserting this in (8), it is enough to prove that:

$$\operatorname{Tr}\left(U^t Z(A_l - b_{l+1}I)^{-2} ZU\right) \leqslant \Delta_l \left(\frac{\Delta_l}{\|U\|^2} + \frac{1}{\delta}\right) - \frac{\Delta_l^P}{\delta}.$$

Since $A_l Z = 0$, we have:

- $\operatorname{Tr}(U^t Z(A_l - b_{l+1}I)^{-2} ZU) = \frac{\|ZU\|_{\mathrm{HS}}^2}{b_{l+1}^2}$ and
- $\Delta_l^Z = -\frac{\|ZU\|_{\mathrm{HS}}^2}{b_l} + \frac{\|ZU\|_{\mathrm{HS}}^2}{b_{l+1}} = \delta \frac{\|ZU\|_{\mathrm{HS}}^2}{b_l b_{l+1}}$,

so taking into account the fact that $\Delta_l \geqslant \Delta_l^Z \geqslant 0$, it remains to prove the following:

$$\frac{\|ZU\|_{\mathrm{HS}}^2}{b_{l+1}^2} \leqslant \delta^2 \frac{\|ZU\|_{\mathrm{HS}}^4}{\|U\|_2^2 b_l^2 b_{l+1}^2} + \frac{\|ZU\|_{\mathrm{HS}}^2}{b_l b_{l+1}}. \tag{9}$$

By Hypothesis (7), this last inequality follows by

$$\frac{\|ZU\|_{\mathrm{HS}}^2}{b_{l+1}^2} \leqslant \delta \frac{\|ZU\|_{\mathrm{HS}}^2}{b_l b_{l+1}^2} + \frac{\|ZU\|_{\mathrm{HS}}^2}{b_l b_{l+1}}, \tag{10}$$

which is trivially true since $b_{l+1} = b_l - \delta$. □

We are now able to complete the proof of Theorem 1.1. To this end, we must verify that conditions (6) and (7) hold at each step. At the beginning we have $A_0 = 0$ and $Z = Id$, so we must choose a barrier b_0 such that:

$$-\frac{\|U\|_{\mathrm{HS}}^2}{b_0} \leqslant -\|D\|_{\mathrm{HS}}^2 - \frac{\|U\|^2}{\delta} \tag{11}$$

and

$$b_0 \leqslant \delta \frac{\|U\|_{\mathrm{HS}}^2}{\|U\|^2}. \tag{12}$$

We choose

$$b_0 := \varepsilon \frac{\|U\|_{\mathrm{HS}}^2}{\|D\|_{\mathrm{HS}}^2} \quad \text{and} \quad \delta := \frac{\varepsilon}{1-\varepsilon} \frac{\|U\|^2}{\|D\|_{\mathrm{HS}}^2},$$

and we note that (11) and (12) are verified. Also, at each step (6) holds because $\phi(A_{l+1}, b_{l+1}) \leqslant \phi(A_l, b_l)$. Since $\|ZU\|_{\mathrm{HS}}^2$ decreases at each step by at most $\|U\|^2$, the right-hand side of (7) decreases by at most δ, and therefore (7) holds once we replace b_l by $b_l - \delta$.

Finally note that, after $k = (1-\varepsilon)^2 \frac{\|U\|_{\mathrm{HS}}^2}{\|U\|^2}$ steps, the barrier will be

$$b_k = b_0 - k\delta = \varepsilon^2 \frac{\|U\|_{\mathrm{HS}}^2}{\|D\|_{\mathrm{HS}}^2}.$$

This completes the proof.

3. Proportionnal Dvoretzky-Rogers factorization

Let $\mathbb{BM}_n$ denote the space of all n-dimensional normed spaces X, known as the Banach-Mazur compactum. If X, Y are in $\mathbb{BM}_n$, we define the Banach-Mazur distance between X and Y as follows:

$$d(X,Y) = \inf\{\|T\| \cdot \|T^{-1}\| \mid T \text{ is an isomorphism between X and Y }\}$$

Remark 3.1. *For K, L two symmetric convex bodies in $\mathbb{R}^n$, one can define the Banach-Mazur distance between K and L as*

$$d(K,L) = \inf\{\alpha/\beta \mid \beta L \subset T(K) \subset \alpha L\}$$

One can easily check that this distance is coherent with the previous one as $d(X,Y) = d(B_X, B_Y)$.

By the classical Dvoretzky-Rogers lemma [6], it is proven that if X is an n-dimensional Banach space then there exist $x_1, ..., x_m \in X$ with $m = \sqrt{n}$ such that for all scalars $(a_j)_{j\leqslant m}$

$$\max_{j\leqslant m} |a_j| \leqslant \left\| \sum_{j\leqslant m} a_j x_j \right\|_X \leqslant c \left(\sum_{j\leqslant m} a_j^2 \right)^{\frac{1}{2}},$$

where c is a universal constant. Bourgain-Szarek [3] proved that the previous statement holds for m proportional to n, and called the result "the proportional Dvoretzky-Rogers factorization":

Theorem B (Proportional Dvoretzky-Rogers factorization)**.** *Let X be an n-dimensional Banach space. $\forall \varepsilon \in (0, 1)$, there exist $x_1, ..., x_k \in X$ with $k \geqslant [(1 - \varepsilon)n]$ such that for all scalars $(a_j)_{j\leqslant k}$*

$$\max_{j\leqslant k} |a_j| \leqslant \left\| \sum_{j\leqslant k} a_j x_j \right\|_X \leqslant c(\varepsilon) \left(\sum_{j\leqslant k} a_j^2 \right)^{\frac{1}{2}},$$

where $c(\varepsilon)$ is a constant depending on ε. Equivalently, the identity operator $i_{2,\infty} : l_2^k \longrightarrow l_\infty^k$ can be written $i_{2,\infty} = \alpha \circ \beta$ with $\beta : l_2^k \longrightarrow X, \alpha : X \longrightarrow l_\infty^k$ and $\|\alpha\| \cdot \|\beta\| \leqslant c(\varepsilon)$.

