



**HAL**  
open science

# Convex relaxations of structured matrix factorizations

Francis Bach

► **To cite this version:**

| Francis Bach. Convex relaxations of structured matrix factorizations. 2013. hal-00861118

**HAL Id: hal-00861118**

**<https://hal.archives-ouvertes.fr/hal-00861118>**

Preprint submitted on 12 Sep 2013

**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.

# Convex relaxations of structured matrix factorizations

Francis Bach  
Sierra Project-team  
Département d'Informatique de l'École Normale Supérieure  
Paris, France  
`francis.bach@ens.fr`

September 12, 2013

## Abstract

We consider the factorization of a rectangular matrix  $X$  into a positive linear combination of rank-one factors of the form  $uv^\top$ , where  $u$  and  $v$  belongs to certain sets  $\mathcal{U}$  and  $\mathcal{V}$ , that may encode specific structures regarding the factors, such as positivity or sparsity. In this paper, we show that computing the optimal decomposition is equivalent to computing a certain gauge function of  $X$  and we provide a detailed analysis of these gauge functions and their polars. Since these gauge functions are typically hard to compute, we present semi-definite relaxations and several algorithms that may recover approximate decompositions with approximation guarantees. We illustrate our results with simulations on finding decompositions with elements in  $\{0, 1\}$ . As side contributions, we present a detailed analysis of variational quadratic representations of norms as well as a new iterative basis pursuit algorithm that can deal with inexact first-order oracles.

## 1 Introduction

Structured matrix factorization has many applications in various areas of science engineering, i.e., clustering and principal component analysis [22, 41], source separation [34, 23], signal processing [1], machine learning [38], and all domains where reduced representations are desired.

Without any structure, traditional principal component analysis may be solved exactly in polynomial time through a singular value decomposition. However, adding additional structure on the components  $U$  and  $V$  of the factorization of  $X = UV^\top$  (e.g., sparsity, non-negativity or discreteness) is most often done through alternating minimization (with respect to  $U$  and  $V$ ). While all steps are usually done through convex optimization, the problem is not jointly convex, and there are typically multiple local minima, and algorithms usually come with no convergence guarantees. For example, in presence of positivity constraints, the problem of non-negative matrix factorization (NMF) may not be solved in polynomial-time in general, and most algorithms for NMF (e.g., [34]) perform a form a block coordinate descent with no guarantees (see hardness results and particular situations of actual solvability in [3, 46]). In this paper, we follow a convex relaxation approach.

We impose some structure on each column of  $U$  and  $V$  and consider a general convex framework which amounts to computing a certain gauge function (such as a norm) at  $X$ , from which the decomposition may be obtained (for example, the nuclear norm leads to the usual singular value decomposition). This convex framework corresponds to removing any rank constraint on the decomposition and has appeared under various forms in the literature, as summing norms [30], decomposition norms [7] or (a special case of) atomic norms [17]. This is presented in details (equivalent representations, rotation-invariant cases, weighted nuclear norm formulations) in Section 4. An interesting aspect is that the gauge functions we consider are polar to generalizations of matrix norms, which are commonly used in many areas of applied mathematics, in particular in robust optimization and control [9].

The statistical and recovery properties of these norms and their relaxations have been studied in several contexts [17, 45]; in this paper, we focus on optimization aspects. The convex framework we introduce in Section 4 only lead to polynomial-time algorithms in few situations (e.g., the nuclear norm based on the singular value decomposition). In Section 5, we consider computable additional relaxations based on semi-definite programming. These may be used to compute the related gauge functions as well as their polar, with constant-factor approximation guarantees in some cases (see Section 5.1). The first setting where one can get dimension-independent guarantees has already been studied by [42, 10] and corresponds to gauge functions that have variational diagonal representations. We also consider a more general setting with dimension-dependent bounds.

A key practical problem is to obtain not only a lower-bound on the value of the gauge function, but also an *explicit* decomposition which preserves the approximation guarantees. We present in Section 6 iterative conditional gradient algorithms and their analysis, which extend existing results in several ways: (a) we obtain convergence guarantees even when the polar gauge function may be approximately computed with a *multiplicative* approximation ratio—earlier work [29] considers only additive approximations, (b) following [5, 27, 52], they may be applied to penalized versions of the problem, i.e., to solve a generalized basis pursuit problem, (c) under some additional assumptions, they may find approximate decompositions of  $X$  that converge linearly.

Finally, our framework relies on variational representations of gauge functions and in particular norms as maxima and minima of quadratic functions. These representations have been already used in several contexts (machine learning, signal processing, optimization, see, e.g., [6] and references therein). In this paper, we provide in Section 3 a thorough analysis of these decompositions (minimal and maximal representations, duality between lower and upper bounds, sufficient and necessary conditions for diagonal or rotation-invariant representations).

**Notation.** Given a positive integer  $d$ , we denote by  $\mathcal{S}_d$  the vector space of symmetric matrices, and by  $\mathcal{S}_d^+$  the subset of positive-semidefinite matrices. For  $x \in \mathbb{R}^d$  and  $p \in [1, +\infty]$ ,  $\|x\|_p$  denotes the  $\ell_p$ -norm of  $x$ , while for a matrix  $X \in \mathbb{R}^{n \times d}$ ,  $\|X\|_p$  denotes the  $\ell_p$ -norm of  $X$ , seen as a vector. That is, if  $\text{vec}(X) \in \mathbb{R}^{nd}$  denotes the vector obtained by stacking the columns of  $X$ ,  $\|X\|_p = \|\text{vec}(X)\|_p$ . The Frobenius norm of  $X$  is denoted  $\|X\|_F = \sqrt{\text{tr } X^\top X} = \|\text{vec}(X)\|_2$ , the nuclear norm (a.k.a. the trace norm) is denoted as  $\|X\|_*$ , and is equal to the sum of the singular values of  $X$ , while the operator norm

(largest singular value of  $X$ ) is denoted  $\|X\|_{\text{op}}$ . Finally,  $1_d$  denotes the vector in  $\mathbb{R}^d$  with all components equal to one.

## 2 Review of gauge function theory

In this section, we present relevant concepts and results from convex analysis. These tools are needed because the type of structure we want to impose go beyond what can be characterized by norms (such as positivity). See [47, 13] for more details on gauge functions and their properties.

**Gauge functions.** Given a closed convex set  $\mathcal{C} \subset \mathbb{R}^d$ , the gauge function  $\gamma_{\mathcal{C}}$  is the function

$$\gamma_{\mathcal{C}}(x) = \inf\{\lambda \geq 0, x \in \lambda\mathcal{C}\}.$$

The domain  $\text{dom}(\gamma_{\mathcal{C}})$  of  $\gamma_{\mathcal{C}}$  is the cone generated by  $\mathcal{C}$ , i.e.,  $\mathbb{R}_+\mathcal{C}$  (that is,  $\gamma_{\mathcal{C}}(x) < +\infty$  if and only if  $x \in \mathbb{R}_+\mathcal{C}$ ). The function  $\gamma_{\mathcal{C}}$  is equivalently defined as the homogenized version of the indicator function  $I_{\mathcal{C}}$  (with values 0 on  $\mathcal{C}$  and  $+\infty$  on its complement), i.e.,  $\gamma_{\mathcal{C}}(x) = \inf_{\lambda \geq 0} \lambda I_{\mathcal{C}}\left(\frac{x}{\lambda}\right)$ . From this interpretation,  $\gamma_{\mathcal{C}}$  is therefore a convex function. Moreover, it is positively homogeneous and has non-negative values. Conversely, any function  $\gamma$  which satisfies these three properties is the gauge function of the set  $\{x \in \mathbb{R}^d, \gamma(x) \leq 1\}$ .

Several closed convex sets  $\mathcal{C}$  lead to the same gauge function. However the unique closed convex set containing the origin is  $\{x \in \mathbb{R}^d, \gamma_{\mathcal{C}}(x) \leq 1\}$ . In general, we have for any closed convex set  $\mathcal{C}$ ,  $\{x \in \mathbb{R}^d, \gamma_{\mathcal{C}}(x) \leq 1\} = \text{hull}(\mathcal{C} \cup \{0\})$ .

Classical examples are norms, which are gauge functions coming from their unit balls: norms are gauge functions  $\gamma$  which (a) have a full domain, (b) are such that  $\gamma(x) = 0 \Leftrightarrow x = 0$ , and (c) are even, which corresponds to sets  $\mathcal{C}$  which (a) have 0 in its interior, (b) are compact and (c) centrally symmetric. In general, the set  $\mathcal{C}$  might neither be compact nor centrally symmetric, for example, when  $\mathcal{C} = \{x \in \mathbb{R}_d^+, 1_d^\top x \leq 1\}$ . Moreover, a gauge function may take infinite values (such as in the previous case, for any vector with a strictly negative component).

**Polar sets and functions.** Given any set  $\mathcal{C}$  (not necessarily convex), the polar of  $\mathcal{C}$  is the set  $\mathcal{C}^\circ$  defined as

$$\mathcal{C}^\circ = \{y \in \mathbb{R}^d, \forall x \in \mathcal{C}, x^\top y \leq 1\}.$$

It is always closed and convex. Moreover, the polar of  $\mathcal{C}$  is equal to the polar of the closure of  $\text{hull}(\mathcal{C} \cup \{0\})$ .

When  $\mathcal{C}$  is the unit ball of a norm  $\Omega$ ,  $\mathcal{C}^\circ$  is the unit ball of the dual norm which we denote  $\Omega^\circ$  (instead of the usual definition  $\Omega^*$ , because the Fenchel conjugate of  $\Omega$  is not the dual norm, but the indicator function of the dual unit ball).

If  $\mathcal{C}$  is a closed convex set containing the origin, then  $\mathcal{C}^{\circ\circ} = \mathcal{C}$ —more generally, for any set  $\mathcal{C}$ ,  $\mathcal{C}^{\circ\circ}$  is the closure of  $\text{hull}(\mathcal{C} \cup \{0\})$ . The polarity is a one-to-one mapping from closed convex sets containing the origin to themselves. In this paper, we will also consider gauge functions associated with closed potentially non convex sets  $\mathcal{C}$ , in which case, we mean the gauge function associated to  $\mathcal{C}^{\circ\circ} = \text{hull}(\mathcal{C} \cup \{0\})$ , i.e.,  $\gamma_{\mathcal{C}} = \gamma_{\mathcal{C}^{\circ\circ}}$ .

