



**HAL**  
open science

# Rank-Based Norms, Capra-Conjugacies and the Rank Function

Paul Barbier, Jean-Philippe Chancelier, Michel de Lara, Valentin Paravy

► **To cite this version:**

Paul Barbier, Jean-Philippe Chancelier, Michel de Lara, Valentin Paravy. Rank-Based Norms, Capra-Conjugacies and the Rank Function. 2023. hal-03240885v2

**HAL Id: hal-03240885**

**<https://hal.science/hal-03240885v2>**

Preprint submitted on 31 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Rank-Based Norms, CAPRA-Conjugacies and the Rank Function

Paul Barbier, Jean-Philippe Chancelier, Michel De Lara, Valentin Paravy,  
CERMICS, École des Ponts, Marne-la-Vallée, France

January 31, 2023

## Abstract

We consider the space of matrices, with given number of rows and of columns, equipped with the classic trace scalar product. With any matrix (source) norm, we associate a coupling, called Capra, between the space of matrices and itself. Then, we compute the Capra conjugate and biconjugate of the rank function. They are expressed in function of a sequence of rank-based norms, more precisely generalized  $r$ -rank and dual  $r$ -rank matrix norms associated with the matrix source norm. We deduce a lower bound of the rank function given by a variational formula which involves the generalized  $r$ -rank norms. In the case of the Frobenius norm, we show that the rank function is equal to the variational formula.

**Keywords.** rank function, matrix norm, rank-based norm, generalized convexity, Capra conjugacy

## 1 Introduction

The rank function is a well-known example of nonconvex and nonsmooth function over matrices (as it is not possible to cover all the references on such a large subject, we refer the reader to a small subset [8, 9] of the literature and to [13] which offers a kind of survey of the rank function). In this paper, we display a variational lower bound of the rank function that involves a sequence of suitable norms.

For this purpose, we introduce sequences of rank-based norms — more precisely, generalized  $r$ -rank and dual  $r$ -rank matrix norms — generated from any (source) norm. This construction, for matrices, can also be found in [12], but in the case of unitarily invariant source norms. With a general source norm, we also define a coupling between the space of matrices and itself, and we compute the biconjugate of the rank function under the associated conjugacy. We deduce a lower bound variational formula for the rank function which involves generalized  $r$ -rank norms. Moreover, when the source norm is the Frobenius norm,

we prove that the inequality is an equality. The conjugacy we use is not the Fenchel conjugacy. This latter has been used for instance in [14] to obtain convex lower envelopes of matrix functions of the form  $M \mapsto \varphi(\text{rk}(M)) + \|M - M_0\|^2$ , or in [12] to obtain convex lower envelopes of matrix functions of the form  $M \mapsto \varphi(\|M\|) + \delta_{\text{rk}(M) \leq r}$ , and then get convex low rank approximation of optimization problems with low rank solutions. This is not the approach we follow in this paper: we use a new CAPRA-conjugacy to analyze, and provide variational formulas for, the rank function, but we are not motivated (at this stage) by possible use in optimization under rank constraint.

The paper is organized as follows. In Sect. 2, we define and study rank-based norms. In Sect. 3, we introduce CAPRA-conjugacies and their relations with the rank function.

## 2 Rank-based norms

In §2.1, we fix notation. In §2.2, we define rank-based norms as, more precisely, generalized  $r$ -rank and dual  $r$ -rank matrix norms. In §2.3, we detail the case of unitarily invariant source matrix norms.

### 2.1 Notation

In all the paper, we consider two fixed positive integers  $m$  (number of rows) and  $n$  (number of columns), and we denote  $d = \min(m, n)$ . We use the notation  $[[k, l]] = \{k, k + 1, \dots, l - 1, l\}$  for any pair of integers such that  $k \leq l$ . We denote by  $\mathcal{M}_{m,n}$  the space of real matrices with  $m$  rows and  $n$  columns, by  $\text{rk} : \mathcal{M}_{m,n} \rightarrow \mathbb{N}$  the rank function (where  $\mathbb{N}$  is the set of non-negative integers) and by  $\mathcal{M}^{\leq r}$  (resp.  $\mathcal{M}^{=r}$ ) the subset of matrices of rank less than or equal to  $r$  (resp. equal to  $r$ ). We recall that the singular values of a matrix  $M \in \mathcal{M}_{m,n}$  are the square root of the (nonnegative) eigenvalues of the square matrix  $M^T M$ , and we denote by  $s(M) = \{s_i(M)\}_{i \in [[1, d]]} \in \mathbb{R}^d$  the vector composed of the singular values of  $M$  arranged in nonincreasing order, that is,

$$s(M) \in K = \{x \in \mathbb{R}^d \mid x_1 \geq \dots \geq x_d \geq 0\} = s(\mathcal{M}_{m,n}), \quad \forall M \in \mathcal{M}_{m,n}. \quad (1)$$

For any  $t \in \mathbb{N}^* = \mathbb{N} \setminus \{0\}$ , we denote by  $\mathcal{O}_t$  the group of orthogonal square  $t \times t$  matrices. It is established that, for any matrix  $M \in \mathcal{M}_{m,n}$ , there exists a *singular value decomposition* [3, p. 6]  $M = U \text{diag}(s(M)) V^T$  of the matrix  $M$ , where  $U \in \mathcal{O}_m$  and  $V \in \mathcal{O}_n$ . It is also readily proven that, for any matrix  $M \in \mathcal{M}_{m,n}$ , for any  $U \in \mathcal{O}_m$  and  $V \in \mathcal{O}_n$ , we have that  $s(M) = s(UMV)$ .

When equipped with the scalar product  $\mathcal{M}_{m,n}^2 \ni M, N \mapsto \text{Tr}(MN^T)$ ,  $\mathcal{M}_{m,n}$  is an Euclidean space which is in duality with itself. As we manipulate functions with values in  $\overline{\mathbb{R}} = [-\infty, +\infty]$ , we adopt the Moreau *lower and upper additions* [15] that extend the usual addition with  $(+\infty) \dot{+} (-\infty) = (-\infty) \dot{+} (+\infty) = -\infty$  or with  $(+\infty) \dot{+} (-\infty) = (-\infty) \dot{+} (+\infty) = +\infty$ . For any subset  $Y \subset \mathcal{M}_{m,n}$ ,  $\sigma_Y : \mathcal{M}_{m,n} \rightarrow \overline{\mathbb{R}}$  denotes the *support*

function of the subset  $Y$ :

$$\sigma_Y(M) = \sup_{N \in Y} \text{Tr}(MN^T), \quad \forall M \in \mathcal{M}_{m,n}. \quad (2)$$

A generic norm on the space  $\mathcal{M}_{m,n}$  of matrices will be denoted by  $\|\cdot\|$  and will be called *matrix norm*<sup>1</sup>. By contrast, a generic norm on the space  $\mathbb{R}^d$  of vectors will be denoted by  $\|\cdot\|$  and will be called *vector norm*. For any matrix norm  $\|\cdot\|$  on the space  $\mathcal{M}_{m,n}$ , we denote by  $\mathbb{B}_{\|\cdot\|} \subset \mathcal{M}_{m,n}$  and  $\mathbb{S}_{\|\cdot\|} \subset \mathbb{B}_{\|\cdot\|} \subset \mathcal{M}_{m,n}$  the associated unit ball and unit sphere. The dual norm  $\|\cdot\|_\star$  of the matrix norm  $\|\cdot\|$  is a matrix norm on the space  $\mathcal{M}_{m,n}$ , defined by  $\|\cdot\|_\star = \sigma_{\mathbb{B}_{\|\cdot\|}}$ .