Finding the right dependence on ε is an important problem and the optimal result is not known yet. In [17], Szarek showed that one cannot hope for a dependence better than $c\varepsilon^{-\frac{1}{10}}$. Szarek-Talagrand [18] proved that the previous result holds with $c(\varepsilon) = c\varepsilon^{-2}$ and in [7] and [8] Giannopoulos improved the dependence to get $c\varepsilon^{-\frac{3}{2}}$ and $c\varepsilon^{-1}$. In all these results, a factorization for the identity operator $i_{1,2} : l_1^k \longrightarrow l_2^k$ was proven and by duality the factorization for $i_{2,\infty}$ was deduced. The previous proofs used some geometric results, technical combinatorics and Grothendieck's factorization theorem. Here we present a direct proof using Theorem 1.1 which allows us to recover the best known dependence on ε and improve the universal constant involved.

Note that Theorem B can be formulated with symmetric convex bodies. In [13], Litvak and Tomczak-Jaegermann proved a nonsymmetric version of the proportional Dvoretzky-Rogers factorization:

Theorem C (Litvak-Tomczak-Jaegermann)**.** *Let $K \subset \mathbb{R}^n$ be a convex body, such that B_2^n is the ellipsoid of minimal volume containing K. Let $\varepsilon \in (0, 1)$ and set $k = [(1 - \varepsilon)n]$. There exist vectors $y_1, y_2, ..., y_k$ in K, and an orthogonal projection P in $\mathbb{R}^n$ with rank $P \geqslant k$ such that for all scalars $t_1, ..., t_k$*

$$c\varepsilon^3 \left(\sum_{j=1}^k |t_j|^2 \right)^{\frac{1}{2}} \leqslant \left\| \sum_{j=1}^k t_j P y_j \right\|_{PK} \leqslant \frac{6}{\varepsilon} \sum_{j=1}^k |t_j|,$$

where $c > 0$ is a universal constant.

Using again Theorem 1.1 combined with some tools developped in [3] and [13], we will be able to improve the dependence on ε in the previous statement.

3.1. **The symmetric case.** Let us start with the original proportional Dvoretzky-Rogers factorization. We will prove the following:

Theorem 3.2. *Let X be an n-dimensional Banach space. $\forall \varepsilon \in (0, 1)$, there exist $x_1, ..., x_k \in X$ with $k \geqslant [(1 - \varepsilon)^2 n]$ such that for all scalars $(a_j)_{j\leqslant m}$*

$$\varepsilon \left(\sum_{j\leqslant k} a_j^2 \right)^{\frac{1}{2}} \leqslant \left\| \sum_{j\leqslant k} a_j x_j \right\|_X \leqslant \sum_{j\leqslant k} |a_j|$$

Equivalently, the identity operator $i_{1,2} : l_1^k \longrightarrow l_2^k$ *can be written as* $i_{1,2} = \alpha \circ \beta$*, where* $\beta : l_1^k \longrightarrow X$*,* $\alpha : X \longrightarrow l_2^k$ *and* $\|\alpha\| \cdot \|\beta\| \leqslant \varepsilon^{-1}$.

Proof. Without loss of generality, we may assume that $X = (\mathbb{R}^n, \|\cdot\|_X)$ and B_2^n is the ellipsoid of minimal volume containing B_X. By John's theorem [10] there exist $x_1, ..., x_m$ contact points of B_X with B_2^n ($\|x_j\|_X = \|x_j\|_{X^*} = \|x_j\|_2 = 1$) and positive scalars $c_1, ..., c_m$ such that

$$Id = \sum_{j \leqslant m} c_j x_j x_j^t \quad \text{and} \quad \|\cdot\|_2 \leqslant \|\cdot\|_X$$

Let $U = \left(\sqrt{c_1}x_1, ..., \sqrt{c_m}x_m\right)$ be the $n \times m$ rectangular matrix whose columns are $\sqrt{c_j}x_j$ and denote $D = diag(\sqrt{c_1}, ..., \sqrt{c_m})$ the $m \times m$ diagonal matrix with $\sqrt{c_j}$ on its diagonal. Let $\varepsilon < 1$, applying Theorem 1.1 to U and D, we find $\sigma \subset \{1, ..., m\}$ such that

$$k = |\sigma| \geqslant \left[(1-\varepsilon)^2 n\right]$$

and for all $a = (a_j)_{j \leqslant m}$

$$\left\|U_\sigma D_\sigma^{-1} a\right\|_2 = \left\|\sum_{j \in \sigma} a_j x_j\right\|_2 \geqslant \varepsilon \left(\sum_{j \in \sigma} |a_j|^2\right)^{\frac{1}{2}} \tag{13}$$