The Fenchel conjugate of  $\gamma_{\mathcal{C}}$  is the indicator function of  $\mathcal{C}^\circ$ , i.e.,  $\gamma_{\mathcal{C}}^* = I_{\mathcal{C}^\circ}$ , which is equivalent to  $\gamma_{\mathcal{C}} = I_{\mathcal{C}^\circ}^*$ , i.e.,  $\forall x \in \mathbb{R}^d$ ,  $\gamma_{\mathcal{C}}(x) = \sup_{y \in \mathcal{C}^\circ} x^\top y$ . Given a gauge function  $\gamma_{\mathcal{C}}$ , we define its polar as the function  $\gamma_{\mathcal{C}}^\circ$  given by

$$\gamma_{\mathcal{C}}^\circ(y) = \inf \left\{ \lambda \geq 0, \forall x \in \mathbb{R}^d, x^\top y \leq \lambda \gamma_{\mathcal{C}}(x) \right\} = \sup_{x \in \mathbb{R}^d} \frac{x^\top y}{\gamma_{\mathcal{C}}(x)},$$

the last inequality being true only if  $\gamma_{\mathcal{C}}(x) = 0 \Leftrightarrow x = 0$  (i.e.,  $\mathcal{C}$  compact). It turns out that  $\gamma_{\mathcal{C}}^\circ = \gamma_{\mathcal{C}^\circ}$ . This implies that  $\gamma_{\mathcal{C}^\circ} = I_{\mathcal{C}^{\circ\circ}}^*$ , i.e.,  $\gamma_{\mathcal{C}^\circ}(y) = \sup_{x \in \mathcal{C}^{\circ\circ}} x^\top y = \sup_{x \in \mathcal{C}} (x^\top y)_+$ . For example, the polar of a norm is its dual norm. We have for all  $x, y \in \mathbb{R}^d$ , the inequality that is well known for norms:  $x^\top y \leq \gamma_{\mathcal{C}}(x) \gamma_{\mathcal{C}^\circ}(y)$ . Finally, the Fenchel-conjugate of  $x \mapsto \frac{1}{2} \gamma_{\mathcal{C}}(x)^2$  is  $y \mapsto \frac{1}{2} \gamma_{\mathcal{C}}^\circ(y)^2$ .

**Operations on gauge functions.** For two closed convex sets  $\mathcal{C}$  and  $\mathcal{D}$  containing the origin, then for all  $x \in \mathbb{R}^d$ ,  $\max\{\gamma_{\mathcal{C}}(x), \gamma_{\mathcal{D}}(x)\} = \gamma_{\mathcal{C} \cap \mathcal{D}}(x)$ . Another combination, is the ‘‘inf-convolution’’ of  $\gamma_{\mathcal{C}}$  and  $\gamma_{\mathcal{D}}$ , i.e.,  $x \mapsto \inf_{x=y+z} \gamma_{\mathcal{C}}(z) + \gamma_{\mathcal{D}}(y)$ , which is equal to  $\gamma_{\text{hull}(\mathcal{C} \cup \mathcal{D})}$ . Moreover,  $\gamma_{\mathcal{C} \cap \mathcal{D}}^\circ = \gamma_{\text{hull}(\mathcal{C}^\circ \cup \mathcal{D}^\circ)}$ , or equivalently,  $(\mathcal{C} \cap \mathcal{D})^\circ = \text{hull}(\mathcal{C}^\circ \cup \mathcal{D}^\circ)$ .

**Links with convex hulls.** Given a compact set  $\mathcal{P}$  and its compact convex hull  $\mathcal{C}$  (for example,  $\mathcal{P}$  might be the set of extreme points of  $\mathcal{C}$ ), we have  $\mathcal{P}^\circ = \mathcal{C}^\circ$ , since maxima of linear functions on  $\mathcal{C}$  or  $\mathcal{P}$  are equal. An alternative definition of  $\gamma_{\mathcal{C}}$  is then

$$\gamma_{\mathcal{C}}(x) = \min \left\{ \sum_{i \in I} \eta_i, (\eta_i)_{i \in I} \in \mathbb{R}_+^I, (x_i)_{i \in I} \in \mathcal{P}^I, I \text{ finite}, x = \sum_{i \in I} \eta_i x_i \right\}.$$

Moreover, in the definition above, by Caratheodory’s theorem for cones, we may restrict the cardinality of  $I$  to be less than or equal to  $d$ .

### 3 Representations of gauge functions through quadratic functions

In this section, we consider a closed set  $\mathcal{U}$  that contains the origin in its convex hull, and explore various representations of the gauge function  $\gamma_{\mathcal{U}} = \gamma_{\mathcal{U}^{\circ\circ}}$  as a maxima or minima of quadratic functions. This corresponds to respectively inner and outer approximations of the convex hull  $\mathcal{U}^{\circ\circ}$  by ellipsoids. Note that unless otherwise stated,  $\mathcal{U}$  might not be convex, might not be compact (i.e.,  $\gamma_{\mathcal{U}}(x) = 0$  even if  $x \neq 0$ ),  $\mathcal{U}$  might not be centrally symmetric, and  $\text{dom}(\gamma_{\mathcal{U}}) = \mathbb{R}_+ \mathcal{U}$  may be strictly included in  $\mathbb{R}^d$ . Since quadratic variational formulations have to be centrally symmetric, we consider symmetrized versions of gauge functions; we consider two ways of ‘‘symmetrizing’’ a gauge function: by intersecting  $\mathcal{U}^{\circ\circ}$  and  $-\mathcal{U}^{\circ\circ}$ , which corresponds to the gauge function  $\gamma_{\mathcal{U}^{\circ\circ} \cap (-\mathcal{U}^{\circ\circ})}(x) = \max\{\gamma_{\mathcal{U}}(x), \gamma_{\mathcal{U}}(-x)\}$ , or by intersecting  $\mathcal{U}^\circ$  and  $-\mathcal{U}^\circ$ , which corresponds to  $\gamma_{\mathcal{U} \cup (-\mathcal{U})}(x) = \gamma_{\mathcal{U}^\circ \cap (-\mathcal{U}^\circ)}(x) = \inf_{x=x_++x_-} \{\gamma_{\mathcal{U}}(x_+) + \gamma_{\mathcal{U}}(-x_-)\}$ . These two gauge functions have domain  $\text{span}(\mathcal{U})$ , the vector space generated by  $\mathcal{U}$ . When  $\mathcal{U}$  is centrally symmetric, then these two gauge functions are equal to  $\gamma_{\mathcal{U}}$ .

### 3.1 Maxima of quadratic functions

We first consider *closed convex* sets  $\mathcal{K} \subset \mathcal{S}_d^+$  of symmetric positive definite matrices such that

$$\forall x \in \text{span}(\mathcal{U}), \gamma_{\mathcal{U} \cup (-\mathcal{U})}(x)^2 = \max_{M \in \mathcal{K}} x^\top M x. \quad (1)$$

When  $\mathcal{U}$  is symmetric, the variational formulation in Eq. (1) leads to a representation of  $\gamma_{\mathcal{U}}(x)^2$  as a convex function  $I_{\mathcal{K}}^*(xx^\top)$  of  $xx^\top$ . Note that in general,  $\mathcal{K}$  is not unique. We now show that there always exists a set  $\mathcal{K}$  satisfying Eq. (1), and provide a description of the largest such set.

**Proposition 1** *Let  $\mathcal{P}_{\mathcal{U}} = \{xx^\top, x \in \mathcal{U}\}$ . Then  $\mathcal{P}_{\mathcal{U}}^\circ \cap \mathcal{S}_d^+$  is the largest closed convex set  $\mathcal{K} \subset \mathcal{S}_d^+$  of positive semidefinite matrices such that Eq. (1) is satisfied.*

**Proof** We first show that  $\mathcal{P}_{\mathcal{U}}^\circ \cap \mathcal{S}_d^+$  satisfies Eq. (1). Let  $x \in \mathbb{R}^d$ , we have  $\sup_{M \in \mathcal{P}_{\mathcal{U}}^\circ} x^\top M x = I_{\mathcal{P}_{\mathcal{U}}^\circ}^*(xx^\top) = \gamma_{\mathcal{P}_{\mathcal{U}}}(xx^\top)$ . Since  $\mathcal{P}_{\mathcal{U}} \subset \mathcal{S}_d^+$ ,  $xx^\top$  may only be decomposed as the sum of matrices of the form  $\lambda xx^\top$  and  $\lambda(-x)(-x)^\top$ , for  $\lambda \geq 0$ . This implies that  $\gamma_{\mathcal{P}_{\mathcal{U}}}(xx^\top) = \min\{\gamma_{\mathcal{U}}(x), \gamma_{\mathcal{U}}(-x)\}^2$ . Note that since  $\mathcal{P}_{\mathcal{U}}^\circ$  does not only include positive semi-definite matrices, it is not incoherent that  $x \mapsto I_{\mathcal{P}_{\mathcal{U}}^\circ}^*(xx^\top) = \sup_{M \in \mathcal{P}_{\mathcal{U}}^\circ} x^\top M x = \min\{\gamma_{\mathcal{U}}(x), \gamma_{\mathcal{U}}(-x)\}^2$  is not convex.

We now have  $\sup_{M \in \mathcal{P}_{\mathcal{U}}^\circ \cap \mathcal{S}_d^+} \sqrt{x^\top M x} \leq \sup_{M \in \mathcal{P}_{\mathcal{U}}^\circ} \sqrt{x^\top M x} = \min\{\gamma_{\mathcal{U}}(x), \gamma_{\mathcal{U}}(-x)\}$ . Since this is a convex function of  $x$ , it must be less than its convex envelope, which is  $\gamma_{\mathcal{U} \cup (-\mathcal{U})}(x)$  (since they have the same Fenchel conjugates).

Moreover, if  $v \in \mathcal{U}^\circ \cap (-\mathcal{U})^\circ$ , then  $vv^\top \in \mathcal{P}_{\mathcal{U}}^\circ$ . This implies that  $\sup_{M \in \mathcal{P}_{\mathcal{U}}^\circ} x^\top M x \geq \sup_{v \in \mathcal{U}^\circ \cap (-\mathcal{U})^\circ} (v^\top x)^2 = I_{\mathcal{U}^\circ \cap (-\mathcal{U})^\circ}^*(x)^2 = \gamma_{\mathcal{U}^\circ \cap (-\mathcal{U})^\circ}(x)^2 = \gamma_{\mathcal{U} \cup (-\mathcal{U})}(x)^2$ . Thus Eq. (1) is indeed satisfied by  $\mathcal{P}_{\mathcal{U}}^\circ \cap \mathcal{S}_d^+$ . Finally, if  $\mathcal{K}$  satisfies Eq. (1), then we must have  $\mathcal{K} \subset \mathcal{P}_{\mathcal{U}}^\circ$  by definition of polar sets, hence  $\mathcal{P}_{\mathcal{U}}^\circ \cap \mathcal{S}_d^+$  is the largest.  $\blacksquare$

Note that the set  $\mathcal{P}_{\mathcal{U}}^\circ \cap \mathcal{S}_d^+$  is equal to  $\{M \in \mathcal{S}_d^+, \forall u \in \mathcal{U}, u^\top M u \leq 1\}$ —this representation was already considered in [6] for norms. In certain situations, the largest possible set is desirable (for example when deriving convex relaxations). In other situations (for example when using these representations for optimization), smallest sets are desirable. However, such a notion is not possible. Indeed, for  $\gamma_{\mathcal{U}} = \|\cdot\|_2$ , the sets  $\mathcal{K} = \{I\}$  and  $\mathcal{K} = \{M \in \mathcal{S}_d^+, \|M\|_F \leq 1\}$ , are two possible sets, and thus there is no single smallest set. One possible small set is the convex hull of the maximal elements of  $\mathcal{P}_{\mathcal{U}}^\circ \cap \mathcal{S}_d^+$  (for the positive semi-definite order).

In the proof of Prop. 1, we have introduced the gauge function  $\gamma_{\mathcal{P}_{\mathcal{U}}}$ . We now provide a representation of  $\gamma_{\mathcal{P}_{\mathcal{U}}}$  related to factorizations of positive semidefinite matrices (note that in the following proposition, we only assume that  $\mathcal{U}$  is closed and contains 0 in its hull).