## 2.2 Definition of generalized $r$ -rank and dual $r$ -rank matrix norms

To define rank-based norms, one could take inspiration from the following construction of vector norms as in [6, Definition 3.2] and in [5, Definition 3]. Given a vector norm  $\|\cdot\|$  on  $\mathbb{R}^d$ , one can define other norms as follows: for any subset  $K \subset \llbracket 1, d \rrbracket$  of indices and for any vector  $x$ , one denotes by  $x_K$  the vector that coincides with  $x$  for the indices in  $K$  and with null entries outside  $K$ ; by taking the supremum of the norm of all these vectors  $x_K$ , for a cardinality  $|K|$  of  $K$  smaller or equal to a fixed integer  $r$ , one obtains  $\max_{|K| \leq r} \|x_K\|$ . Such a construction indeed defines a norm, which has been studied in [6, 5]. Unfortunately, this procedure does not work with the rank, as we illustrate below.

Let  $\|\cdot\|$  be the  $\ell_1$ -norm on the space  $\mathcal{M}_{m,n}$  of matrices, that is,  $\|M\|$  is the sum of the modules of all the components of the matrix  $M \in \mathcal{M}_{m,n}$ . Define, for any matrix  $M \in \mathcal{M}_{m,n}$  and  $r \in \llbracket 1, d \rrbracket$ ,  $\|M\|_{(r)}^{\text{tn}} = \sup_{X \subset M, \text{rk}(X) \leq r} \|X\|$ , where  $X \subset M$  is a shorthand for matrices  $X$  of  $\mathcal{M}_{m,n}$  for which there exists  $K \subset \llbracket 1, m \rrbracket$  and  $L \subset \llbracket 1, n \rrbracket$  such that  $X$  coincides with  $M$ , except for entries  $X_{k,l} = 0$  for  $(k,l) \notin K \times L$ . We now show that the function  $\|\cdot\|_{(r)}^{\text{tn}}$  is not a norm by contradicting the triangular inequality. Indeed, consider  $m = n = 2$  and the matrix  $M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ . As  $\text{rk}(M) = 1$ , we get that  $\|M\|_{(1)}^{\text{tn}} = 4$ . We can write  $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

and we easily get that  $\left\| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\|_{(1)}^{\text{tn}} = \left\| \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\|_{(1)}^{\text{tn}} = 1$ . However, the triangular inequality

does not hold true as we have that  $4 = \|M\|_{(1)}^{\text{tn}} > \left\| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\|_{(1)}^{\text{tn}} + \left\| \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\|_{(1)}^{\text{tn}} = 1 + 1 = 2$ .

This is why we turn to the following definition, inspired by the properties of dual coordinate- $r$  vector norms in [6, Equation (16), Proposition 3.3]. This construction, for matrices, can also be found in [12] in the unitarily invariant norm case (see the discussion at the beginning of §2.3).

---

<sup>1</sup>In some books, the terminology *matrix norm* is reserved for submultiplicative norms over matrices, which is not the case here.

**Proposition 1.** Let  $\|\cdot\|$  be a norm on the space  $\mathcal{M}_{m,n}$  of matrices. We denote by  $\mathbb{B}_{\|\cdot\|}$  and  $\mathbb{S}_{\|\cdot\|}$  the associated unit ball and unit sphere, as well as, for any  $r \in \llbracket 0, d \rrbracket$ ,

$$\mathbb{B}_{\|\cdot\|}^{\leq r} = \mathbb{B}_{\|\cdot\|} \cap \mathcal{M}^{\leq r}, \quad \mathbb{B}_{\|\cdot\|}^{\overline{r}} = \mathbb{B}_{\|\cdot\|} \cap \mathcal{M}^{\overline{r}}, \quad \mathbb{S}_{\|\cdot\|}^{\leq r} = \mathbb{S}_{\|\cdot\|} \cap \mathcal{M}^{\leq r}, \quad \mathbb{S}_{\|\cdot\|}^{\overline{r}} = \mathbb{S}_{\|\cdot\|} \cap \mathcal{M}^{\overline{r}}. \quad (3)$$

The following expressions  $\|\cdot\|_{(r),\star}^{\text{rk}}$  define a nondecreasing sequence  $\left\{ \|\cdot\|_{(r),\star}^{\text{rk}} \right\}_{r \in \llbracket 1, d \rrbracket}$  of norms on  $\mathcal{M}_{m,n}$

$$\|N\|_{(r),\star}^{\text{rk}} = \sigma_{\mathbb{B}_{\|\cdot\|}^{\leq r}}(N) = \sup_{M \in \mathbb{B}_{\|\cdot\|}^{\leq r}} \text{Tr}(MN^T), \quad \forall N \in \mathcal{M}_{m,n}, \quad \forall r \in \llbracket 1, d \rrbracket, \quad (4)$$

which satisfy

$$\|\cdot\|_{(r),\star}^{\text{rk}} = \sigma_{\mathbb{B}_{\|\cdot\|}^{\overline{r}}} = \sigma_{\mathbb{S}_{\|\cdot\|}^{\leq r}} = \sigma_{\mathbb{S}_{\|\cdot\|}^{\overline{r}}}, \quad \forall r \in \llbracket 1, d \rrbracket. \quad (5)$$

*Proof.* The sequence  $\left\{ \|\cdot\|_{(r),\star}^{\text{rk}} \right\}_{r \in \llbracket 1, d \rrbracket}$  in (4) is nondecreasing since the sequence  $\left\{ \mathbb{B}_{\|\cdot\|}^{\leq r} \right\}_{r \in \llbracket 1, d \rrbracket}$  of unit balls in (3) is nondecreasing, as so is the sequence  $\left\{ \mathcal{M}^{\leq r} \right\}_{r \in \llbracket 1, d \rrbracket}$ .