To simplify the notations, we may assume that $\sigma = \{1, \dots, k\}$. Denote P the orthogonal projection of X onto $Y = \operatorname{span}\{(x_j)_{j \leqslant k}\}$. Now note that (13) guarantees that the $(x_j)_{j \leqslant k}$ are linearly independent and therefore that P is of rank k. Define T and β as follows:

$$\begin{array}{ccccc} \beta & : & l_1^k & \to & X \\ & & e_j & \mapsto & x_j \quad \text{for } j \leqslant k \end{array} \qquad \text{and} \qquad \begin{array}{ccccc} T & : & Y & \to & l_2^k \\ & & x_j & \mapsto & e_j \quad \text{for } j \leqslant k \end{array}$$

and write $\alpha = TP$. For $a = (a_j)_{j \leqslant k} \in \mathbb{R}^k$, by the triangle inequality we have

$$\|\beta(a)\|_X = \left\|\sum_{j \leqslant k} a_j x_j\right\|_X \leqslant \sum_{j \leqslant k} |a_j| \cdot \|x_j\|_X = \sum_{j \leqslant k} |a_j|,$$

and therefore $\|\beta\| \leqslant 1$. Now let $x \in X$, then $Px \in Y$ and one can write $Px = \sum_{j \leqslant k} a_j x_j$. Using (13) we get

$$\|\alpha(x)\|_2 = \left(\sum_{j \leqslant k} a_j^2\right)^{\frac{1}{2}} \leqslant \frac{1}{\varepsilon}\left\|\sum_{j \leqslant k} a_j x_j\right\|_2 = \frac{1}{\varepsilon}\|Px\|_2 \leqslant \frac{1}{\varepsilon}\|x\|_2 \leqslant \frac{1}{\varepsilon}\|x\|_X,$$

and therefore $\|\alpha\| \leqslant \varepsilon^{-1}$ which finishes the proof. □

As a direct application of the previous result, we have

Corollary 3.3. *Let* X *be an* n*-dimensional Banach space. For any* $\varepsilon \in (0,1)$*, there exists* Y *a subspace of* X *of dimension* $k \geqslant [(1-\varepsilon)^2 n]$ *such that* $d(Y, l_1^k) \leqslant \frac{\sqrt{n}}{\varepsilon}$.

3.2. **The nonsymmetric case.** Let us now turn to the nonsymmetric version of Theorem 3.2. We will prove the following:

Theorem 3.4. *Let* $K \subset \mathbb{R}^n$ *be a convex body, such that* B_2^n *is the ellipsoid of minimal volume containing* K*.* $\forall \varepsilon \in (0,1)$*, there exist* $x_1, ..., x_k$ *with* $k \geqslant [(1-\varepsilon)n]$ *contact points and there exists* P *an orthogonal projection of rank* $\geqslant k$ *such that for all* $(a_j)_{j \leqslant k}$

$$\frac{\varepsilon^2}{16}\left(\sum_{j=1}^k |a_j|^2\right)^{\frac{1}{2}} \leqslant \left\|\sum_{j=1}^k a_j P x_j\right\|_{PK} \leqslant \frac{4}{\varepsilon}\sum_{j=1}^k |a_j|$$

Proof. By John's Theorem [10], we get an identity decomposition in $\mathbb{R}^n$

$$Id = \sum_{j\leqslant m} c_j x_j x_j^t \quad \text{and} \quad \sum_{j\leqslant m} c_j x_j = 0$$

where $x_1, ..., x_m$ are contact points of K and B_2^n and $(c_j)_{j\leqslant m}$ positive scalars. Note that we will not use the second assertion i.e the fact that $\sum_{j\leqslant m} c_j x_j = 0$.

As before, take $U = \left(\sqrt{c_1}x_1, ..., \sqrt{c_m}x_m\right)$ the $n \times m$ rectangular matrix whose columns are $\sqrt{c_j}x_j$. Denote $D = diag(\sqrt{c_1}, ..., \sqrt{c_m})$ the $m \times m$ diagonal matrix with $\sqrt{c_j}$ on its diagonal. Applying Theorem 1.1 to U and D with $\frac{\varepsilon}{4}$, we find $\sigma_1 \subset \{1, ..., m\}$ such that

$$s = |\sigma_1| \geqslant \left(1 - \frac{\varepsilon}{4}\right)^2 n \geqslant (1 - \frac{\varepsilon}{2})n$$

and for all $a = (a_j)_{j\leqslant m}$

$$\left\|U_{\sigma_1} D_{\sigma_1}^{-1} a\right\|_2 = \left\|\sum_{j\in\sigma_1} a_j x_j\right\|_2 \geqslant \frac{\varepsilon}{4}\left(\sum_{j\in\sigma_1} |a_j|^2\right)^{\frac{1}{2}} \tag{14}$$

Define $Y = \mathrm{span}\{x_j\}_{j\in\sigma_1}$. We will now use the argument of Litvak and Tomczak-Jaegermann to construct the projection P. First partition σ_1 into $\left[\frac{\varepsilon}{2}s\right]$ disjoint subsets A_l of equal size. Clearly

$$|A_l| \leqslant \left[\frac{s}{[\frac{\varepsilon}{2}s]}\right] + 1 \leqslant \left[\frac{2}{\varepsilon} \cdot \frac{\frac{\varepsilon}{2}s}{[\frac{\varepsilon}{2}s]}\right] + 1 \leqslant \left[\frac{4}{\varepsilon}\right] + 1$$

Let $z_l = \sum_{i\in A_l} x_i$ and take $P : Y \longrightarrow Y$ the orthogonal projection onto $\mathrm{span}\{z_l\}^\perp$. For every l, we have $Pz_l = 0$ so that for $j \in A_l$ we can write

$$-Px_j = \sum_{i\in A_l, i\neq j} Px_i = (|A_l| - 1) \cdot \frac{1}{|A_l| - 1} \sum_{i\in A_l, i\neq j} Px_i$$

We deduce that for every l and every $j \in A_l$, we have $-Px_j \in (|A_l| - 1)\, PK \subset \frac{4}{\varepsilon}PK$.