**Proposition 2** *Let  $\mathcal{P}_{\mathcal{U}} = \{xx^\top, x \in \mathcal{U}\} \subset \mathcal{S}_d^+$ . We have, for all positive semidefinite matrix  $M \in \mathcal{S}_d^+$ :*

$$\gamma_{\mathcal{P}_{\mathcal{U}}}(M) = \inf_{r \geq 0} \inf_{M = \sum_{m=1}^r x_m x_m^\top} \sum_{m=1}^r \gamma_{\mathcal{U}}(x_m)^2 = \inf_{r \geq 0} \inf_{M = \sum_{m=1}^r \lambda_m x_m x_m^\top, x_m \in \mathcal{U}} \sum_{m=1}^r \lambda_m.$$

Moreover, we may choose  $r \leq d(d+1)/2$ .

**Proof** This is a direct application of the representation of gauge functions and the property  $\gamma_{\mathcal{P}_U}(xx^\top) = \min\{\gamma_U(x), \gamma_U(-x)\}^2$ , that was shown in the proof of Prop. 1. ■

Note that  $\mathbb{R}_+\mathcal{U} = \text{dom}(\gamma_U)$  may not be equal to  $\mathbb{R}^d$ , and that the domain of  $\gamma_{\mathcal{P}_U}$  may not be equal to  $\mathcal{S}_d^+$ , i.e., some positive matrices may not be decomposed as positive linear combinations of dyads  $uu^\top$  obtained from elements  $u$  of  $\mathcal{U}$ ; for example, when  $\mathcal{U} = \{x \in \mathbb{R}_+^d, 1_d^\top x = 1\}$  is the simplex, then  $\text{dom}(\gamma_U)$  is the set of completely positive matrices (see, e.g., [11]).

The last proposition provides a structured decomposition framework for positive semi-definite matrices, that will be considered for rectangular matrices in Section 4. Obtaining explicitly the decomposition  $M = UU^\top$  from  $M$  may be done with the iterative algorithms presented in Section 6. Note that by considering a representation of  $U$  as  $U = M^{1/2}S$  where  $SS^\top = I$ , computing  $\gamma_{\mathcal{P}_U}$  may be seen as a factorization problem with two factors (which can then be used in alternating minimization procedures).

### 3.2 Minima of quadratic functions

We now consider closed convex sets  $\mathcal{L} \subset \mathcal{S}_d^+$  such that

$$\forall x \in \text{span}(\mathcal{U}), \max\{\gamma_U(x), \gamma_U(-x)\}^2 = \inf_{M \in \mathcal{L}} x^\top M^{-1}x. \quad (2)$$

Here, we define  $x^\top M^{-1}x$  as  $x^\top M^{-1}x = \inf t$  such that  $\begin{pmatrix} M & x \\ x^\top & t \end{pmatrix} \succ 0 \Leftrightarrow tM \succ xx^\top$ . This implies that the value may be finite even when  $M$  is not invertible.

When  $\mathcal{U}$  is symmetric, the variational formulation in Eq. (2) leads to a representation of  $\gamma_U(x)^2$  as a concave  $\inf_{M \in \mathcal{L}} \text{tr} M^{-1}xx^\top$  of  $xx^\top$ . This is to be contrasted with the fact that it is also a *convex* function of  $xx^\top$  because of the representation discussed in Section 3.1. The two properties are in fact related through a duality argument:

**Proposition 3** *Let  $\mathcal{L} \subset \mathcal{S}_d^+$  be a closed convex set. Then the following two properties are equivalent:*

- (a)  $\forall x \in \text{span}(\mathcal{U}), \inf_{M \in \mathcal{L}} x^\top M^{-1}x = \max\{\gamma_U(x), \gamma_U(-x)\}^2,$
- (b)  $\forall y \in \text{span}(\mathcal{U}^\circ), \sup_{M \in \mathcal{L}} y^\top M y = \gamma_{\mathcal{U} \cup (-\mathcal{U})}(y)^2.$

*This implies that the largest set  $\mathcal{L}$  such that (a) is valid is  $\mathcal{P}_{\mathcal{U}^\circ} \cap \mathcal{S}_d^+$  defined in Prop. 1.*

**Proof** The two functions  $y \mapsto \frac{1}{2}\gamma_{\mathcal{U} \cup (-\mathcal{U})}(y)^2 = \frac{1}{2}\inf_{y=y_++y_-} \{\gamma_{\mathcal{U}}^\circ(y_+) + \gamma_{\mathcal{U}}^\circ(-y_-)\}^2$  and  $x \mapsto \frac{1}{2}\max\{\gamma_U(x), \gamma_U(-x)\}^2$  are Fenchel-conjugate to each other. We thus need to show that this is the same for  $x \mapsto \frac{1}{2}\inf_{M \in \mathcal{L}} x^\top M^{-1}x$  and  $y \mapsto \frac{1}{2}\sup_{M \in \mathcal{L}} y^\top M y$ , which is straightforward to check. Note that  $\mathcal{P}_{\mathcal{U}^\circ} \subset \mathcal{P}_{\mathcal{U}^\circ}^\circ \subset \mathcal{P}_{\mathcal{U}}^\circ \cap \mathcal{S}_d^+$ , but that the reverse inclusions are typically not true (take, e.g., the unit  $\ell_2$ -ball). ■

**Duality between representations.** For a given gauge function  $\gamma_{\mathcal{U}}$ , there are two possible variational representations, as maxima or minima of quadratic functions. The two possible variational representations are linked to each other. Indeed, if  $\mathcal{K} \subset \mathcal{S}_d^+$  is a closed convex set such that for all  $x \in \mathbb{R}^d$ ,  $\sup_{M \in \mathcal{K}} x^\top M x \leq \gamma_{\mathcal{U} \cup (-\mathcal{U})}(x)^2$  (i.e., we have a lower-bound which is convex in  $xx^\top$ ), then  $\mathcal{K} \subset \mathcal{P}_{\mathcal{U}}^\circ$ , which implies by duality  $\mathcal{P}_{\mathcal{U}}^{\circ\circ} \subset \mathcal{K}^\circ \cap \mathcal{S}_d^+$ . Thus for any  $y \in \mathbb{R}^d$ ,  $I_{\mathcal{K}^\circ \cap \mathcal{S}_d^+}^*(yy^\top) \geq I_{\mathcal{P}_{\mathcal{U}}^\circ}^*(yy^\top) = \max\{\gamma_{\mathcal{U}}^\circ(y), \gamma_{\mathcal{U}}^\circ(-y)\}^2$ . This then implies by Fenchel duality (i.e., using arguments from the proof of Prop. 3) that for all  $x \in \mathbb{R}^d$ ,  $\gamma_{\mathcal{U} \cup (-\mathcal{U})}(x)^2 \geq \inf_{M \in \mathcal{K}^\circ \cap \mathcal{S}_d^+} x^\top M^{-1} x$ , i.e., we have a concave lower-bound based on  $\mathcal{K}^\circ \cap \mathcal{S}_d^+$ .

Similarly, if we have a concave upper-bound of  $\max\{\gamma_{\mathcal{U}}^\circ(y), \gamma_{\mathcal{U}}^\circ(-y)\}^2$  based on  $\mathcal{K}$ , we also have a convex upper-bound based on  $\mathcal{K}^\circ \cap \mathcal{S}_d^+$ . Note that it is not true in general that exact representations of one kind transfer to exact representation of the other kind.

### 3.3 Examples

For the  $\ell_2$ -norm ball, then we may characterize exactly  $\mathcal{P}_{\mathcal{U}}^\circ$  as  $\mathcal{P}_{\mathcal{U}}^\circ = \{M \in \mathbb{R}^{d \times d}, M \preceq I\}$ . This provides an example  $\mathcal{K} = \mathcal{P}_{\mathcal{U}}^\circ \cap \mathcal{S}_d^+ = \{M \in \mathcal{S}_d^+, M \preceq I\}$ , for which  $\mathcal{K}^\circ \cap \mathcal{S}_d^+ = \{M \in \mathcal{S}_d^+, \text{tr } M \leq 1\}$ . Another example is  $\mathcal{K} = \{I\}$  with  $\mathcal{K}^\circ \cap \mathcal{S}_d^+ = \{M \in \mathcal{S}_d^+, \text{tr } M \leq 1\}$ . Yet another one is  $\mathcal{K} = \{M \in \mathcal{S}_d^+, \|M\|_F \leq 1\}$  with  $\mathcal{K}^\circ \cap \mathcal{S}_d^+ = \{M \in \mathcal{S}_d^+, \|M\|_F \leq 1\}$ .

For the  $\ell_1$ -norm ball, we also have a representation in closed form of the largest set  $\mathcal{K} = \mathcal{P}_{\mathcal{U}}^\circ \cap \mathcal{S}_d^+ = \{M \in \mathbb{R}^{d \times d}, M \succeq 0, \|\text{diag}(M)\|_\infty \leq 1\}$ .

Apart from these two sets, a simple (manageable in polynomial time) description of  $\mathcal{P}_{\mathcal{U}}^\circ \cap \mathcal{S}_d^+$  is not available and smaller sets are generally available. For example, for the  $\ell_p$ -norm,  $p \in [1, \infty]$ , the set  $\mathcal{K} = \{M \succeq 0, \|M\|_q \leq 1\}$  satisfies Eq. (1). Moreover, if  $p \geq 2$ ,  $\mathcal{K} = \{\text{Diag}(\eta), \eta \in \mathbb{R}_+^d, \|\eta\|_{p/(p-2)} \leq 1\}$  also does [10] (because  $p/(p-2) = 1/(1-2/p)$  and  $\sup_{\eta \geq 0, \|\eta\|_{p/(p-2)} \leq 1} \sum_i \eta_i x_i^2 = \|x \circ x\|_{p/2}$ , where  $\circ$  denotes the pointwise product of vectors). See additional example of diagonal representations in Section 3.4.

### 3.4 Diagonal representations

We now consider cases where the set of matrices  $\mathcal{K}$  and  $\mathcal{L}$  are diagonal, i.e., all principal axes of ellipsoids are aligned with the canonical basis. For simplicity, we consider only sets  $\mathcal{U}$  which are compact, have zero in their interior, and are invariant by sign flips of any components. The corresponding gauge functions are then *absolute norms*, which are functions of the absolute values of each component [49]. The following proposition provides several characterizations (see related work in [10, 6]).