First, we prove that  $\sigma_{\mathbb{S}_{\|\cdot\|}^{\overline{r}}} = \sigma_{\mathbb{S}_{\|\cdot\|}^{\leq r}}$ , for  $r \in \llbracket 1, d \rrbracket$ . For this purpose, we show that  $\overline{\mathbb{S}_{\|\cdot\|}^{\overline{r}}} = \mathbb{S}_{\|\cdot\|}^{\leq r}$ , where  $\overline{\cdot}$  denotes the topological closure. The inclusion  $\overline{\mathbb{S}_{\|\cdot\|}^{\overline{r}}} \subset \mathbb{S}_{\|\cdot\|}^{\leq r}$  is straightforward because it is well known [13, Theorem 2] that  $\overline{\mathcal{M}^{\overline{r}}} = \mathcal{M}^{\leq r}$ , from which we deduce that  $\overline{\mathbb{S}_{\|\cdot\|}^{\overline{r}}} = \overline{\mathbb{S}_{\|\cdot\|} \cap \mathcal{M}^{\overline{r}}} \subset \overline{\mathbb{S}_{\|\cdot\|}} \cap \overline{\mathcal{M}^{\overline{r}}} = \mathbb{S}_{\|\cdot\|} \cap \mathcal{M}^{\leq r}$ , the inclusion being a property of the topological closure.

To prove the reverse inclusion  $\mathbb{S}_{\|\cdot\|} \cap \mathcal{M}^{\leq r} \subset \overline{\mathbb{S}_{\|\cdot\|}^{\overline{r}}}$ , we consider  $M \in \mathbb{S}_{\|\cdot\|} \cap \mathcal{M}^{\leq r}$ . As  $\overline{\mathcal{M}^{\overline{r}}} = \mathcal{M}^{\leq r}$ , there exists a sequence  $\{M_n\}_{n \in \mathbb{N}}$  in  $\mathcal{M}^{\overline{r}}$  such that  $M_n \rightarrow M$  when  $n \rightarrow +\infty$ . Since  $M \in \mathbb{S}_{\|\cdot\|}$ , we can always suppose that  $M_n \neq 0$ , for all  $n \in \mathbb{N}$ . Therefore  $\frac{M_n}{\|M_n\|}$  is well defined, and when  $n \rightarrow +\infty$  we have that  $\frac{M_n}{\|M_n\|} \rightarrow \frac{M}{\|M\|} = M$  since  $M \in \mathbb{S}_{\|\cdot\|}$ . Now, for all  $n \in \mathbb{N}$ , on the one hand,  $\frac{M_n}{\|M_n\|} \in \mathcal{M}^{\overline{r}}$  and, on the other hand,  $\frac{M_n}{\|M_n\|} \in \mathbb{S}_{\|\cdot\|}$ . As a consequence, we get that the sequence  $\left\{ \frac{M_n}{\|M_n\|} \right\}_{n \in \mathbb{N}} \in \mathbb{S}_{\|\cdot\|} \cap \mathcal{M}^{\overline{r}}$ , and we conclude that the limit of the sequence  $M \in \overline{\mathbb{S}_{\|\cdot\|}^{\overline{r}}}$ . Thus, we have proven that  $\overline{\mathbb{S}_{\|\cdot\|}^{\overline{r}}} = \mathbb{S}_{\|\cdot\|}^{\leq r}$ , hence that  $\sigma_{\mathbb{S}_{\|\cdot\|}^{\leq r}} = \sigma_{\overline{\mathbb{S}_{\|\cdot\|}^{\overline{r}}}} = \sigma_{\mathbb{S}_{\|\cdot\|}^{\overline{r}}}$  by [2, Proposition 7.13].

Second, we prove that  $\sigma_{\mathbb{B}_{\|\cdot\|}^{\leq r}} = \sigma_{\mathbb{S}_{\|\cdot\|}^{\leq r}}$ . It is readily established that  $\mathbb{S}_{\|\cdot\|}^{\leq r} \subset \mathbb{B}_{\|\cdot\|}^{\leq r} \subset \text{co}\mathbb{S}_{\|\cdot\|}^{\leq r}$  (the convex hull of  $\mathbb{S}_{\|\cdot\|}^{\leq r}$ ) as any point in  $\mathbb{B}_{\|\cdot\|}^{\leq r} = \mathbb{B}_{\|\cdot\|} \cap \mathcal{M}^{\leq r}$  is the convex combination of a point and its opposite in  $\mathbb{S}_{\|\cdot\|}^{\leq r} = \mathbb{S}_{\|\cdot\|} \cap \mathcal{M}^{\leq r}$ . Therefore, by property [2, Proposition 7.13] of the support function (2), we get that  $\sigma_{\mathbb{B}_{\|\cdot\|}^{\leq r}} = \sigma_{\mathbb{S}_{\|\cdot\|}^{\leq r}}$ . By the same reasoning, we also obtain that  $\sigma_{\mathbb{B}_{\|\cdot\|}^{\overline{r}}} = \sigma_{\mathbb{S}_{\|\cdot\|}^{\overline{r}}}$  which, combined with the first part, gives  $\sigma_{\mathbb{B}_{\|\cdot\|}^{\leq r}} = \sigma_{\mathbb{S}_{\|\cdot\|}^{\leq r}} = \sigma_{\mathbb{S}_{\|\cdot\|}^{\overline{r}}} = \sigma_{\mathbb{B}_{\|\cdot\|}^{\overline{r}}}$ .

Third, we prove that (4) defines norms. We consider a fixed  $r \in \llbracket 1, d \rrbracket$ . As the set  $\mathbb{B}_{\|\cdot\|}^{\leq r}$  is easily seen to be bounded and symmetric,  $\|\cdot\|_{(r),\star}^{\text{rk}} = \sigma_{\mathbb{B}_{\|\cdot\|}^{\leq r}}$  is a 1-homogeneous subadditive function with values in  $[0, +\infty[$ . It remains to prove that, for any  $N \in \mathcal{M}_{m,n}$ ,  $\sigma_{\mathbb{B}_{\|\cdot\|}^{\leq r}}(N) = 0 \Leftrightarrow N = 0$ . For this purpose, we consider a matrix  $N \in \mathcal{M}_{m,n}$  which satisfies  $\sigma_{\mathbb{B}_{\|\cdot\|}^{\leq r}}(N) = 0$ , and

we prove that  $N = 0$ . We consider the singular value decomposition  $N = U \text{diag}(s(N))V^T$  of the matrix  $N$ . Defining  $M = U \text{diag}(s_1(N), 0, \dots, 0)V^T$ , the matrix  $M$  has rank less than or equal to 1. Thus, we obtain that

$$|s_1(N)|^2 = |s_1(M)|^2 = \text{Tr}(MN^T) \leq \|M\| \sup_{M' \in \mathbb{S}_{\|\cdot\|}^{\leq r}} \text{Tr}(M'^T N) = \|M\| \sigma_{\|\cdot\|}^{\leq r}(N) = 0,$$

hence that  $s_1(N) = 0$ . This implies that all the singular values of  $N$  are null because  $s_1(N)$  is the largest one. Hence, we get that  $N = 0$ .

This ends the proof.  $\square$

Now, we define rank-based norms as follows.