Let $T : \mathbb{R}^{|\sigma_1|} \longrightarrow Y$ a linear operator defined by $Te_j = x_j$ for all $j \in \sigma_1$, where $(e_j)_{j\in\sigma_1}$ denotes the canonical basis of $\mathbb{R}^{|\sigma_1|}$ and Y is equipped with the euclidean norm. Since $(x_j)_{j\leqslant s}$ are linearly independent, T is an isomorphism. Moreover, by (14), we have $\|T^{-1}\| \leqslant \frac{4}{\varepsilon}$. Take $P' = T^{-1}PT$ and P'' the orthogonal projection onto $\mathrm{Im}P'$. It is easy to check that $P''P' = P'$ and

$$k = \mathrm{rank}P'' = \mathrm{rank}P \geqslant \left(1 - \frac{\varepsilon}{2}\right) s \geqslant (1 - \varepsilon)\, n$$

For all scalars $(a_j)_{j\in\sigma_1}$,

$$\begin{aligned}\left\|\sum_{j\in\sigma_1} a_j P x_j\right\|_2 &= \left\|\sum_{j\in\sigma_1} a_j PTe_j\right\|_2 \\ &= \left\|\sum_{j\in\sigma_1} T\left(a_j P' e_j\right)\right\|_2 \\ &\geqslant \frac{1}{\|T^{-1}\|}\cdot\left\|\sum_{j\in\sigma_1} a_j P' e_j\right\|_2 \\ &\geqslant \frac{\varepsilon}{4}\cdot\left\|\sum_{j\in\sigma_1} a_j P'' e_j\right\|_2\end{aligned}$$

Now take $U' = (P''e_1, ..., P''e_s)$ the $s\times s$ matrix whose columns are $(P''e_j)$. Apply Theorem 1.1 with U' and Id as diagonal matrix and $\frac{\varepsilon}{4}$ as parameter, then there exists $\sigma \subset \sigma_1$ of size

$$|\sigma| \geqslant \left(1-\frac{\varepsilon}{4}\right)^2 s \geqslant (1-\varepsilon)n$$

such that for all scalars $(a_j)_{j\in\sigma}$,

$$\left\|\sum_{j\in\sigma} a_j P'' e_j\right\|_2 \geqslant \frac{\varepsilon}{4}\left(\sum_{j\in\sigma}|a_j|^2\right)^{\frac{1}{2}}$$

This gives us the following

$$\left\|\sum_{j\in\sigma} a_j P x_j\right\|_2 \geqslant \frac{\varepsilon}{4}\cdot\left\|\sum_{j\in\sigma} a_j P'' e_j\right\|_2 \geqslant \frac{\varepsilon^2}{16}\left(\sum_{j\in\sigma}|a_j|^2\right)^{\frac{1}{2}}$$

On the other hand, since $K \subset B_2^n$ we have $PK \subset B_2^k$ and therefore

$$\left\|\sum_{j\in\sigma} a_j P x_j\right\|_2 \leqslant \left\|\sum_{j\in\sigma} a_j P x_j\right\|_{PK}$$

Denoting $A = -PK \cap PK$ which is a centrally symmetric convex body and using the fact that $-Py_j \in \frac{4}{\varepsilon}PK$ alongside the triangle inequality, one can write

$$\left\|\sum_{j\in\sigma} a_j P x_j\right\|_A \leqslant \frac{4}{\varepsilon}\sum_{j\in\sigma}|a_j|$$

Finally, we have

$$\frac{\varepsilon^2}{16}\left(\sum_{j\in\sigma}|a_j|^2\right)^{\frac{1}{2}} \leqslant \left\|\sum_{j\in\sigma} a_j P x_j\right\|_2 \leqslant \left\|\sum_{j\in\sigma} a_j P x_j\right\|_{PK} \leqslant \left\|\sum_{j\in\sigma} a_j P x_j\right\|_A \leqslant \frac{4}{\varepsilon}\sum_{j\in\sigma}|a_j|$$

□

One can interpret the previous result geometrically as follows:

Corollary 3.5. *Let $K \subset \mathbb{R}^n$ be a convex body such that B_2^n is the ellipsoid of minimal volume containing K. $\forall \varepsilon \in (0,1)$, there exists P an orthogonal projection of rank $k \geqslant [(1-\varepsilon)n]$ such that*

$$\frac{\varepsilon}{4}B_1^k \subset PK \subset \frac{16}{\varepsilon^2}B_2^k.$$

Moreover, $d(PK, B_1^k) \leqslant \frac{64\sqrt{n}}{\varepsilon^3}$.

By duality, this means that there exists a subspace $E \subset \mathbb{R}^n$ *of dimension* $k \geqslant [(1-\varepsilon)n]$ *such that*

$$\frac{\varepsilon^2}{16} B_2^k \subset K \cap E \subset \frac{4}{\varepsilon} B_\infty^k.$$

Moreover, $d(K \cap E, B_\infty^k) \leqslant \frac{64\sqrt{n}}{\varepsilon^3}$.