**Proposition 4** *Let  $\Omega$  be an absolute norm on  $\mathbb{R}^d$  and  $\Omega^\circ$  its dual (also absolute) norm. The following three conditions are equivalent:*

- (a) *There exists a non-empty closed convex subset  $\mathcal{H}$  of  $\mathbb{R}_+^d$  such that  $\forall x \in \mathbb{R}^d, \Omega(x)^2 = \inf_{\eta \in \mathcal{H}} x^\top \text{Diag}(\eta)^{-1} x$ .*
- (b) *The function  $t \mapsto \Omega(t^{1/2})^2$  is concave on  $\mathbb{R}_+^d$ .*
- (c) *The function  $t \mapsto \Omega^\circ(t^{1/2})^2$  is convex on  $\mathbb{R}_+^d$ .*



**Proof** It is straightforward to see that (a) implies both (b) and (c).

Assume (b). The Fenchel conjugate  $t \mapsto -\Omega(t^{1/2})^2$  on  $\mathbb{R}_+^d$ , defined as  $f(u) = \sup_{t \in \mathbb{R}_+^d} t^\top u - \Omega(t^{1/2})^2$  is the indicator function of a convex set  $\mathcal{C}$  (because  $t \mapsto \Omega(t^{1/2})^2$  is positively homogeneous). Moreover  $-\mathcal{C} \subset \mathbb{R}_+^d$ , and  $u \in \mathcal{C}$  if and only if for all  $t \in \mathbb{R}_+^d$ ,  $\Omega(t^{1/2})^2 \leq -u^\top t$ , i.e., for all  $w \in \mathbb{R}^d$ ,  $\Omega(w)^2 \leq w^\top \text{Diag}(-u)w$ , which is equivalent to, for all  $s \in \mathbb{R}^d$ ,  $\Omega^\circ(s)^2 \geq s^\top \text{Diag}(-u)^{-1}s$ . Thus  $-1/\mathcal{C} = \{-(1/u), u \in \mathcal{C}\}$  is also convex.

By Fenchel duality, we get, for all  $t \in \mathbb{R}_+^d$ ,  $-\Omega(t^{1/2})^2 = \sup_{u \in \mathcal{C}} t^\top u$ , i.e., for all  $w \in \mathbb{R}^d$ ,  $\Omega(w)^2 = \Omega(|w|)^2 = \inf_{u \in -\mathcal{C}} w^\top \text{Diag}(u)w$ , which shows that we can take  $\mathcal{H} = -1/\mathcal{C}$  and obtain (a).

Assume (c). Then the Fenchel conjugate  $t \mapsto \Omega^\circ(t^{1/2})^2$  on  $\mathbb{R}_+^d$ , defined as  $f(u) = \sup_{t \in \mathbb{R}_+^d} t^\top u - \Omega^\circ(t^{1/2})^2$  is the indicator function of a convex set  $\mathcal{C}$  (because  $t \mapsto \Omega^\circ(t^{1/2})^2$  is positively homogeneous). By duality, we get, for all  $t \in \mathbb{R}_+^d$ ,  $\Omega^\circ(t^{1/2})^2 = \sup_{u \in \mathcal{C}} t^\top u$ , i.e., for all  $s \in \mathbb{R}^d$ ,  $\Omega^\circ(s)^2 = \Omega^\circ(|s|)^2 = \sup_{u \in \mathcal{C}} w^\top \text{Diag}(u)w$ , which leads to (a) when computing Fenchel conjugates.  $\blacksquare$

The construction above gives a way of constructing the largest set  $\mathcal{H} \subset \mathbb{R}_+^d$ , such that  $\forall w \in \mathbb{R}^d, \Omega(w)^2 = \inf_{\eta \in \mathcal{H}} w^\top \text{Diag}(\eta)^{-1}w$ :  $\mathcal{H}$  is the set of  $\eta \in \mathbb{R}_+^d$  such that for all  $s \in \mathbb{R}^d$ ,  $s^\top \text{Diag}(\eta)s \leq \Omega^\circ(s)^2$ , then, we have shown above that  $1/\mathcal{H}$  is the set of  $\zeta \in \mathbb{R}_+^d$  such that for all  $w \in \mathbb{R}^d$ ,  $\Omega(w)^2 \leq w^\top \text{Diag}(\zeta)w$ , and is also convex.

**Examples.** Among  $\ell_p$ -norms, only the ones for which  $p \leq 2$  have diagonal representations, with  $\mathcal{H} = \{\eta \geq 0, \|\eta\|_{p/(2-p)} \leq 1\}$ . Indeed, we have for  $\Omega = \|\cdot\|_p$  and  $1/p + 1/q = 1$ :  $\Omega^\circ(t^{1/2})^2 = \|t\|_{q/2} = \sup_{\|s\|_{1/(1-2/q)} \leq 1} s^\top t$ , with  $1/(1-2/q) = 1/(1-2(1-1/p)) = 1/(2/p-1) = p/(2-p)$ .

Beyond  $\ell_p$ -norms, the set  $\mathcal{H}$  have also been used to define norms with specific structured sparsity-inducing properties [40, 6] or optimality properties in terms of convex relaxations [43]. Note moreover, that such representations are useful for optimization as they lead to simple reweighted- $\ell_2$  algorithms for all of these norms [19, 6].

### 3.5 Matrix norms invariant by rotation

We now consider norms on  $\mathbb{R}^{n \times d}$  which are invariant by right-multiplication by a rotation, i.e., which are functions of  $WW^\top$ ,  $W \in \mathbb{R}^{n \times d}$ . The corresponding unit ball  $\mathcal{U}$  is invariant by right-multiplication. The following proposition may be shown using the same arguments than Prop. 4.

**Proposition 5** *Let  $\Omega$  be a norm on  $\mathbb{R}^{n \times d}$  which is invariant by right-multiplication by a rotation, and  $\Omega^\circ$  its dual norm. The following three conditions are equivalent:*

- (a) *There exists a non-empty convex subset  $\mathcal{K}$  of  $\mathcal{S}_d^+$  such that  $\forall W \in \mathbb{R}^{n \times d}$ ,  $\Omega(W)^2 = \inf_{M \in \mathcal{K}} \text{tr } W^\top M^{-1}W$ .*
- (b)  *$W \mapsto \Omega(W)^2$  is a concave function of  $WW^\top$ .*
- (c)  *$V \mapsto \Omega^\circ(V)^2$  is a concave function of  $VV^\top$ .*

All norms obtained in Section 4 for  $\mathcal{V}$  the  $\ell_2$ -ball will be instances of these norms. Moreover, all norms which are invariant by right and left multiplication by a rotation, must be obtained as a spectral function, i.e.,  $\Omega(W)$  is a symmetric norm on the singular values of  $W$ . Quadratic representations based on the spectrum could be derived as well using tools from [36]. Note that for Prop. 5, the set  $\mathcal{K}$  is then a spectral set (defined through constraints on eigenvalues).

## 4 Gauge functions and structured matrix factorizations

In this section, we consider two compact (non-necessarily convex) sets  $\mathcal{U} \subset \mathbb{R}^n$  and  $\mathcal{V} \subset \mathbb{R}^d$ . We will always assume that 0 is in the convex hulls of both  $\mathcal{U}$  and  $\mathcal{V}$ .

We consider the decomposition of matrices  $X \in \mathbb{R}^{n \times d}$  as a positive linear combination of matrices of the form  $uv^\top$  for  $u \in \mathcal{U}$  and  $v \in \mathcal{V}$ , i.e.,  $X = \sum_{m=1}^r \lambda_m u_m v_m^\top$ , with a priori no constraints on the rank  $r$  (it will always be less than  $nd$ ). Finding the decomposition such that  $\sum_{m=1}^r \lambda_m$  is minimum leads to a gauge function, with interesting properties.

### 4.1 Definitions

Let  $\mathcal{U} \subset \mathbb{R}^n$  and  $\mathcal{V} \subset \mathbb{R}^d$  be two compact sets. We define the function  $\Theta$  on  $\mathbb{R}^{n \times d}$  as follows

$$\Theta(X) = \inf_{r \geq 0} \inf_{X = \sum_{m=1}^r \lambda_m u_m v_m^\top, \lambda_m \in \mathbb{R}_+, u_m \in \mathcal{U}, v_m \in \mathcal{V}} \sum_{m=1}^r \lambda_m. \quad (3)$$

This function is exactly the gauge function of the set  $\{uv^\top, u \in \mathcal{U}, v \in \mathcal{V}\}$ . It is therefore convex, non-negative and positively homogeneous. Moreover, the minimization problem defining it is attained for  $r \leq nd$ . There are several equivalent formulations which we are going to use, which relies on the fact that  $u_m v_m^\top = (u_m \lambda_m)(v_m \lambda_m^{-1})^\top$  for any  $\lambda_m > 0$  (in the following expressions, we always minimize with respect to the rank and thus omit the notation  $\inf_{r \geq 0}$ ).

**Proposition 6** *We have*

$$\begin{aligned} \Theta(X) &= \inf_{X = \sum_{m=1}^r u_m v_m^\top} \frac{1}{2} \sum_{m=1}^r \{\gamma_{\mathcal{U}}(u_m)^2 + \gamma_{\mathcal{V}}(v_m)^2\} = \inf_{X = \sum_{m=1}^r u_m v_m^\top} \sum_{m=1}^r \gamma_{\mathcal{U}}(u_m) \gamma_{\mathcal{V}}(v_m) \\ &= \inf_{X = \sum_{m=1}^r u_m v_m^\top, u_m \in \mathcal{U}} \sum_{m=1}^r \gamma_{\mathcal{V}}(v_m) = \inf_{X = \sum_{m=1}^r u_m v_m^\top, v_m \in \mathcal{V}} \sum_{m=1}^r \gamma_{\mathcal{U}}(u_m). \end{aligned}$$

Moreover,  $\Theta$  is a norm as soon as  $\mathcal{U}$  and  $\mathcal{V}$  are norm balls—in this case, these norms were studied in several settings [30, 7, 17]. The next proposition shows that the polar of  $\Theta$  has a simple form (the proof is straightforward from the gauge function interpretation). It is a matrix norm when  $\mathcal{U}$  and  $\mathcal{V}$  are norm balls; moreover, in all cases, since  $\mathcal{U}$  and  $\mathcal{V}$  are assumed compact,  $\Theta^\circ$  has full domain.

**Proposition 7** *The polar of  $\Theta$  is equal to*

$$\Theta^\circ(Y) = \max_{u \in \mathcal{U}, v \in \mathcal{V}} u^\top Y v = \max_{u \in \mathcal{U}} \gamma_{\mathcal{V}}^\circ(Y^\top u) = \max_{v \in \mathcal{V}} \gamma_{\mathcal{U}}^\circ(Y v) = \max_{\gamma_{\mathcal{U}}(u) \leq 1, \gamma_{\mathcal{V}}(v) \leq 1} u^\top Y v.$$

The gauge function  $\Theta$  or its polar may not be computed in closed form in general. There are two special cases where this is possible, namely when one of the sets is the  $\ell_1$ -norm ball, and when the two sets are the  $\ell_2$ -norm balls.