**Definition 2.** Let  $\|\cdot\|$  be a norm on the space  $\mathcal{M}_{m,n}$  of matrices, that we call source (matrix) norm. The matrix norms in the nondecreasing sequence  $\{\|\cdot\|_{(r),\star}^{\text{rk}}\}_{r \in \llbracket 1,d \rrbracket}$ , given by Proposition 1, are called generalized dual  $r$ -rank matrix norms. By taking their dual norms  $\|\cdot\|_{(r)}^{\text{rk}} = \left(\|\cdot\|_{(r),\star}^{\text{rk}}\right)_{\star}$ , we obtain a nonincreasing sequence  $\{\|\cdot\|_{(r)}^{\text{rk}}\}_{r \in \llbracket 1,d \rrbracket}$  of norms on  $\mathcal{M}_{m,n}$  called generalized  $r$ -rank matrix norms.

Notice that, by (4) for  $r = d$ , and then by taking the dual norms, we get that

$$\|\cdot\|_{(1),\star}^{\text{rk}} \leq \dots \leq \|\cdot\|_{(d),\star}^{\text{rk}} = \|\cdot\|_{\star} \quad \text{and} \quad \|\cdot\|_{(1)}^{\text{rk}} \geq \dots \geq \|\cdot\|_{(d)}^{\text{rk}} = \|\cdot\|. \quad (6)$$

When the source norm  $\|\cdot\|$  is unitarily invariant (see §2.3), the norms above have been introduced and studied in [12]. Thus, we provide an extension (hence, the term *generalized*) of the so-called *rank constrained dual norm* in [12, Equation (7)] to generalized dual  $r$ -rank matrix norm  $\|\cdot\|_{(r),\star}^{\text{rk}}$  in Equation (4) in Proposition 1, and of the so-called *low-rank inducing norm* in [12, Equation (8)] to generalized  $r$ -rank matrix norm  $\|\cdot\|_{(r)}^{\text{rk}}$  in Definition 2. This extension is justified as our main result — namely, a lower bound for the rank function in Theorem 7 — holds for any source norm, unitarily invariant or not, and involves generalized  $r$ -rank matrix norms. Some common norms are not unitarily invariant, such as the  $\ell_p$ -norm for  $p \neq 2$  (including the supremum norm when  $p = \infty$ ). However, we have not been able to obtain explicit formulas for generalized  $r$ -rank matrix norms in these special non unitarily invariant cases.

## 2.3 The case of unitarily invariant source matrix norms

As just said, our main result (variational lower bound of the rank function) does not require unitarily invariant norms. However, we devote this §2.3 to unitarily invariant source matrix norms for two reasons: to stress proximity and difference with [12]; to provide a special case where the inequality in the forthcoming Theorem 7 is an equality.

In §2.3.1, we provide background on unitarily invariant matrix norms. In §2.3.2, we make the link between generalized  $r$ -rank and dual  $r$ -rank matrix norms, on the one hand, and generalized coordinate and dual coordinate- $r$  norms and the  $\ell_0$  pseudonorm, on the other hand.

### 2.3.1 Background on unitarily invariant matrix norms

We recall that a *unitarily invariant norm* on  $\mathcal{M}_{m,n}$  is a matrix norm such that  $\| \|UMV\| \| = \| \|M\| \|$ , for any matrix  $M \in \mathcal{M}_{m,n}$  and orthogonal matrices  $U \in \mathcal{O}_m, V \in \mathcal{O}_n$ .

We recall that a *symmetric absolute norm* is a vector norm  $\|\cdot\|$  on  $\mathbb{R}^d$  which satisfies the following properties:  $\|\cdot\|$  is *absolute* in the sense that  $\| \| |x| \| \| = \| \|x\| \|$ , for any  $x \in \mathbb{R}^d$ , where  $|x| = (|x_1|, \dots, |x_d|)$ ;  $\|\cdot\|$  is *symmetric* (or *permutation invariant*), that is,  $\| \| (x_{\nu(1)}, \dots, x_{\nu(d)}) \| \| = \| \| (x_1, \dots, x_d) \| \|$ , for any  $x \in \mathbb{R}^d$  and for any permutation  $\nu$  of the indices in  $\llbracket 1, d \rrbracket$ . In the literature, a symmetric absolute norm is also often called a *symmetric gauge function* (this is the vocabulary used in [12]). These two notions are linked by the following property (see [3, Theorem IV.2.1]).

**Proposition 3** (Von Neumann). *A norm  $\| \| \cdot \| \|$  on the space  $\mathcal{M}_{m,n}$  of matrices is unitarily invariant if and only if there exists a symmetric absolute norm  $\|\cdot\|$  on  $\mathbb{R}^d$  such that*

$$\| \| \cdot \| \| = \|\cdot\| \circ s \text{ that is, } \| \| M \| \| = \left\| \| (s_1(M), \dots, s_d(M)) \| \|, \forall M \in \mathcal{M}_{m,n}. \quad (7)$$

In that case, one has the following relation between dual norms

$$\| \| \cdot \| \|_{\star} = \|\cdot\|_{\star} \circ s. \quad (8)$$

We call Equations (7) and (8) *factorization equations* as the unitarily invariant matrix norms  $\| \| \cdot \| \|$  and  $\| \| \cdot \| \|_{\star}$  are factorized by means of the symmetric absolute vector norms  $\|\cdot\|$  and  $\|\cdot\|_{\star}$ . The proof relies on the so-called *Von Neumann inequality trace theorem* [7]:

$$\sup_{U \in \mathcal{O}_m, V \in \mathcal{O}_n} \text{Tr}(UMVN^T) = \langle s(M), s(N) \rangle, \forall M, N \in \mathcal{M}_{m,n}, \quad (9)$$

where  $\langle \cdot, \cdot \rangle$  is the scalar product on  $\mathbb{R}^d$ .

### 2.3.2 Links with generalized coordinate- $r$ norms and the $\ell_0$ pseudonorm

In [12, Lemma 3], it is shown that, when the source norm  $\| \| \cdot \| \|$  is unitarily invariant, the rank constrained dual norms (corresponding to the generalized  $r$ -rank matrix norms) are unitarily invariant, and a factorization equation like (8) is given in [12, Equation (9)]. As a consequence, the low-rank inducing norms (corresponding to the generalized dual  $r$ -rank matrix norms) are also unitarily invariant, and a factorization equation like (7) is given in [12, Equation (10)].

In Proposition 4, we will complete this result by providing additional characterizations of the factorization equations (7) and (8) for the generalized  $r$ -rank matrix norms and the generalized dual  $r$ -rank matrix norms.