3.3. **Estimate of the Banach-Mazur distance to the Cube.** In [3], Bourgain-Szarek showed how to estimate the Banach-Mazur distance to the cube once a proportional Dvoretzky-Rogers factorization is proven. This technique was again used in [7] and [18]. Since we are able to obtain a proportional Dvoretzky-Rogers factorization with a better constant, using the same argument we will recover the best known asymptotic for the Banach-Mazur distance to the cube and improve the constants involved. Let us start defining

$$R_\infty^n = \max \{ d(X, l_\infty^n) / \ X \in \mathbb{BM}_n \}$$

Similarly one can define R_1^n, and since the Banach-Mazur distance is invariant by duality then $R_1^n = R_\infty^n$. It follows from John's theorem [10] that the diameter of $\mathbb{BM}_n$ is less than n and therefore a trivial estimate is $R_\infty^n \leqslant n$. In [17], Szarek showed the existence of an n-dimensional Banach space X such that $d(X, l_\infty^n) \geqslant c\sqrt{n}\log(n)$. Bourgain-Szarek proved in [3] that $R_\infty^n \leqslant o(n)$ while Szarek-Talagrand [18] and Giannopoulos [7] improved this upper bound to $cn^{\frac{7}{8}}$ and $cn^{\frac{5}{6}}$ respectively. Here, we will prove the following estimate:

Theorem 3.6. *Let X be an n-dimensional Banach space. Then*

$$d(X, l_1^n) \leqslant 2^{\frac{5}{6}} \sqrt{n} \cdot d(X, l_2^n)^{\frac{2}{3}}.$$

Since by John's theorem [10], for any $X \in \mathbb{BM}_n$ we have $d(X, l_2^n) \leqslant \sqrt{n}$ then we get the following:

Corollary 3.7. $R_1^n = R_\infty^n \leqslant (2n)^{\frac{5}{6}}$.

Proof of Theorem 3.6. We denote $d_X = d(X, l_2^n)$. In order to bound $d(X, l_1^n)$, we need to define an isomorphism $T : l_1^n \longrightarrow X$ and estimate $\|T\| \cdot \|T^{-1}\|$. A natural way is to find a basis of X and then define T the operator which sends the canonical basis of $\mathbb{R}^n$ to this basis of X. The main idea is to find a "large" subspace Y of X which is "not too far" from l_1 (actually more is needed), then complement the basis of Y to obtain a basis of X. Finding the "large" subspace is the heart of the method and is given by the proportional Dvoretzky-Rogers factorization. The proof is mainly divided in four steps:

-First step: Place B_X into a "good" position.

Since the Banach-Mazur distance is invariant under linear transformation, we may change the position of B_X. Therefore without loss of generality we may assume that

$$\| \cdot \|_2 \leqslant \| \cdot \|_X \leqslant d_X \| \cdot \|_2$$

-Second step: Let $\varepsilon > 0$ and set $k = (1-2\varepsilon)n$. Apply Theorem 3.2 to find $x_1, ..., x_k$ in X such that for all scalars $(a_j)_{j \leqslant k}$

$$\varepsilon \left(\sum_{j \leqslant k} a_j^2 \right)^{\frac{1}{2}} \leqslant \left\| \sum_{j \leqslant k} a_j x_j \right\|_X \leqslant \sum_{j \leqslant k} |a_j| \tag{15}$$

Note that $(x_j)_{j \leqslant k}$ are linearly independent and are a good candidate to be part of the basis of X.

-Third step: To form a basis of X, we simply take $y_{k+1}, .., y_n$ an orthogonal basis of $\mathrm{span}\{(x_j)_{j \leqslant k}\}^\perp$

such that $\|y_j\|_2 = \frac{1}{d_X}$. By (15), we have

$$\forall j > k, \quad \|y_j\|_X \leqslant 1$$

-Fourth step Define $T : l_1^k \longrightarrow X$ by $T(e_j) = x_j$ if $j \leqslant k$ and $T(e_j) = y_j$ if $j > k$. Let $a = (a_j)_{j \leqslant n} \in \mathbb{R}^n$ and

$$Ta = \sum_{j=1}^{k} a_j x_j + \sum_{j=k+1}^{n} a_j y_j.$$

Then using the triangle inequality and (15), one can write

$$\|a\|_1 = \sum_{j \leqslant k} |a_j| + \sum_{j > k} |a_j| \geqslant \left\| \sum_{j \leqslant k} a_j x_j + \sum_{j > k} a_j y_j \right\|_X \geqslant \left\| \sum_{j \leqslant k} a_j x_j + \sum_{j > k} a_j y_j \right\|_2 .$$

We also have

$$\begin{aligned}
\|Ta\|_2 &\geqslant \left[\left\| \sum_{j \leqslant k} a_j x_j \right\|_2^2 + \left\| \sum_{j > k} a_j y_j \right\|_2^2 \right]^{\frac{1}{2}} \qquad \text{by orthogonality} \\
&\geqslant \left[\varepsilon^2 \sum_{j \leqslant k} a_j^2 + \sum_{j > k} a_j^2 \|y_j\|_2^2 \right]^{\frac{1}{2}} \\
&\geqslant \left[\frac{\varepsilon^2}{n} \left(\sum_{j \leqslant k} |a_j| \right)^2 + \frac{1}{d_X^2 (n-k)} \left(\sum_{j > k} |a_j| \right)^2 \right]^{\frac{1}{2}} \quad \text{by Cauchy-Shwarz} \\
&\geqslant \left[\frac{\varepsilon^2}{n} \left(\sum_{j \leqslant k} |a_j| \right)^2 + \frac{1}{2\varepsilon n d_X^2} \left(\sum_{j > k} |a_j| \right)^2 \right]^{\frac{1}{2}} \\
&\geqslant \frac{1}{\sqrt{2}} \left[\frac{\varepsilon}{\sqrt{n}} \sum_{j \leqslant k} |a_j| + \frac{1}{d_X \sqrt{2\varepsilon n}} \sum_{j > k} |a_j| \right] \\
&\geqslant \frac{1}{2^{\frac{5}{6}} \sqrt{n} d_X^{\frac{2}{3}}} \sum_{j=1}^{n} |a_j| \qquad \text{taking } \varepsilon = (\sqrt{2} d_X)^{-\frac{2}{3}}.
\end{aligned}$$