## 4.2 Examples

In this paper, we consider several sets  $\mathcal{U}$  and  $\mathcal{V}$  with interesting properties which are displayed in Table 1. Note that when  $\mathcal{V}$  is the  $\ell_1$ -ball, then the norm  $\Theta$  may be computed in closed form as  $\Theta(X) = \sum_{i=1}^d \gamma_{\mathcal{U}}(X(:, i))$ , where  $X(:, i)$  is the  $i$ -th column of  $X$ . The case where  $\mathcal{V}$  is the  $\ell_2$ -ball is also specific and will be treated in Section 4.3. Some specific combinations of  $\mathcal{U}$  and  $\mathcal{V}$  are particularly interesting:

- $\mathcal{U}$  and  $\mathcal{V}$  are unit  $\ell_2$ -balls: the norm  $\Theta$  is the nuclear norm (sum of singular values), leading to a regularizer inducing low-rank [45].
- $\mathcal{V}$  is the unit  $\ell_2$ -ball and  $\mathcal{U}^\circ$  is the intersection of the unit  $\ell_2$ -ball and a scaled  $\ell_\infty$ -ball, i.e.,  $\mathcal{U}^\circ = \{w \in \mathbb{R}^n, \|w\|_2 \leq 1, \|w\|_\infty \leq \nu\}$ . The norm  $\Theta$  may then be computed in closed form as  $\Theta(X) = \inf_{X=Y+Z} \|Y\|_* + \nu \sum_{i=1}^n \|Z(i, \cdot)\|_2$ , and has been used in robust versions of principal component analysis where some observations may be corrupted (through a “low-rank + group-sparse” model) [50].
- $\gamma_{\mathcal{V}} = \|\cdot\|_2$  and  $\gamma_{\mathcal{U}} = \sqrt{\nu\|\cdot\|_2^2 + (1-\nu)\|\cdot\|_1^2}$ : this is a convex relaxation of sparse coding [7], where a decomposition with a sparse factor  $U$  and low rank (small number of columns for  $U$  and  $V$ ) is looked for.
- $\mathcal{U}$  and  $\mathcal{V}$  are  $\ell_\infty$ -unit balls: the norm  $\Theta$  has been considered as a complexity measure for sign matrices [37]. It is moreover related to the cut-norm [2]; the relaxation presented in Section 5 is the max-norm [48, 35]. Note here that our conditional gradient algorithms presented in Section 6 significantly improves the number of rank-one factor to obtain a  $\varepsilon$ -approximate decomposition from  $O(1/\varepsilon^2)$  [2] to  $O(\log \frac{1}{\varepsilon})$ .
- $\gamma_{\mathcal{V}} = \|\cdot\|_2$  and  $\gamma_{\mathcal{U}} = \|\cdot\|_{4/3}$ : we have  $\gamma_{\mathcal{U}}^\circ = \|\cdot\|_4$ , and the dual norm is such that  $\Theta^\circ(Y)^4 = \max_{\|v\|_2 \leq 1} \sum_i (Yv)_i^4$ , and corresponds to maximum kurtosis projections [28].
- $\mathcal{U}$  and  $\mathcal{V}$  are  $\ell_p$ -norm balls. This is a case considered by [10], where approximation guarantees are derived. Values of  $p$  for which dimension-independent guarantees are obtained are exactly the ones corresponding to diagonal representations in Prop. 11.
- $\mathcal{U}$  and  $\mathcal{V}$  are  $\ell_p$ -norm balls intersected with the positive orthant. We obtain convex relaxations of non-negative matrix factorization. Note that for  $p = \infty$ , it comes with approximation guarantees in terms of the value of the gauge function  $\Theta$ .

Table 1: Examples of structured sets  $\mathcal{U}$ , together with a computable surrogate  $\mathcal{C}_{\mathcal{U}}$  of  $\mathcal{P}_{\mathcal{U}} = \{uu^\top, u \in \mathcal{U}\}$ , that will be used in Section 5. Non-negativity constraints  $u \geq 0$  may be added to  $\mathcal{U}$  with an additional constraint  $M \geq 0$  for  $\mathcal{C}_{\mathcal{U}}$ .

	$\mathcal{U}$	$\mathcal{C}_{\mathcal{U}}$
$\ell_p$ -ball, $p \in [1, 2]$	$\ u\ _p \leq 1$	$M \succcurlyeq 0, \ M\ _p \leq 1, \text{tr } M \leq 1$
$\ell_p$ -ball, $p \in [2, +\infty]$	$\ u\ _p \leq 1$	$M \succcurlyeq 0, \ \text{diag}(M)\ _{p/2} \leq 1$
sparse coding	$\nu\ u\ _2^2 + (1-\nu)\ u\ _1^2 \leq 1$	$M \succcurlyeq 0, \nu \text{tr } M + (1-\nu)\ M\ _1 \leq 1$
binary coding	$u \in \{0, 1\}^d$	$M \succcurlyeq mm^\top, M \geq 0, m = \text{diag}(M) \leq 1$

–  $\mathcal{U} = \{0, 1\}^n$  and  $\gamma_{\mathcal{V}} = \|\cdot\|_2$ : this corresponds to decomposing the matrix  $X$  with a factor  $U$  with binary values. We will consider this example in our experiments in Section 7. Extensions may be considered by (a) adding also a constraint that less than  $k$  components of  $u$  are non-zeros (i.e., by adding a constraint  $u^\top 1_n \leq k$ ), or (b) considering  $\mathcal{U} = \{\frac{1}{F(A)}\}$ ,  $A \subset V$ ,  $A \neq \emptyset$ , for a non-decreasing submodular function  $F$ , leading to  $\gamma_{\mathcal{U}}(x) = f(x) + I_{\mathbb{R}_+^n}(x)$  where  $f$  is the Lovász extension of  $F$  [24, 4].

### 4.3 Special case: $\gamma_{\mathcal{V}} = \|\cdot\|_2$

Computing the polar of  $\Theta$  then corresponds to a quadratic maximization problem:

$$\Theta^\circ(Y)^2 = \max_{u \in \mathcal{U}} u^\top Y Y^\top u = \max_{u \in \mathcal{U}^{\circ\circ}} u^\top Y Y^\top u = \max_{\gamma_{\mathcal{U}}(u) \leq 1} u^\top Y Y^\top u. \quad (4)$$

We also have a variational quadratic representation corresponding to a rotation-invariant gauge function (Section 3.5); following Section 3.1, we denote by  $\mathcal{P}_{\mathcal{U}}$  the set  $\{uu^\top, u \in \mathcal{U}\}$ . There is an explicit representation of  $\Theta$  and  $\Theta^\circ$  in terms of  $\mathcal{P}_{\mathcal{U}}^{\circ\circ}$  (the convex hull of  $\mathcal{P}_{\mathcal{U}} \cup \{0\}$ ).

**Proposition 8** *We have, for all  $X, Y \in \mathbb{R}^{n \times d}$ :*

$$\Theta(X)^2 = \inf_{M \in \mathcal{P}_{\mathcal{U}}^{\circ\circ}} \text{tr } X^\top M^{-1} X \quad \text{and} \quad \Theta^\circ(Y)^2 = \max_{M \in \mathcal{P}_{\mathcal{U}}^{\circ\circ}} \text{tr } M Y Y^\top.$$

**Proof** The representation of  $\Theta^\circ$  is obvious from Eq. (4). We then have, for any  $X \in \mathbb{R}^{n \times d}$ :

$$\begin{aligned} \Theta(X)^2 &= \sup_{Y \in \mathbb{R}^{n \times d}} \text{tr } X^\top Y - 2\Theta^\circ(Y)^2 \text{ like all polar pairs,} \\ &= \sup_{Y \in \mathbb{R}^{n \times d}} \text{tr } X^\top Y - 2 \max_{M \in \mathcal{P}_{\mathcal{U}}^{\circ\circ}} \text{tr } Y Y^\top M \text{ from Eq. (4),} \\ &= \inf_{M \in \mathcal{P}_{\mathcal{U}}^{\circ\circ}} \sup_{Y \in \mathbb{R}^{n \times d}} \text{tr } X^\top Y - 2 \text{tr } Y Y^\top M \text{ by Fenchel duality,} \\ &= \inf_{M \in \mathcal{P}_{\mathcal{U}}^{\circ\circ}} \text{tr } X^\top M^{-1} X. \end{aligned}$$

Note that  $\Theta$  may not have full domain. ■

From the earlier representation, we get  $\Theta(X) = \inf_{M \succcurlyeq 0} \left\{ \frac{1}{2} \gamma_{\mathcal{P}_U}(M) + \frac{1}{2} \text{tr} X^\top M^{-1} X \right\}$ , by simply optimizing over the half line generated by  $M$ . We also have a representation involving all decompositions of  $X$  as  $UV^\top$ :

$$\Theta(X) = \min_{X=UV^\top} \frac{1}{2} \text{tr} VV^\top + \frac{1}{2} \gamma_{\mathcal{P}_U}(UU^\top), \quad (5)$$

which is straightforward from the definitions and interpretations of  $\Theta$  and  $\gamma_{\mathcal{P}_U}$  as a gauge function (Prop. 2). In the next section, we show how to compute  $U$  and  $V$  from the solution of a certain convex problem—note however that this does not lead to the optimal  $U$  and  $V$  in Eq. (3). See more details in Section 6.

#### 4.4 General case

We now consider all possible cases, beyond  $\gamma_{\mathcal{V}} = \|\cdot\|_2$  (where  $\gamma_{\mathcal{P}_V}(M) = \text{tr} M$ ). We only assume that  $\mathcal{U}$  and  $\mathcal{V}$  are compact and contain zero in their hulls. It is tempting to consider an extension of Eq. (5), by considering

$$\tilde{\Theta}(X) = \min_{X=UV^\top} \frac{1}{2} [\gamma_{\mathcal{P}_V}(VV^\top) + \gamma_{\mathcal{P}_U}(UU^\top)] = \min_{X=UV^\top} \sqrt{\gamma_{\mathcal{P}_V}(VV^\top) \gamma_{\mathcal{P}_U}(UU^\top)}. \quad (6)$$

However, when neither  $\mathcal{U}$  nor  $\mathcal{V}$  are unit  $\ell_2$ -balls, this only provides a lower-bound to  $\Theta$ , because the optimal decompositions of  $UU^\top$  and  $VV^\top$  for respectively  $\gamma_{\mathcal{P}_U}$  and  $\gamma_{\mathcal{P}_V}$  may not lead to an optimal decomposition of  $X$ . Since all our further relaxations in Section 5 will work directly on  $\tilde{\Theta}$ , we study precisely the properties of  $\tilde{\Theta}$ . We first show that Eq. (6) defines a gauge function and give a novel expression of its polar, as a maximum of weighted nuclear norms (note that the function  $(M, N) \mapsto \|M^{1/2} Y N^{1/2}\|_*$  is concave on  $\mathcal{S}_n^+ \times \mathcal{S}_d^+$ ).

**Proposition 9** *Let  $\tilde{\Theta}$  be defined in Eq. (6). Then,  $\tilde{\Theta}$  is a gauge function (convex, positively homogeneous, and non-negative), and its polar has the expression:*

$$\forall Y \in \mathbb{R}^{n \times d}, \quad \tilde{\Theta}^\circ(Y) = \sup_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}} \|M^{1/2} Y N^{1/2}\|_*. \quad (7)$$

Moreover, for all  $X, Y \in \mathbb{R}^{n \times d}$ ,  $\Theta(X) \geq \tilde{\Theta}(X)$  and  $\Theta^\circ(Y) \leq \tilde{\Theta}^\circ(Y)$ .