For this purpose, we recall that the so-called  $\ell_0$  *pseudonorm* on  $\mathbb{R}^d$  is the function  $\ell_0 : \mathbb{R}^d \rightarrow \llbracket 0, d \rrbracket$  defined by

$$\ell_0(x) = \text{number of nonzero components of } x, \quad \forall x \in \mathbb{R}^d. \quad (10)$$

It is clear that rank and  $\ell_0$  pseudonorm are related through the relation

$$\text{rk}(M) = \ell_0(s(M)) , \quad \forall M \in \mathcal{M}_{m,n} . \quad (11)$$

In [5, Definition 2] (see also [6, Definition 3.2]) we introduce, for any vector norm  $\|\cdot\|$  on  $\mathbb{R}^d$ , the sequence  $\{\|\cdot\|_{(r)}^{\mathcal{R}}\}_{r \in \llbracket 1, d \rrbracket}$  of *generalized coordinate- $r$  norms* on  $\mathbb{R}^d$ , and the sequence  $\{\|\cdot\|_{(r),\star}^{\mathcal{R}}\}_{r \in \llbracket 1, d \rrbracket}$  of *generalized dual coordinate- $r$  norms*, their dual norms. We do not detail their definition as we will only need the forthcoming characterization (13): the norms  $\|\cdot\|_{(r),\star}^{\mathcal{R}}$ , for any  $r \in \llbracket 1, d \rrbracket$ , are related to the  $\ell_0$  pseudonorm by means of

$$\text{the level sets} \quad \ell_0^{\leq r} = \{x \in \mathbb{R}^d \mid \ell_0(x) \leq r\} , \quad \forall r \in \llbracket 0, d \rrbracket , \quad (12a)$$

$$\text{and the level curves} \quad \ell_0^= r = \{x \in \mathbb{R}^d \mid \ell_0(x) = r\} , \quad \forall r \in \llbracket 0, d \rrbracket , \quad (12b)$$

as it is proven in [6, Equation (16)] that, for any  $r \in \llbracket 1, d \rrbracket$ , the generalized dual coordinate- $r$  norm satisfies

$$\|\cdot\|_{(r),\star}^{\mathcal{R}} = \sigma_{\ell_0^{\leq r} \cap \mathbb{S}_{\|\cdot\|}} = \sigma_{\ell_0^= r \cap \mathbb{S}_{\|\cdot\|}} , \quad (13)$$

where  $\mathbb{S}_{\|\cdot\|} \subset \mathbb{R}^d$  denotes the unit sphere of the norm  $\|\cdot\|$ . The expression (13) is reminiscent, using (3), of the property (5) of the generalized rank-based norms.

In [5, Definition 3] we introduce, for any vector norm  $\|\cdot\|$  on  $\mathbb{R}^d$ , the sequence  $\{\|\cdot\|_{\star,(r)}^{\text{sn}}\}_{r \in \llbracket 1, d \rrbracket}$  of *generalized  $r$ -support dual norms* and the sequence  $\{\|\cdot\|_{\star,(r)}^{\text{tn}}\}_{r \in \llbracket 1, d \rrbracket}$  of *generalized top- $r$  dual norms*. As with generalized coordinate- $r$  norms, we do not detail their definition. However, we recall their expression when  $\|\cdot\| = \|\cdot\|_{\ell_p}$  is the  $\ell_p$  norm. We establish in [5, Table 1] that the associated generalized coordinate- $r$  norm  $\|\cdot\|_{(r)}^{\mathcal{R}}$  is the  $(p,r)$ -*support norm*  $\|x\|_{p,r}^{\text{sn}}$ , and the generalized dual coordinate- $r$  norm  $\|\cdot\|_{(r),\star}^{\mathcal{R}}$  is the *top- $(q,r)$  norm*  $\|\cdot\|_{q,r}^{\text{tn}}$ , where  $1/p + 1/q = 1$ . For  $y \in \mathbb{R}^d$ , letting  $\nu$  denote a permutation of  $\{1, \dots, d\}$  such that  $|y_{\nu(1)}| \geq |y_{\nu(2)}| \geq \dots \geq |y_{\nu(d)}|$ , we have that  $\|y\|_{q,r}^{\text{tn}} = \left(\sum_{l=1}^r |y_{\nu(l)}|^q\right)^{\frac{1}{q}}$ .

In the next Proposition 4, we show relationships between, on the one hand, the four above sequences — generalized coordinate- $r$  norms, generalized dual coordinate- $r$  norms, generalized  $r$ -support dual norms, generalized  $r$ -support dual norms — of vector norms (related to the  $\ell_0$  pseudonorm [5]) and, on the other hand, generalized  $r$ -rank matrix norms and the generalized dual  $r$ -rank matrix norms (related to the rank function [12]), through factorization equations like (7) and (8). As discussed at the beginning of §2.3.2, these relationships are new (in comparison with [12, Lemma 3]).

**Proposition 4.** *When the source norm  $\|\cdot\|$  on  $\mathcal{M}_{m,n}$  is unitarily invariant, with associated symmetric absolute norm  $\|\cdot\|$  on  $\mathbb{R}^d$  as in Proposition 3, then the generalized  $r$ -rank matrix norms  $\{\|\cdot\|_{(r)}^{\text{rk}}\}_{r \in \llbracket 1, d \rrbracket}$  and the generalized dual  $r$ -rank matrix norms  $\{\|\cdot\|_{(r),\star}^{\text{rk}}\}_{r \in \llbracket 1, d \rrbracket}$  (see Definition 2) are unitarily invariant and the factorization equations (7) and (8) are given by*



$$\|\cdot\|_{(r)}^{\text{rk}} = \|\cdot\|_{(r)}^{\mathcal{R}} \circ s = \|\cdot\|_{\star, (r)}^{\text{sn}} \circ s, \quad \forall r \in \llbracket 1, d \rrbracket, \quad (14a)$$

$$\|\cdot\|_{(r), \star}^{\text{rk}} = \|\cdot\|_{(r), \star}^{\mathcal{R}} \circ s = \|\cdot\|_{\star, (r)}^{\text{tn}} \circ s, \quad \forall r \in \llbracket 1, d \rrbracket. \quad (14b)$$

*Proof.* The first part of the Proposition — the generalized  $r$ -rank matrix norms  $\left\{ \|\cdot\|_{(r)}^{\text{rk}} \right\}_{r \in \llbracket 1, d \rrbracket}$  and the generalized dual  $r$ -rank matrix norms  $\left\{ \|\cdot\|_{(r), \star}^{\text{rk}} \right\}_{r \in \llbracket 1, d \rrbracket}$  are unitarily invariant — can be found in [12, Lemma 3]. This is why, we now turn to our contribution, namely Equation (14).