As a conclusion,

$$\frac{1}{2^{\frac{5}{6}} \sqrt{n} d_X^{\frac{2}{3}}} \|a\|_1 \leqslant \|Ta\|_X \leqslant \|a\|_1$$

and therefore $d(X, l_1^n) \leqslant 2^{\frac{5}{6}} \sqrt{n} d_X^{\frac{2}{3}}$ for all $X \in \mathbb{BM}_n$. $\square$

Remark 3.8. Here we are interested in high dimensional results; this is why the constant is not that important. If we want an estimate for "small" dimensions, then the value of the constant becomes important. In [7], Giannopoulos proved that $R_\infty^n \leqslant c n^{\frac{5}{6}}$ with $c = \frac{2^{\frac{7}{6}}}{(\sqrt{2}-1)^{\frac{1}{3}}} \sim 3,0116$, and thus his result becomes nontrivial when the dimension is larger than 747. On the other hand, our result becomes nontrivial whenever the dimension is bigger than 32. Moreover, if we are interested in small dimensions, we can obtain a better result by choosing ε in the last inequality in a different way: in fact we have chosen $\varepsilon = (2n)^{-\frac{1}{3}}$ (replacing d_X with $\sqrt{n}$) in the asymptotic regime, otherwise one just need to optimize on ε so that it satisfies $\frac{\varepsilon}{\sqrt{(1-\varepsilon)^2 n}} = \frac{1}{n\sqrt{1-(1-\varepsilon)^2}}$; then our result becomes nontrivial when the dimension is larger than 16. In [19], Taschuk has also obtained an estimate for the Banach-Mazur distance to the cube of "small"-dimensional spaces. Precisely, he proved the following

$$R_\infty^n \leqslant \sqrt{n^2 - 2n + 2 + \frac{2}{\sqrt{n+2} - 1}}$$

One can check that our result improves on that whenever the dimension is larger than 22.

4. Projection on coordinate subspaces

Given an $n \times m$ matrix U and an integer $k \leqslant m$, our aim is to find a coordinate projection of U of rank k which gives the best minimal operator norm among all coordinate projections. First results were obtained by Lunin [14], and a complete answer to this question was given by Kashin-Tzafriri [11] who proved the following:

Theorem D (Kashin-Tzafriri)**.** *Let U be an $n \times m$ matrix. Fix λ with $1/m \leqslant \lambda \leqslant \frac{1}{4}$. Then, there exists a subset ν of $\{1, \ldots, m\}$ of cardinality $|\nu| \geqslant \lambda m$ such that*

$$\|U_\nu\| \leqslant c\left(\sqrt{\lambda}\|U\|_2 + \frac{\|U\|_{\mathrm{HS}}}{\sqrt{m}}\right),$$

where $U_\nu = UP_\nu$ and P_ν denotes the coordinate projection onto $\mathbb{R}^\nu$.

The conclusion of the Theorem states that for a fixed $\lambda < \frac{1}{4}$ we have

$$\min_{\substack{\sigma \subset \{1,\ldots,m\} \\ |\sigma| = \lambda m}} \|U_\sigma\| \leqslant c\left(\sqrt{\lambda}\|U\| + \frac{\|U\|_{\mathrm{HS}}}{\sqrt{m}}\right), \tag{16}$$

and this estimate is optimal in the sense that the dependence on the parameters in the right hand side cannot be improved.

Kashin-Tzafriri's proof (see [21]) uses the selectors with some other probabilistic arguments and the Grothendieck's factorization Theorem. In [20], Tropp gave a randomized algorithm to realize Grothendieck's factorization theorem and therefore he was able to give a randomized algorithm to find the subset σ promised in Theorem C.

Our aim here is to give a deterministic algorithm to find the subset σ. Our method uses tools from the work of Batson-Spielman-Srivastava [2] and allows us to improve Kashin-Tzafriri's result by extending the size of the coordinate projection and getting better constants in the result.

Theorem 4.1. *Let U be an $n \times m$ matrix and let $1/m \leqslant \lambda \leqslant \eta < 1$. Then, there exists $\sigma \subset \{1, \ldots, m\}$ with $|\sigma| = k \geqslant \lambda m$ such that*

$$\|U_\sigma\| \leqslant \frac{1}{\sqrt{1-\lambda}}\left(\sqrt{\lambda + \eta}\|U\| + \sqrt{1 + \frac{\lambda}{\eta}}\frac{\|U\|_{\mathrm{HS}}}{\sqrt{m}}\right),$$

In particular,

$$\|U_\sigma\| \leqslant \frac{\sqrt{2}}{\sqrt{1-\lambda}}\left(\sqrt{\lambda}\|U\| + \frac{\|U\|_{\mathrm{HS}}}{\sqrt{m}}\right),$$

where U_σ denotes the selection of the columns of U with indices in σ.