**Proof** The three properties of  $\tilde{\Theta}$  are straightforward. We have, for  $Y \in \mathbb{R}^{n \times d}$ :

$$\begin{aligned} \tilde{\Theta}^\circ(Y) &= \sup_{\gamma_{\mathcal{P}_V}(VV^\top) \gamma_{\mathcal{P}_U}(UU^\top) \leq 1} \text{tr} Y^\top UV^\top = \sup_{\gamma_{\mathcal{P}_V}(VV^\top) \leq 1, \gamma_{\mathcal{P}_U}(UU^\top) \leq 1} \text{tr} Y^\top UV^\top \\ &= \sup_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}} \sup_{UU^\top=M, VV^\top=N} \text{tr} Y^\top UV^\top \\ &= \sup_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}} \sup_{UU^\top=I, VV^\top=I} \text{tr} Y^\top M^{1/2} UV^\top N^{1/2} = \sup_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}} \|M^{1/2} Y N^{1/2}\|_*, \end{aligned}$$

because  $\sup_{UU^\top=I, VV^\top=I} \text{tr} Z^\top UV^\top = \|Z\|_*$ . Finally, we have for some  $u_m, v_m$ ,  $\Theta(X) = \frac{1}{2} \sum_m \{ \gamma_U(u_m)^2 + \gamma_V(v_m)^2 \}$ . Thus  $\Theta(X) \geq \tilde{\Theta}(X)$  because of Prop. 2, which implies the

reverse inequality for polars. ■

We may now give several equivalent expressions for  $\tilde{\Theta}$  and  $\tilde{\Theta}^\circ$  which are obtained using usual convex duality.

**Proposition 10** *Let  $\tilde{\Theta}$  be defined in Eq. (6). We have:*

$$\begin{aligned}\tilde{\Theta}(X) &= \inf_{M \in \mathcal{S}_n, N \in \mathcal{S}_d} \frac{1}{2} \gamma_{\mathcal{P}_U}(M) + \frac{1}{2} \gamma_{\mathcal{P}_V}(N) \text{ s.t. } \begin{pmatrix} M & X \\ X^\top & N \end{pmatrix} \succcurlyeq 0 \\ &= \sup_{Q \in \mathcal{P}_U^\circ, S \in \mathcal{P}_V^\circ, Z \in \mathbb{R}^{n \times d}} \frac{1}{2} \text{tr} \begin{pmatrix} Q & Z \\ Z^\top & S \end{pmatrix} \begin{pmatrix} 0 & X \\ X^\top & 0 \end{pmatrix} \text{ s.t. } \begin{pmatrix} Q & Z \\ Z^\top & S \end{pmatrix} \succcurlyeq 0, \\ \tilde{\Theta}^\circ(Y) &= \inf_{Q \in \mathcal{S}_n, S \in \mathcal{S}_d} \frac{1}{2} \gamma_{\mathcal{P}_U^\circ}(Q) + \frac{1}{2} \gamma_{\mathcal{P}_V^\circ}(S) \text{ s.t. } \begin{pmatrix} Q & Y \\ Y^\top & S \end{pmatrix} \succcurlyeq 0 \\ &= \sup_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}, Z \in \mathbb{R}^{n \times d}} \frac{1}{2} \text{tr} \begin{pmatrix} M & Z \\ Z^\top & N \end{pmatrix} \begin{pmatrix} 0 & Y \\ Y^\top & 0 \end{pmatrix} \text{ s.t. } \begin{pmatrix} M & Z \\ Z^\top & N \end{pmatrix} \succcurlyeq 0.\end{aligned}$$

**Proof** The four expressions are equivalent by taking polars or using Lagrangian duality for the semi-definite cone. The only element left to prove is that any of these four statements are correct. We have:

$$\max_{Z \in \mathbb{R}^{n \times d}} \text{tr} Z^\top Y \text{ s.t. } \begin{pmatrix} M & Z \\ Z^\top & N \end{pmatrix} \succcurlyeq 0 = \max_{\|T\|_{\text{op}} \leq 1} \text{tr} Y^\top M^{1/2} T N^{1/2} = \|M^{1/2} Y N^{1/2}\|_*,$$

which leads to desired result from Prop. 9. Note that we may obtain directly a relationship between  $\tilde{\Theta}^\circ$  and  $\Theta^\circ$  as follows:

$$\begin{aligned}\Theta^\circ(Y) &= \max_{u \in \mathcal{U}, v \in \mathcal{V}} v^\top Y^\top u = \max_{u \in \mathcal{U}, v \in \mathcal{V}} \frac{1}{2} \text{tr} \begin{pmatrix} uu^\top & uv^\top \\ vu^\top & vv^\top \end{pmatrix} \begin{pmatrix} 0 & Y \\ Y^\top & 0 \end{pmatrix} \\ &= \max_{M \in \mathcal{P}_U, N \in \mathcal{P}_V} \frac{1}{2} \text{tr} \begin{pmatrix} M & Z \\ Z^\top & N \end{pmatrix} \begin{pmatrix} 0 & Y \\ Y^\top & 0 \end{pmatrix} \text{ s.t. } \begin{pmatrix} M & Z \\ Z^\top & N \end{pmatrix} \succcurlyeq 0 \\ &\leq \max_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}} \frac{1}{2} \text{tr} \begin{pmatrix} M & Z \\ Z^\top & N \end{pmatrix} \begin{pmatrix} 0 & Y \\ Y^\top & 0 \end{pmatrix} \text{ s.t. } \begin{pmatrix} M & Z \\ Z^\top & N \end{pmatrix} \succcurlyeq 0 = \tilde{\Theta}^\circ(Y).\end{aligned}$$

■

Note that the representations of  $\tilde{\Theta}$  and  $\tilde{\Theta}^\circ$  have the same form, replacing  $\mathcal{P}_U$  by  $\mathcal{P}_U^\circ \cap \mathcal{S}_d^+$ . We thus get that for all  $X \in \mathbb{R}^{n \times d}$ :

$$\Theta(X) \geq \tilde{\Theta}(X) = \sup_{Q \in \mathcal{P}_U^\circ \cap \mathcal{S}_n^+, S \in \mathcal{P}_V^\circ \cap \mathcal{S}_d^+} \|Q^{1/2} Y S^{1/2}\|_*, \quad (8)$$

that is a variational formulation of  $\tilde{\Theta}$  through nuclear norms (note again that  $\mathcal{P}_U^\circ$  and  $\mathcal{P}_V^\circ$  may not be compact and  $\tilde{\Theta}(X)$  potentially infinite). This is to be contrasted with the following representation:

$$\Theta^\circ(Y)^2 = \sup_{M \in \mathcal{P}_U, N \in \mathcal{P}_V} \text{tr} M Y N Y^\top = \sup_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}} \text{tr} M Y N Y^\top = \sup_{M \in \mathcal{P}_U, N \in \mathcal{P}_V} \|M^{1/2} Y N^{1/2}\|_F^2,$$

which leads to

$$\begin{aligned}
\Theta(X)^2 &= \sup_{Y \in \mathbb{R}^{n \times d}} \inf_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}} \text{tr } X^\top Y - 2 \text{tr } M Y N Y^\top \\
&\leq \inf_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}} \sup_{Y \in \mathbb{R}^{n \times d}} \text{tr } X^\top Y - 2 \text{tr } M Y N Y^\top \text{ by weak duality,} \\
&= \inf_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}} \text{tr } M^{-1} X N^{-1} Y^\top, \\
\Theta(X) &\leq \inf_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}} \sqrt{\text{tr } M^{-1} X N^{-1} Y^\top} = \inf_{M \in \mathcal{P}_U^{\circ\circ}, N \in \mathcal{P}_V^{\circ\circ}} \|M^{-1/2} X N^{-1/2}\|_F,
\end{aligned}$$

which is a variational upper-bound as a weighted Frobenius norm, which is not computable as a convex program. Instead, alternate minimization could be used, each of the steps being doable as a convex program.

**Relationships between  $Q$ ,  $S$ ,  $M$ ,  $N$ ,  $U$  and  $V$ .** Given the primal/dual solutions  $(M, N)$ ,  $(Q, S)$  of the dual convex optimization problems defining  $\tilde{\Theta}$  in Prop. 10, we have the following optimality conditions: Assume  $Q = AA^\top$  with  $A$  full rank (i.e.,  $A^\top A$  invertible), and  $S = BB^\top$  with  $B$  full rank. Then an optimal  $Z$  is  $Z = AGH^\top B^\top$ , with  $A^\top X B = G \text{Diag}(s) H^\top$  is the singular value decomposition of  $A^\top X B$ . We have

$$\begin{pmatrix} M & -X \\ -X^\top & N \end{pmatrix} \succcurlyeq 0 \text{ and } \begin{pmatrix} M & -X \\ -X^\top & N \end{pmatrix} \begin{pmatrix} AA^\top & AGH^\top B^\top \\ BHG^\top A^\top & BB^\top \end{pmatrix} = 0,$$

which implies (since  $A$  and  $B$  have full rank):

$$\begin{pmatrix} M & -X \\ -X^\top & N \end{pmatrix} \begin{pmatrix} A & AGH^\top \\ BVU^\top & B \end{pmatrix} = 0.$$

This leads to  $A^\top M A = A^\top X B H G^\top = G \text{Diag}(s) G^\top$ . We then take as candidates  $U = (AA^\top)^\dagger A G \text{Diag}(s)^{1/2}$  and  $V = (BB^\top)^\dagger B H \text{Diag}(s)^{1/2}$ , leading to  $U U^\top = M$  and  $V V^\top = N$ . Moreover,  $X = U V^\top$  and  $\tilde{\Theta}(X) = \|A^\top X B\|_* = \text{tr } \text{Diag}(s)$ . Finally, by optimality of  $Q$ , we have  $\gamma_{\mathcal{P}_U}(U U^\top) = \text{tr } U U^\top Q = \text{tr } A^\top M A = \text{tr } \text{Diag}(s)$  and  $\gamma_{\mathcal{P}_V}(V V^\top) = \text{tr } V V^\top S = \text{tr } B^\top N B = \text{tr } \text{Diag}(s)$ , and thus  $(U, V)$  is an optimal decomposition for Eq. (6).

**Validity for other sets than  $\mathcal{P}_U$  and  $\mathcal{P}_V$ .** The duality properties presented in Prop. 10 are valid for all convex sets of positive semi-definite matrices that contain zero (and in particular for the ones in Section 5). This will be used in Section 5 to get computable approximations (with guarantees).

## 4.5 Non-convex approaches

Finding approximations of the gauge function  $\Theta$  or its polar  $\Theta^\circ$  has been tackled thoroughly through non-convex approaches, which may only find stationary points of the associated optimization problems. In this paper, we briefly review two approaches based on alternating minimization.

**Computing polar function through a generalized power method.** The task is to find maximizers  $(u, v) \in \mathcal{U} \times \mathcal{V}$  of  $u^\top Y v$ , for a certain  $Y \in \mathbb{R}^{n \times d}$ . The generalized power method iterates the recursion  $u \in \arg \max_{u \in \mathcal{U}} u^\top Y v$  and  $v \in \arg \max_{v \in \mathcal{V}} u^\top Y v$  (i.e., alternate maximization), started from a certain  $u$  (typically random, or obtained from the relaxation in Section 5). See an analysis in the context of sparse principal component analysis in [32].

Each iteration increases  $u^\top Y v$ , and hence the values are converging. However, the iterates themselves do not converge in general, and when they do, may not converge to a global maximizer. There are certain situations where they do: when  $\mathcal{U}$  and  $\mathcal{V}$  are  $\ell_2$ -balls (in this case, this is the usual power method [26]), or when  $\mathcal{U}$  and  $\mathcal{V}$  are  $\ell_p$ -balls and  $Y$  has only non-negative elements [15].