We suppose that the norm  $\|\cdot\|$  is unitarily invariant on  $\mathcal{M}_{m,n}$  and that  $\|\cdot\|$  is the associated symmetric absolute norm. For any  $N \in \mathcal{M}_{m,n}$ , we have that<sup>2</sup>

$$\begin{aligned} \|N\|_{(r), \star}^{\text{rk}} &= \sup_{\|M\|=1, \text{rk}(M) \leq r} \text{Tr}(MN^T) && \text{(by definition (4) and property (5))} \\ &= \sup_{\|M\|=1, \text{rk}(M) \leq r, U \in \mathcal{O}_m, V \in \mathcal{O}_n} \text{Tr}(UMVN^T) \end{aligned}$$

by change of variable  $M \rightarrow UMV$ , and using the properties that  $\|UMV\| = \|M\|$  and that  $\text{rk}(UMV) = \text{rk}(M)$

$$\begin{aligned} &= \sup_{\|M\|=1, \text{rk}(M) \leq r} \left\{ \sup_{U \in \mathcal{O}_m, V \in \mathcal{O}_n} \text{Tr}(UMVN^T) \right\} \\ &= \sup_{\|M\|=1, \text{rk}(M) \leq r} \langle s(N), s(M) \rangle \quad \text{(using Von Neumann inequality trace theorem (9))} \\ &= \sup_{\|s(M)\|=1, \ell_0(s(M)) \leq r} \langle s(N), s(M) \rangle \quad \text{(by (7) and (11))} \\ &= \sup_{\|x\|=1, \ell_0(x) \leq r, x \in K} \langle s(N), x \rangle \end{aligned}$$

as easily seen from the definition (1) of the cone  $K = s(\mathcal{M}_{m,n})$ , which is in one-to-one correspondence with the image of the singular values mapping  $s$

$$= \sup_{\|x\|=1, \ell_0(x) \leq r} \langle s(N), x \rangle$$

because  $s(N) \in K$ , hence the supremum is achieved on the cone  $K$  by the well-known Hardy-Littlewood-Pólya rearrangement inequality

$$= \sigma_{\ell_0 \leq r \cap \mathbb{S}_{\|\cdot\|}}(s(N))$$

---

<sup>2</sup>The proof starts like in [12, Proof to Lemma 3, A.1], but then goes on a different direction as we explicitly introduce the  $\ell_0$  pseudonorm.

by definition (12a) of the level sets  $\ell_0^{\leq r}$ , and as  $\mathbb{S}_{\|\cdot\|} \subset \mathbb{R}^d$  is the unit sphere of the norm  $\|\cdot\|$

$$= \|s(N)\|_{(r),\star}^{\mathcal{R}} \quad (\text{as } \|\cdot\|_{(r),\star}^{\mathcal{R}} = \sigma_{\ell_0^{\leq r} \cap \mathbb{S}_{\|\cdot\|}} \text{ by (13)})$$

Thus, we have proven that  $\|\cdot\|_{(r),\star}^{\text{rk}} = \|\cdot\|_{(r),\star}^{\mathcal{R}} \circ s$ , that is, the first equality in (14b).

Now, it is easily established by (13) that the vector norm  $\|\cdot\|_{(r),\star}^{\mathcal{R}}$  is a symmetric absolute norm (hence so is its dual norm  $\|\cdot\|_{(r)}^{\mathcal{R}}$ ). As a consequence, the first equality in (14) easily follows by using (8) (see [3, Proposition IV.2.11]), giving

$$\|\cdot\|_{(r)}^{\text{rk}} = \left( \|\cdot\|_{(r),\star}^{\text{rk}} \right)_{\star} = \left( \|\cdot\|_{(r),\star}^{\mathcal{R}} \circ s \right)_{\star} = \left( \|\cdot\|_{(r),\star}^{\mathcal{R}} \right)_{\star} \circ s = \|\cdot\|_{(r)}^{\mathcal{R}} \circ s. \quad (15)$$

Thus, we have proven the first equality in (14a).

There remains to prove the second equalities in (14a) and (14b). Because the symmetric absolute norm  $\|\cdot\|$  in Proposition 4 is a symmetric monotonic norm [1, Theorem 2], it is a so-called orthant-monotonic norm [10, 11]. As a consequence, by [5, Proposition 7], we get that, for any  $r \in \llbracket 1, d \rrbracket$ ,  $\|\cdot\|_{(r)}^{\mathcal{R}} = \|\cdot\|_{\star,(r)}^{\text{sn}}$  and  $\|\cdot\|_{(r),\star}^{\mathcal{R}} = \|\cdot\|_{\star,(r)}^{\text{tn}}$ . This gives the second equalities in (14a) and (14b), and ends the proof.  $\square$

As an illustration, when the source norm  $\|\cdot\|$  is the *Frobenius norm* given by

$$\|M\|_F = \sqrt{\text{Tr}(MM^T)} = \|s(M)\|_{\ell_2} = \sqrt{\sum_{i=1}^d s_i(M)^2}, \quad \forall M \in \mathcal{M}_{m,n}, \quad (16)$$

Equations (14a) and (14b) are

$$\|M\|_{F(r)}^{\text{rk}} = \|s(M)\|_{2,r}^{\text{sn}}, \quad \forall r \in \llbracket 1, d \rrbracket, \quad \forall M \in \mathcal{M}_{m,n}, \quad (17a)$$

$$\|N\|_{F(r),\star}^{\text{rk}} = \|s(N)\|_{2,r}^{\text{tn}} = \sqrt{\sum_{i=1}^r s_i(N)^2}, \quad \forall r \in \llbracket 1, d \rrbracket, \quad \forall N \in \mathcal{M}_{m,n}. \quad (17b)$$

### 3 CAPRA-conjugacies and the rank function

In §3.1, we adapt the definition of CAPRA-couplings in [6] to the case of matrices instead of vectors. In §3.2, we provide a variational lower bound of the rank function.

#### 3.1 CAPRA-couplings and conjugacies for matrices

We adapt the definition of CAPRA-couplings in [6] to the space  $\mathcal{M}_{m,n}$  of matrices.

**Definition 5.** Let  $\|\cdot\|$  be a source matrix norm on  $\mathcal{M}_{m,n}$ . The CAPRA-coupling  $\dot{\phi}$ , between  $\mathcal{M}_{m,n}$  and  $\mathcal{M}_{m,n}$ , associated with  $\|\cdot\|$ , is defined by:

$$\forall M, N \in \mathcal{M}_{m,n}, \quad \dot{\phi}(M, N) = \begin{cases} \frac{\text{Tr}(MN^T)}{\|M\|} & \text{if } M \neq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (18)$$

For any function  $F : \mathcal{M}_{m,n} \rightarrow \overline{\mathbb{R}}$ , the  $\dot{\zeta}$ -Fenchel-Moreau conjugate, or CAPRA-conjugate, is the function  $F^{\dot{\zeta}} : \mathcal{M}_{m,n} \rightarrow \overline{\mathbb{R}}$  defined by

$$F^{\dot{\zeta}}(N) = \sup_{M \in \mathcal{M}_{m,n}} \left( \dot{\zeta}(M, N) \dot{+} (-F(M)) \right), \quad \forall N \in \mathcal{M}_{m,n}, \quad (19)$$

and the  $\dot{\zeta}$ -Fenchel-Moreau biconjugate, or CAPRA-biconjugate, is the function  $F^{\dot{\zeta}\dot{\zeta}'} : \mathcal{M}_{m,n} \rightarrow \overline{\mathbb{R}}$  defined by

$$F^{\dot{\zeta}\dot{\zeta}'}(M) = \sup_{N \in \mathcal{M}_{m,n}} \left( \dot{\zeta}(M, N) \dot{+} (-F^{\dot{\zeta}}(N)) \right), \quad \forall M \in \mathcal{M}_{m,n}. \quad (20)$$

Then, we show below that the CAPRA-conjugate and biconjugate of the rank function are expressed in function of the generalized dual  $r$ -rank matrix norms  $\left\{ \|\cdot\|_{(r),\star}^{\text{rk}} \right\}_{r \in \llbracket 1, d \rrbracket}$  and the generalized  $r$ -rank matrix norms  $\left\{ \|\cdot\|_{(r)}^{\text{rk}} \right\}_{r \in \llbracket 1, d \rrbracket}$ , given by Definition 2. We do not give the proof as it is a simple adaptation, to the matrix case, of the proofs of [6, Propositions 4.4, 4.5].