Proof. We denote by $(e_j)_{j \leqslant m}$ the canonical basis of $\mathbb{R}^m$. Since

$$U_\sigma \cdot U_\sigma^t = \sum_{j \leqslant \sigma} (Ue_j) \cdot (Ue_j)^t,$$

our problem reduces to the question of estimating the largest eigenvalue of this sum of rank one matrices. We will follow the same procedure as in the proof of the restricted invertibility theorem: at each step, we would like to add a column of the original matrix and then study the evolution of the largest eigenvalue. However, it will be convenient for us to add suitable multiples of the columns of U in order to construct the l-th matrix; for each l we will choose a subset σ_k of cardinality $|\sigma_l| = l$ and consider the matrix $A_l = \sum_{j\in\sigma_l} s_j\,(Ue_j)\cdot(Ue_j)^t$ where $(s_j)_{j\in\sigma}$ will be positive numbers which will be suitably chosen. At the step l, the barrier will be denoted by u_l, namely the eigenvalues of A_l will be all smaller than u_l. The corresponding potential is $\psi(A_l,u_l) := \mathrm{Tr}\,(U^t(u_lI-A_l)^{-1}U)$. We set $A_0 = 0$, while u_0 will be determined later.

As we did before, at each step the value of the potential $\psi(A_l,u_l)$ will decrease so that we can continue the iteration, while the value of the barrier will increase by a constant δ, i.e. $u_{l+1} = u_l + \delta$. We will use a lemma which appears as Lemma 3.4 in [16]. We state it here in the notation introduced above.

Lemma 4.2. *Assume that $\lambda_{\max}(A_l) \leqslant u_l$. Let v be a vector in $\mathbb{R}^n$ satisfying*

$$F_l(v) := \frac{v^t(u_{l+1}I-A_l)^{-2}v}{\psi(A_l,u_l)-\psi(A_l,u_{l+1})}\|U\|^2 + v^t(u_{l+1}I-A_l)^{-1}v \leqslant \frac{1}{s}.$$

Then, if we define $A_{l+1} = A_l + svv^t$ we have

$$\lambda_{\max}(A_{l+1}) \leqslant u_{l+1} \quad \textit{and} \quad \psi(A_{l+1},u_{l+1}) \leqslant \psi(A_l,u_l).$$

Proof. Using Sherman-Morrison formula we have:

$$\begin{aligned}
\psi(A_{l+1},u_{l+1}) &= \mathrm{Tr}\left(U^t\left(u_{l+1}I-A_l-svv^t\right)U\right)\\
&= \mathrm{Tr}\left(U^t\left(u_{l+1}I-A_l\right)U\right) + \frac{sv^t(u_{l+1}I-A_l)^{-1}UU^t(u_{l+1}I-A_l)^{-1}v}{1-sv^t(u_{l+1}I-A_l)^{-1}v}\\
&\leqslant \psi(A_l,u_l) - (\psi(A_l,u_l)-\psi(A_l,u_{l+1})) + \frac{v^t(u_{l+1}I-A_l)^{-2}v}{\frac{1}{s}-v^t(u_{l+1}I-A_l)^{-1}v}\|U\|^2
\end{aligned}$$

Since $v^t(u_{l+1}I-A_l)^{-1}v < F_l(v)$ and $F_l(v) \leqslant \frac{1}{s}$ we deduce that the quantity above is finite. This implies that $\lambda_{\max}(A_{l+1}) < u_{l+1}$, since otherwise one would find $s' < s$ such that $\lambda_{\max}(A_l + s'vv^t) = u_{l+1}$ and therefore $\psi(A_l + s'vv^t, u_{l+1})$ would blow up which contradicts the fact that it is finite.
On the other hand, rearranging the inequality above using the fact that $F_l(v) \leqslant \frac{1}{s}$ we get $\psi(A_{l+1},u_{l+1}) \leqslant \psi(A_l,u_l)$. □

We write α for the initial potential, i.e. $\alpha = \frac{\|U\|_{\mathrm{HS}}^2}{u_0}$. Suppose that $A_l = \sum_{j\in\sigma_l} s_j\,(Ue_j)\cdot(Ue_j)^t$ is constructed so that $\psi(A_l,u_l) \leqslant \psi(A_{l-1},u_{l-1}) \leqslant \alpha$ and $\lambda_{\max}(A_l) \leqslant u_l$. We will now use Lemma 4.2 in order to construct A_{l+1}. To this end, we must find a vector Ue_j not chosen before and a scalar s_{l+1} so that $F_l(Ue_j) \leqslant \frac{1}{s_{l+1}}$, and then use the lemma. Since $(u_lI-A_l)^{-1}$ and $(u_{l+1}I-A_l)^{-1}$ are positive semidefinite, one can easily check that

$$(u_lI-A_l)^{-1} - (u_{l+1}I-A_l)^{-1} \succeq \delta(u_{l+1}I-A_l)^{-2}.$$

Therefore,

$$\mathrm{Tr}\left(U^t(u_{l+1}I-A_l)^{-2}U\right) \leqslant \frac{1}{\delta}\left(\psi(A_l,u_l)-\psi(A_l,u_{l+1})\right).$$