In Section 6, obtaining approximate maximizers  $(u, v)$  will be key for the generalized conditional gradient algorithm and simplicial methods. In our experiments in Section 7, the power method leads to better values once initialized from results of the convex relaxations presented in Section 5.

**Alternating optimization.** We consider  $r \leq np$  elements and minimize either the sum  $\sum_{m=1}^r \gamma_{\mathcal{U}}(u_m) \gamma_{\mathcal{V}}(v_m)$  or  $\sum_{m=1}^r \left\{ \frac{1}{2} \gamma_{\mathcal{U}}(u_m)^2 + \frac{1}{2} \gamma_{\mathcal{V}}(v_m)^2 \right\}$ , subject to  $X = \sum_{m=1}^r u_m v_m^\top$ . This may be done by alternating minimization, with no guarantees, beyond decreasing the value of the upper-bound on the norm.

Note that in many applications, what needs to be solved is an approximation problem of the form  $\min_{Z \in \mathbb{R}^{n \times d}} \frac{1}{2} \|X - Z\|_F^2 + \lambda \Theta(Z)$ , which may be tackled with similar algorithms. In fact, in our simulations in Section 7, we solve this problem as it leads to empirically easier optimization problems and more stable results.

**Non-convex optimization without local minima.** As all optimization problems on positive semi-definite matrices, computing the lower-bound  $\tilde{\Theta}(X)$  through Prop. 10 may be done using a low-rank representation of the matrix  $\begin{pmatrix} M & X \\ X^\top & N \end{pmatrix} = \begin{pmatrix} U \\ V \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix}^\top$ . As shown by [16], if the number of columns of  $U$  is greater than the rank of any solution of the semi-definite program, then the non-convex optimization problem has no local minima. However, this result should still be taken with caution in our context: (a) the non-convex problem has no local minima but still have many stationary points and care should still be taken when using iterative algorithms (see, e.g., [31], for approaches based on trust-region methods and manifold-based optimization); moreover, (b) the result only applies to  $\tilde{\Theta}(X)$  and not to  $\Theta(X)$ , (c) even in the situations where  $\tilde{\Theta}(X) = \Theta(X)$  (e.g., when of the two sets is an unit  $\ell_2$ -ball), such an approach requires to be able to compute  $\gamma_{\mathcal{P}_U}$ , which is not possible in many cases (the max-norm considered in [35] corresponds to the convex relaxation  $\gamma_{\mathcal{C}_U}$  of  $\gamma_{\mathcal{P}_U}$  presented in Section 5); (d) such local-search non-convex techniques cannot be used for finding a decomposition of  $M$  as  $UU^\top$  where all gauge functions  $\gamma_{\mathcal{U}}$  of columns of  $U$  are small (which can however be done using the convex optimization techniques presented in Section 6).

Nevertheless, in practice, alternating optimization techniques, although they come with no guarantees, still perform well when they can be applied (see Section 7).



## 5 Computable convex relaxations of decomposition norms

In the previous section, we have given several representations of the gauge function  $\Theta$  and its polar  $\Theta^\circ$ . They rely on the convex hull  $\mathcal{P}_\mathcal{U}^{\circ\circ}$  of the union of  $\{0\}$  and  $\mathcal{P}_\mathcal{U} = \{uu^\top, u \in \mathcal{U}\}$ . Thus the algorithms may have polynomial time if this convex hull may be described in polynomial time. This is notably the case where  $\mathcal{U}$  is the unit  $\ell_2$ -ball, in which case  $\mathcal{P}_\mathcal{U}^{\circ\circ} = \{M \in \mathcal{S}_n^+, \text{tr } M \leq 1\}$ . We are not aware of any other interesting case for which we have such semi-definite representations.

Before describing computable relaxations, we first refine the computability requirement which is weaker than having a good representation of  $\mathcal{P}_\mathcal{U}^{\circ\circ}$  or its polar  $\mathcal{P}_\mathcal{U}^\circ$ . We start with a simple lemma:

**Lemma 1** *Let  $\mathcal{A}, \mathcal{B} \subset \mathcal{S}_n^+$ . The following statements are equivalent:*

- (a)  $\forall M \in \mathcal{S}_n^+, \sup_{N \in \mathcal{A}} \text{tr } MN \leq \sup_{N \in \mathcal{B}} \text{tr } MN$ ,
- (b)  $\mathcal{A}^{\circ\circ} - \mathcal{S}_n^+ \subset \mathcal{B}^{\circ\circ} - \mathcal{S}_n^+$ ,
- (c)  $\mathcal{A}^\circ \cap \mathcal{S}_n^+ \supset \mathcal{B}^\circ \cap \mathcal{S}_n^+$ .

*Moreover, if  $\mathcal{A}$  is composed of rank-one matrices, then they are also equivalent to*

- (d)  $\forall x \in \mathbb{R}^n, \sup_{N \in \mathcal{A}} x^\top N x \leq \sup_{N \in \mathcal{B}} x^\top N x$ .

**Proof** (a) is equivalent to (c) by simple properties of gauge functions. (b) is equivalent to (c) because of polar calculus and  $(\mathcal{S}_n^+)^\circ = -\mathcal{S}_n^+$ , which imply  $(\mathcal{A}^\circ \cap \mathcal{S}_n^+)^\circ = \text{hull}(\mathcal{A}^{\circ\circ} \cup (-\mathcal{S}_n^+)) = \mathcal{A}^{\circ\circ} - \mathcal{S}_n^+$ . (a) trivially implies (d). We only need to show (d) implies (a). Assume (d) is true and let  $M = YY^\top \in \mathcal{S}_n^+$ , we have

$$\sup_{N \in \mathcal{A}} \text{tr } MN = \sup_{uu^\top \in \mathcal{A}} \sup_{\|v\|_2 \leq 1} (u^\top Y v)^2 \leq \sup_{\|v\|_2 \leq 1} \sup_{M \in \mathcal{B}} \text{tr } Y v v^\top Y^\top M \leq \sup_{M \in \mathcal{B}} \text{tr } Y^\top M Y,$$

hence the result. ■

Given the representation of  $\tilde{\Theta}$  in Prop. 10, a key consequence of the previous lemma is that in the representation of  $\tilde{\Theta}(X)$  and  $\tilde{\Theta}^\circ(Y)$  in Section 4, we may replace  $\mathcal{P}_\mathcal{U}^{\circ\circ}$  by any set  $\mathcal{C}_\mathcal{U}$  such that

$$\forall M \in \mathcal{S}_n^+, \sup_{u \in \mathcal{U}} u^\top M u = \sup_{N \in \mathcal{C}_\mathcal{U}} \text{tr } MN.$$

This allows us to treat the case where  $\mathcal{U}$  is the unit  $\ell_1$ -ball, for which  $\forall M \in \mathcal{S}_n^+, \max_{u \in \mathcal{U}} u^\top M u = \|\text{diag}(M)\|_\infty = \|M\|_\infty$ . We may thus take  $\mathcal{C}_\mathcal{U} = \{M \in \mathcal{S}_n^+, \|M\|_1 \leq 1\}$ . Note however that it is not true that  $\mathcal{P}_\mathcal{U}^{\circ\circ} = \{M \in \mathcal{S}_n^+, \|M\|_1 \leq 1\}$  (counter-examples may be found in [33]).

For all other cases, we need computable convex approximations  $\mathcal{C}_\mathcal{U}$  of  $\mathcal{P}_\mathcal{U}^{\circ\circ}$  such that for

$$\forall M \in \mathcal{S}_n^+, \sup_{u \in \mathcal{U}} u^\top M u \leq \sup_{N \in \mathcal{C}_\mathcal{U}} \text{tr } MN$$

(if we impose the property above for all symmetric matrices  $M$ , it is equivalent to  $\mathcal{P}_\mathcal{U}^{\circ\circ} \subset \mathcal{C}_\mathcal{U}$ ). Given Lemma 1, this is equivalent to

$$\forall y \in \mathbb{R}^n, \max\{\gamma_\mathcal{U}^\circ(y), \gamma_\mathcal{U}^\circ(-y)\}^2 = \sup_{u \in \mathcal{U}} (u^\top y)^2 \leq \sup_{N \in \mathcal{C}_\mathcal{U}} y^\top N y = \gamma_{\mathcal{C}_\mathcal{U}}^\circ(y y^\top), \quad (9)$$

that is, we need an *upper-bound* on  $\max\{\gamma_{\mathcal{U}}^{\circ}(y), \gamma_{\mathcal{U}}^{\circ}(-y)\}^2$ , which is convex in  $yy^{\top}$ . Note that given two sets  $\mathcal{C}_{\mathcal{U}}$  that satisfy Eq. (9), their intersection also does and the relaxation is then always tighter. We can therefore use several properties of the set  $\mathcal{U}$ , that may then be combined together:

- *Representation of  $\gamma_{\mathcal{U}}$  as minima of quadratic functions* (Section 3.2): this corresponds to the case where  $\gamma_{\mathcal{U}}^{\circ}$  is represented as a maximum of quadratic functions, i.e., Eq. (9) is an equality. In this situation, the set  $\mathcal{C}_{\mathcal{U}}$  is used directly. For example, when  $\mathcal{U}$  is the unit  $\ell_p$ -ball, we may choose  $\mathcal{C}_{\mathcal{U}} = \{M \in \mathcal{S}_n^+, \|M\|_p \leq 1\}$ .
- *Representation of  $\gamma_{\mathcal{U}}$  as maxima of quadratic functions* (Section 3.1): as shown in Section 3.2, if  $\gamma_{\mathcal{U} \cap (-\mathcal{U})}(x)^2 = \sup_{M \in \mathcal{K}_{\mathcal{U}}} x^{\top} M x$ , then Eq. (9) is satisfied for  $\mathcal{C}_{\mathcal{U}} = \mathcal{K}_{\mathcal{U}}^{\circ} \cap \mathcal{S}_n^+$ . Note that reduced representations may also be useful. For example, when  $\gamma_{\mathcal{U}}^{\circ}$  has a diagonal representation outlined in Section 3.4, we may choose  $\mathcal{K}_{\mathcal{U}}$  to be of the form  $\{\text{Diag}(\eta), \eta \in \mathcal{H}_{\mathcal{U}}\}$ , then  $\mathcal{K}_{\mathcal{U}}^{\circ} \cap \mathcal{S}_n^+ = \{M \in \mathcal{S}_n^+, \text{diag}(M) \in \mathcal{H}_{\mathcal{U}}\}$  may be considered as a set  $\mathcal{C}_{\mathcal{U}}$ . For example, when  $\mathcal{U}$  is the  $\ell_p$ -norm unit ball, for  $p \geq 2$ , we obtain  $\mathcal{C}_{\mathcal{U}} = \{M \in \mathcal{S}_n^+, \|\text{diag}(M)\|_{p/2} \leq 1\}$ .
- *Positivity*: if  $\mathcal{U} \subset \mathbb{R}_+^n$ , then  $\mathcal{P}_{\mathcal{U}}^{\circ\circ} \subset \{M \in \mathcal{S}_n^+, M \geq 0\}$ .
- *Linear inequalities*: any matrix  $M$  such that  $\max_{\gamma_{\mathcal{U}}(u) \leq 1} u^{\top} M u$  may be computed or efficiently upper-bounded by  $h$  adds another constraint of the form  $\text{tr} M U \leq h$ . Typically,  $M = I$  is a good candidate, as this defines the equivalence between the gauge function  $\gamma_{\mathcal{U}}$  and  $\|\cdot\|_2$ .