**Proposition 6.** *Let  $\|\cdot\|$  be a source matrix norm on  $\mathcal{M}_{m,n}$ , and  $\dot{\zeta}$  be the associated CAPRA-coupling as in Definition 5.*

*For any function  $\varphi : \llbracket 0, d \rrbracket \rightarrow \overline{\mathbb{R}}$ , we have that (with the convention that  $\|\cdot\|_{(0),\star}^{\text{rk}} = 0$ )*

$$(\varphi \circ \text{rk})^{\dot{\zeta}}(N) = \sup_{i \in \llbracket 0, d \rrbracket} \left\{ \|\|N\|_{(i),\star}^{\text{rk}} - \varphi(i) \right\}, \quad \forall N \in \mathcal{M}_{m,n}, \quad (21)$$

*and, for any function  $\varphi : \llbracket 0, d \rrbracket \rightarrow \mathbb{R}_+$  (that is, with nonnegative finite values) and such that  $\varphi(0) = 0$ , we have that*

$$(\varphi \circ \text{rk})^{\dot{\zeta}\dot{\zeta}'}(M) = \frac{1}{\|\|M\|\|} \min_{\substack{M^{(1)} \in \mathcal{M}_{m,n}, \dots, M^{(d)} \in \mathcal{M}_{m,n} \\ \sum_{r=1}^d \|M^{(r)}\|_{(r)}^{\text{rk}} \leq \|M\| \\ \sum_{r=1}^d M^{(r)} = M}} \sum_{r=1}^d \varphi(r) \|\|M^{(r)}\|_{(r)}^{\text{rk}}, \quad \forall M \in \mathcal{M}_{m,n} \setminus \{0\}. \quad (22)$$

### 3.2 Variational lower bound and expression of the rank function

Now, thanks to Proposition 6, we obtain a variational lower bound of the rank function, and also a variational expression when the source norm is the Frobenius norm.

**Theorem 7.** *Let  $\|\cdot\|$  be a source norm on the space  $\mathcal{M}_{m,n}$  of matrices, with associated sequence  $\left\{ \|\cdot\|_{(r)}^{\text{rk}} \right\}_{r \in \llbracket 1, d \rrbracket}$  of generalized  $r$ -rank matrix norms as in Definition 2. Then, we have the following variational lower bound of the rank function*

$$\text{rk}(M) \geq \frac{1}{\|\|M\|\|} \min_{\substack{M^{(1)} \in \mathcal{M}_{m,n}, \dots, M^{(d)} \in \mathcal{M}_{m,n} \\ \sum_{r=1}^d \|M^{(r)}\|_{(r)}^{\text{rk}} \leq \|M\| \\ \sum_{r=1}^d M^{(r)} = M}} \sum_{r=1}^d r \|\|M^{(r)}\|_{(r)}^{\text{rk}}, \quad \forall M \in \mathcal{M}_{m,n} \setminus \{0\}. \quad (23)$$

Moreover, if the source norm is the Frobenius norm  $\|\cdot\|_F$  given by (16), the inequality in (23) is an equality:

$$\text{rk}(M) = \frac{1}{\|M\|_F} \min_{\substack{M^{(1)} \in \mathcal{M}_{m,n}, \dots, M^{(d)} \in \mathcal{M}_{m,n} \\ \sum_{r=1}^d \|M^{(r)}\|_F^{\text{rk}} \leq \|M\|_F \\ \sum_{r=1}^d M^{(r)} = M}} \sum_{r=1}^d r \|M^{(r)}\|_F^{\text{rk}}, \quad \forall M \in \mathcal{M}_{m,n} \setminus \{0\}. \quad (24)$$

*Proof.* From the expression (22) of  $\text{rk}^{\dot{C}\dot{C}'}$ , with  $\varphi$  the identity function, and from the (true for any coupling) inequality  $\text{rk} \geq \text{rk}^{\dot{C}\dot{C}'}$ , we readily deduce (23).

In the rest of the proof — which follows that of [4, Theorem 3.5] —  $\|\cdot\|$  denotes the Frobenius norm (16) (instead of  $\|\cdot\|_F$  to alleviate notation). We consider a fixed matrix  $M \in \mathcal{M}_{m,n} \setminus \{0\}$  and we are going to show that  $\text{rk}(M) = \text{rk}^{\dot{C}\dot{C}'}(M)$ . We denote by  $r = \text{rk}(M) \geq 1$  the rank of  $M$ . By the factorization Equation (17b), and as  $s_i(M) = 0 \iff i > r$ , we have that

$$\|M\|_{(k),\star}^{\text{rk}} = \|s(M)\|_{2,k}^{\text{tn}} = \sqrt{\sum_{i=1}^k s_i(M)^2} \begin{cases} = \sqrt{\sum_{i=1}^r s_i(M)^2} = \|M\|, & \forall k \geq r, \\ < \sqrt{\sum_{i=1}^r s_i(M)^2} = \|M\|, & \forall k < r. \end{cases} \quad (25)$$

We consider the function  $\phi : ]0, +\infty[ \rightarrow \mathbb{R}$  defined by

$$\phi(\lambda) = \frac{\text{Tr}(\lambda M M^T)}{\|M\|} - \sup_{k \in \llbracket 0, d \rrbracket} \left\{ \| \lambda M \|_{(k),\star}^{\text{rk}} - k \right\}, \quad \forall \lambda > 0, \quad (26)$$

and we will show that  $\lim_{\lambda \rightarrow +\infty} \phi(\lambda) = r$ . We have that

$$\phi(\lambda) = \lambda \|M\| - \sup_{k \in \llbracket 1, d \rrbracket} \left\{ \lambda \|M\|_{(k),\star}^{\text{rk}} - k \right\}$$

by definition (26) of  $\phi$ , by the convention that  $\|M\|_{(0),\star}^{\text{rk}} = 0$  and by  $\|M\|^2 = \text{Tr}(M M^T)$