It follows that

$$\begin{aligned}
\sum_{j\notin\sigma_l} F_l(Ue_j) \leqslant \sum_{j\leqslant m} F_l(Ue_j) &= \frac{Tr\,(U^t(u_{l+1}I-A_l)^{-2}U)}{\psi(A_l,u_l)-\psi(A_l,u_{l+1})}\|U\|^2 + \psi(A_l,u_{l+1})\\
&\leqslant \frac{\|U\|^2}{\delta} + \alpha,
\end{aligned}$$

and therefore one can find $i \notin \sigma_l$ such that

$$F_l(Ue_i) \leqslant \frac{1}{|\sigma_l^c|}\left(\frac{\|U\|^2}{\delta}+\alpha\right) \leqslant \frac{1}{|\sigma_k^c|}\left(\frac{\|U\|^2}{\delta}+\alpha\right),$$

where k is the maximum number of steps (which is in our case λm).
We are going to choose all s_j equal to $s := \frac{(1-\lambda)m}{\alpha+\frac{\|U\|^2}{\delta}}$. By the previous lemma, it is sufficient to take $A_{l+1} = A_l + s\,(Ue_i)\cdot(Ue_i)^t$. After $k=\lambda m$ steps, we get $\sigma=\sigma_k$ such that

$$\begin{aligned}\lambda_{\max}\left(\sum_{j\in\sigma_k}(Ue_j)\cdot(Ue_j)^t\right) &\leqslant \frac{1}{s}(u_0+k\delta) = \frac{\alpha+\frac{\|U\|^2}{\delta}}{(1-\lambda)m}(u_0+k\delta)\\ &= \frac{1}{1-\lambda}\left[\frac{\|U\|_{\mathrm{HS}}^2}{m}+\lambda\|U\|^2+\lambda\|U\|_{\mathrm{HS}}^2\frac{\delta}{u_0}+\frac{\|U\|^2}{m}\frac{u_0}{\delta}\right]\end{aligned}$$

The result follows by taking $u_0 = \eta m\delta$. The second part of the theorem follows by taking $\lambda=\eta$. $\square$

Aknowledgement. I am grateful to my PhD advisor Olivier Guédon for many helpful discussions. I would also like to thank the doctoral school of Paris-Est for giving me the opportunity to visit the University of Athens and Apostolos Giannopoulos for his hospitality and his precious help.

References

[1] K. Ball. Ellipsoids of maximal volume in convex bodies. *Geom. Dedicata*, 41(2):241–250, 1992.
[2] J. D. Batson, D. A. Spielman, and N. Srivastava. Twice-Ramanujan sparsifiers. In *STOC'09—Proceedings of the 2009 ACM International Symposium on Theory of Computing*, pages 255–262. ACM, New York, 2009.
[3] J. Bourgain and S. J. Szarek. The Banach-Mazur distance to the cube and the Dvoretzky-Rogers factorization. *Israel J. Math.*, 62(2):169–180, 1988.
[4] J. Bourgain and L. Tzafriri. Invertibility of large submatrices with applications to the geometry of banach spaces and harmonic analysis. *Israel J. Math.*, 57:137–224, 1987.
[5] J. Diestel, H. Jarchow, and A. Tonge. *Absolutely summing operators*, volume 43 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.
[6] A. Dvoretzky and C. A. Rogers. Absolute and unconditional convergence in normed linear spaces. *Proc. Nat. Acad. Sci. U. S. A.*, 36:192–197, 1950.
[7] A. A. Giannopoulos. A note on the Banach-Mazur distance to the cube. In *Geometric aspects of functional analysis (Israel, 1992–1994)*, volume 77 of *Oper. Theory Adv. Appl.*, pages 67–73. Birkhäuser, Basel, 1995.
[8] A. A. Giannopoulos. A proportional Dvoretzky-Rogers factorization result. *Proc. Amer. Math. Soc.*, 124(1):233–241, 1996.
[9] E. D. Gluskin. The diameter of the Minkowski compactum is roughly equal to n. *Funktsional. Anal. i Prilozhen.*, 15(1):72–73, 1981.
[10] F. John. Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948*, pages 187–204. Interscience Publishers, Inc., New York, N. Y., 1948.
[11] B. Kashin and L. Tzafriri. Some remarks on the restriction of operators to coordinate subspaces. preprint.
[12] M. Ledoux and M. Talagrand. *Probability in Banach spaces*. Springer, Berlin, 1991.
[13] A. E. Litvak and N. Tomczak-Jaegermann. Random aspects of high-dimensional convex bodies. In *Geometric aspects of functional analysis*, volume 1745 of *Lecture Notes in Math.*, pages 169–190. Springer, Berlin, 2000.
[14] A. A. Lunin. On operator norms of submatrices. *Mat. Zametki*, 45(3):94–100, 128, 1989.
[15] D. A. Spielman and N. Srivastava. An elementary proof of the restricted invertibility theorem. *Israel J. Math.*, 190:83–91, 2012.
[16] N. Srivastava. *Spectral Sparsification and Restricted Invertibility*. PhD thesis, Yale University, Mar. 2010.
[17] S. J. Szarek. Spaces with large distance to l_∞^n and random matrices. *Amer. J. Math.*, 112(6):899–942, 1990.

[18] S. J. Szarek and M. Talagrand. An "isomorphic" version of the Sauer-Shelah lemma and the Banach-Mazur distance to the cube. In *Geometric aspects of functional analysis (1987–88)*, volume 1376 of *Lecture Notes in Math.*, pages 105–112. Springer, Berlin, 1989.

[19] S. Taschuk. The Banach-Mazur distance to the cube in low dimensions. *Discrete Comput. Geom.*, 46(1):175–183, 2011.

[20] J. A. Tropp. Column subset selection, matrix factorization, and eigenvalue optimization. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 978–986, Philadelphia, PA, 2009. SIAM.

[21] R. Vershynin. John's decompositions: selecting a large part. *Israel J. Math.*, 122:253–277, 2001.

PIERRE YOUSSEF,
UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE
LABORATOIRE D'ANALYSE ET MATHÉMATIQUES APPLIQUÉES,
5, BOULEVARD DESCARTES, CHAMPS SUR MARNE,
77454 MARNE-LA-VALLÉE, CEDEX 2, FRANCE
E-MAIL: PIERRE.YOUSSEF@UNIV-MLV.FR