$$\begin{aligned} &= \lambda \|M\|_{(r),\star}^{\text{rk}} + \inf \left\{ 0, - \sup_{k \in \llbracket 1, d \rrbracket} \left[ \lambda \|M\|_{(k),\star}^{\text{rk}} - k \right] \right\} \quad (\text{as } \|M\| = \|M\|_{(r),\star}^{\text{rk}} \text{ by (25)}) \\ &= \inf \left\{ \lambda \|M\|_{(r),\star}^{\text{rk}}, \lambda \|M\|_{(r),\star}^{\text{rk}} + \inf_{k \in \llbracket 1, d \rrbracket} \left( - \left[ \lambda \|M\|_{(k),\star}^{\text{rk}} - k \right] \right) \right\} \\ &= \inf \left\{ \lambda \|M\|_{(r),\star}^{\text{rk}}, \inf_{k \in \llbracket 1, d \rrbracket} \left( \lambda \left( \|M\|_{(r),\star}^{\text{rk}} - \|M\|_{(k),\star}^{\text{rk}} \right) + k \right) \right\} \\ &= \inf \left\{ \lambda \|M\|_{(r),\star}^{\text{rk}}, \inf_{k \in \llbracket 1, r-1 \rrbracket} \left( \lambda \left( \|M\|_{(r),\star}^{\text{rk}} - \|M\|_{(k),\star}^{\text{rk}} \right) + k \right), \right. \\ &\quad \left. \inf_{k \in \llbracket r, d \rrbracket} \left( \lambda \left( \|M\|_{(r),\star}^{\text{rk}} - \|M\|_{(k),\star}^{\text{rk}} \right) + k \right) \right\} \\ &= \inf \left\{ \lambda \|M\|_{(r),\star}^{\text{rk}}, \inf_{k \in \llbracket 1, r-1 \rrbracket} \left( \lambda \left( \|M\|_{(r),\star}^{\text{rk}} - \|M\|_{(k),\star}^{\text{rk}} \right) + k \right), r \right\} \end{aligned}$$

as  $\|M\|_{(k),\star}^{\text{rk}} = \|M\|_{(r),\star}^{\text{rk}}$  for  $k \geq r$  by (25). Let us show that the two first terms in the infimum go to  $+\infty$  when  $\lambda \rightarrow +\infty$ . The first term  $\lambda \|M\|_{(r),\star}^{\text{rk}}$  goes to  $+\infty$  because  $\|M\|_{(r),\star}^{\text{rk}} = \|M\| > 0$  by assumption ( $M \neq 0$ ). The second term  $\inf_{k \in \llbracket 1, r-1 \rrbracket} \left( \lambda \left( \|M\|_{(r),\star}^{\text{rk}} - \|M\|_{(k),\star}^{\text{rk}} \right) + k \right)$  also goes to  $+\infty$  because  $\text{rk}(M) = r$ , so that  $\|M\| = \|M\|_{(r),\star}^{\text{rk}} > \|M\|_{(k),\star}^{\text{rk}}$  for  $k \in \llbracket 1, r-1 \rrbracket$  as shown in (25). Therefore, we get that  $\lim_{\lambda \rightarrow +\infty} \phi(\lambda) = \inf\{+\infty, +\infty, r\} = r$ . This concludes the proof since

$$\begin{aligned} r = \lim_{\lambda \rightarrow +\infty} \phi(\lambda) &\leq \sup_{N \in \mathcal{M}_{m,n}} \left( \frac{\text{Tr}(MN^T)}{\|M\|} - \sup_{k \in \llbracket 0, d \rrbracket} \left\{ \|N\|_{(k),\star}^{\text{rk}} - k \right\} \right) \quad (\text{by definition (26) of } \phi) \\ &= \sup_{N \in \mathcal{M}_{m,n}} \left( \frac{\text{Tr}(MN^T)}{\|M\|} - \text{rk}^{\dot{\zeta}}(N) \right) \end{aligned}$$

by the formula (21) for the conjugate  $\text{rk}^{\dot{\zeta}}$

$$\begin{aligned} &= \text{rk}^{\dot{\zeta}'}(M) && (\text{by the biconjugate formula (20)}) \\ &\leq \text{rk}(M) && (\text{as } \text{rk}^{\dot{\zeta}'} \leq \text{rk}) \\ &= r. && (\text{by assumption}) \end{aligned}$$

Therefore, we have obtained that  $r = \text{rk}^{\dot{\zeta}'}(M) = \text{rk}(M)$ .

This ends the proof.  $\square$

## 4 Conclusion

In this paper, we have shown how to obtain a variational lower bound of the rank function (Theorem 7). Interestingly, the formula depends on a (source) matrix norm and on the derived generalized  $r$ -rank matrix norms, that we introduce (Definition 2). This is made possible by the versatility of the CAPRA-couplings, themselves depending on a matrix norm (Definition 5). Moreover, we show that the variational expression we obtain is equal to the rank function when the source norm is the Frobenius norm (Theorem 7).

Thus, we hope to offer a general framework to derive matrix norms suitable for optimization problems involving the rank function, as well as variational formulations.

## References

- [1] F. L. Bauer, J. Stoer, and C. Witzgall. Absolute and monotonic norms. *Numer. Math.*, 3:257–264, 1961.
- [2] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer-Verlag, New York, second edition, 2017.

- [3] R. Bhatia. *Matrix Analysis*. Springer-Verlag, New York, 1997.
- [4] J.-P. Chancelier and M. De Lara. Hidden convexity in the  $l_0$  pseudonorm. *Journal of Convex Analysis*, 28(1):203–236, 2021.
- [5] J.-P. Chancelier and M. De Lara. Capra-convexity, convex factorization and variational formulations for the  $l_0$  pseudonorm. *Set-Valued and Variational Analysis*, 30:597–619, 2022.
- [6] J.-P. Chancelier and M. De Lara. Constant along primal rays conjugacies and the  $l_0$  pseudonorm. *Optimization*, 71(2):355–386, 2022.
- [7] K. Fan and A. J. Hoffman. Some metric inequalities in the space of matrices. *Proceedings of the American Mathematical Society*, 6(1):111–116, 1955.
- [8] M. Fazel, H. Hindi, and S. Boyd. A rank minimization heuristic with application to minimum order system approximation. In *Proceedings of the 2001 American Control Conference. (Cat. No.01CH37148)*, volume 6, pages 4734–4739, 2001.
- [9] M. Fazel, H. Hindi, and S. Boyd. Rank minimization and applications in system theory. In *Proceedings of the 2004 American Control Conference*, volume 4, pages 3273–3278, 2004.
- [10] D. Gries. Characterization of certain classes of norms. *Numerische Mathematik*, 10:30–41, 1967.
- [11] D. Gries and J. Stoer. Some results on fields of values of a matrix. *SIAM Journal on Numerical Analysis*, 4(2):283–300, 1967.
- [12] C. Grussler and P. Giselsson. Low-rank inducing norms with optimality interpretations. *SIAM Journal on Optimization*, 28(4):3057–3078, Jan 2018.
- [13] J.-B. Hiriart-Urruty and H. Le. A variational approach of the rank function. *TOP: An Official Journal of the Spanish Society of Statistics and Operations Research*, 21(2):207–240, 2013.
- [14] V. Larsson and C. Olsson. Convex low rank approximation. *Int. J. Comput. Vision*, 120(2):194–214, nov 2016.
- [15] J. J. Moreau. Inf-convolution, sous-additivité, convexité des fonctions numériques. *J. Math. Pures Appl. (9)*, 49:109–154, 1970.