In Table 1, we present relaxations for the examples we consider in this paper. For all relaxations we consider (and also with positivity constraints), we have:  $\mathcal{C}_{\mathcal{U}} \cap \{M, \text{rank}(M) = 1\} = \mathcal{P}_{\mathcal{U}}$ . This implies that for all  $u \in \text{dom}(\gamma_{\mathcal{U}})$ ,  $\gamma_{\mathcal{P}_{\mathcal{U}}}(uu^{\top}) = \gamma_{\mathcal{C}_{\mathcal{U}}}(uu^{\top}) = \min\{\gamma_{\mathcal{U}}(u), \gamma_{\mathcal{U}}(-u)\}^2$ . Moreover, for all relaxations in Table 1 (without positivity constraints), we have equality in Eq. (9), a property that will be useful when deriving performance guarantees in the next sections (Prop. 12).

## 5.1 Performance guarantees

Our computable relaxations, will be based on replacing the convex sets  $\mathcal{P}_{\mathcal{U}}^{\circ\circ}$  and  $\mathcal{P}_{\mathcal{V}}^{\circ\circ}$  by surrogates  $\mathcal{C}_{\mathcal{U}} \subset \mathcal{S}_n^+$  and  $\mathcal{C}_{\mathcal{V}} \subset \mathcal{S}_d^+$  such that  $\mathcal{C}_{\mathcal{U}} - \mathcal{S}_n^+ \subset \mathcal{P}_{\mathcal{U}} - \mathcal{S}_n^+$  and  $\mathcal{C}_{\mathcal{V}} - \mathcal{S}_d^+ \subset \mathcal{P}_{\mathcal{V}} - \mathcal{S}_d^+$  (see Lemma 1). This leads to the relaxations, for  $X, Y \in \mathbb{R}^{n \times d}$ :

$$\Theta(X) \geq \inf_{M \in \mathcal{S}_n^+, N \in \mathcal{S}_d^+} \frac{1}{2} \gamma_{\mathcal{C}_{\mathcal{U}}}(M) + \frac{1}{2} \gamma_{\mathcal{C}_{\mathcal{V}}}(N) \text{ s.t. } \begin{pmatrix} M & X \\ X^{\top} & N \end{pmatrix} \succcurlyeq 0 \quad (10)$$

$$= \sup_{Q \in \mathcal{C}_{\mathcal{U}}^{\circ} \cap \mathcal{S}_n^+, S \in \mathcal{C}_{\mathcal{V}}^{\circ} \cap \mathcal{S}_d^+} \|Q^{1/2} X S^{1/2}\|_*, \quad (11)$$

$$\Theta^{\circ}(Y) \leq \sup_{M \in \mathcal{C}_{\mathcal{U}}, N \in \mathcal{C}_{\mathcal{V}}} \|M^{1/2} Y N^{1/2}\|_*, \quad (12)$$

which specializes to the following expressions when  $\mathcal{V} = \{v \in \mathbb{R}^d, \|v\|_2 \leq 1\}$ :

$$\Theta(X) \geq \inf_{M \in \mathcal{S}_n^+} \frac{1}{2} \gamma_{\mathcal{C}_U}(M) + \frac{1}{2} \text{tr} X^\top M^{-1} X \quad (13)$$

$$= \sup_{Q \in \mathcal{C}_U^\circ \cap \mathcal{S}_n^+} \|Q^{1/2} X\|_*, \quad (14)$$

$$\Theta^\circ(Y) \leq \sup_{M \in \mathcal{C}_U} \sqrt{\text{tr} Y^\top M Y}. \quad (15)$$

Because the bounds are also polar to each other (from Prop. 10), if we have a performance guarantee for  $\Theta^\circ$ , i.e.,

$$\forall Y \in \mathbb{R}^{n \times d}, \Theta^\circ(Y) \leq \sup_{M \in \mathcal{C}_U, N \in \mathcal{C}_V} \|M^{1/2} Y N^{1/2}\|_* \leq \kappa \Theta^\circ(Y), \quad (16)$$

for some  $\kappa \geq 1$ , we immediately get

$$\forall X \in \mathbb{R}^{n \times d}, \Theta(X) \geq \sup_{Q \in \mathcal{C}_U^\circ \cap \mathcal{S}_n^+, S \in \mathcal{C}_V^\circ \cap \mathcal{S}_d^+} \|Q^{1/2} X S^{1/2}\|_* \geq \frac{1}{\kappa} \Theta(X). \quad (17)$$

We now consider two situations, where we do have such performance guarantees like Eq. (16) and Eq. (17). The first situation corresponds to constraints on only diagonal elements of matrices (such as for  $\mathcal{U}$  and  $\mathcal{V}$  equal to  $\ell_p$ -norm balls for  $p \geq 2$ ), and leads to dimension-independent guarantees; it was derived by [42]:

**Proposition 11** *Assume  $\mathcal{C}_U = \{M \in \mathcal{S}_n^+, \text{diag}(M) \in \mathcal{H}_U\}$  and  $\mathcal{C}_V = \{M \in \mathcal{S}_d^+, \text{diag}(M) \in \mathcal{H}_V\}$  for some convex sets  $\mathcal{H}_U \subset \mathbb{R}_+^n$  and  $\mathcal{H}_V \subset \mathbb{R}_+^d$ , and that  $\mathcal{C}_U \cap \{M, \text{rank}(M) = 1\} = \mathcal{P}_U$ ,  $\mathcal{C}_V \cap \{M, \text{rank}(M) = 1\} = \mathcal{P}_V$ . Then Eq. (16) and Eq. (17) are valid for  $\kappa = 1/(2\sqrt{3}/\pi - 2/3) < 2.3$ . If moreover,  $\mathcal{V}$  is the unit  $\ell_2$ -ball, they are also valid for  $\kappa = \sqrt{\pi/2} < 1.3$ .*

**Proof** We only consider the case where  $\mathcal{V}$  is the unit disk, which gives the main idea of the proof of [42]. We consider a maximizer  $M$  of Eq. (15) and a vector  $u = D^{1/2} \text{sign}(v)$  where  $v$  is a random sample from a normal distribution with mean 0 and covariance matrix  $D^{-1/2} M D^{-1/2}$ , with  $D = \text{Diag}(\text{diag}(M))$ . We have, using standard arguments from [42, 25],  $uu^\top \in \mathcal{C}_U \cap \{\text{rank}(M) = 1\} = \mathcal{P}_U$  and

$$\begin{aligned} \Theta^\circ(Y)^2 &\geq \mathbb{E}[u^\top Y Y^\top u] = \mathbb{E} \text{tr} D^{1/2} Y Y^\top D^{1/2} \text{sign}(v) \text{sign}(v)^\top \\ &= \frac{2}{\pi} \text{tr} D^{1/2} Y Y^\top D^{1/2} \arcsin[D^{-1/2} M D^{-1/2}] \\ &\geq \frac{2}{\pi} \text{tr} D^{1/2} Y Y^\top D^{1/2} [D^{-1/2} M D^{-1/2}] = \frac{2}{\pi} \text{tr} Y^\top M Y = \frac{2}{\pi} \sup_{M \in \mathcal{C}_U} \text{tr} Y^\top M Y. \end{aligned}$$

The general case is also based on sampling and normalizing Gaussian random variables (see more details in [42]). ■

The other situation corresponds to situations where the gauge functions  $\gamma_U$ ,  $\gamma_V$  and their polars have certain *exact* quadratic representations (for example, for  $\mathcal{U}$  and  $\mathcal{V}$  equal to  $\ell_p$ -norm balls for any  $p \in [1, \infty]$ ).

**Proposition 12** *Assume that for all  $w \in \text{span}(\mathcal{U})$  and  $z \in \text{span}(\mathcal{V})$ , we have the representations  $\max\{\gamma_{\mathcal{U}}^{\circ}(w), \gamma_{\mathcal{U}}^{\circ}(-w)\}^2 = \gamma_{\mathcal{C}_{\mathcal{U}}}^{\circ}(ww^{\top})$  and  $\max\{\gamma_{\mathcal{V}}^{\circ}(z), \gamma_{\mathcal{V}}^{\circ}(-z)\}^2 = \gamma_{\mathcal{C}_{\mathcal{V}}}^{\circ}(zz^{\top})$ . Then Eq. (16) and Eq. (17) are valid for  $\kappa = \min\{n, d\}$ . If moreover,  $\mathcal{V}$  is the unit disk, they are also valid for  $\kappa = \sqrt{\min\{n, d\}}$ .*

**Proof** Our assumptions imply by duality that for all  $w \in \text{span}(\mathcal{U})$ ,  $\gamma_{\mathcal{U} \cup (-\mathcal{U})}(w)^2 = \inf_{N \in \mathcal{C}_{\mathcal{U}}} w^{\top} N^{-1} w = G_{\mathcal{U}}(ww^{\top})$ , with  $G_{\mathcal{U}}(M) = \inf_{N \in \mathcal{C}_{\mathcal{U}}} MN^{-1}$ , which is defined as, with  $M = WW^{\top} \in \mathcal{C}_{\mathcal{U}}$ , the minimum of  $\text{tr } B$  s.t.  $\begin{pmatrix} A & W \\ W^{\top} & N \end{pmatrix} \succcurlyeq 0$ . By taking  $N = WW^{\top}$  and  $A = I$ , we obtain that  $G_{\mathcal{U}}(M) \leq \text{rank}(M)$ .

We first consider the case where  $\mathcal{V}$  is the unit disk. We consider a solution  $M$  of Eq. (15) and a vector  $u$  which is random sample from a normal distribution with mean 0 and covariance matrix  $M$ . By definition of  $\mathcal{C}_{\mathcal{U}}$ ,  $u$  must belong to  $\text{span}(\mathcal{U})$ . Let  $r \leq \min\{n, d\}$  be the rank of  $M$ . We have:  $\mathbb{E}[u^{\top} Y Y^{\top} u] = \text{tr } Y^{\top} M Y$ . Moreover, by Jensen's inequality (and concavity of  $G_{\mathcal{U}}$ ):

$$\mathbb{E} \gamma_{\mathcal{U} \cup (-\mathcal{U})}(u)^2 = \mathbb{E} G_{\mathcal{U}}(uu^{\top}) \leq G_{\mathcal{U}}(\mathbb{E} uu^{\top}) \leq r \leq \min\{n, d\}.$$

Thus for  $\alpha = \frac{1}{\min\{n, d\}} \sup_{N \in \mathcal{C}_{\mathcal{U}}} \text{tr } Y^{\top} N Y = \frac{1}{\min\{n, d\}} \text{tr } Y^{\top} M Y$ , we have  $\mathbb{E}[u^{\top} Y Y^{\top} u - \alpha \gamma_{\mathcal{U} \cup (-\mathcal{U})}(u)^2] \geq 0$ . Therefore, since  $\gamma_{\mathcal{U}}$  is not uniformly equal to zero (because we have assumed  $\mathcal{U}$  compact), there exists  $u$  such that  $\gamma_{\mathcal{U}}(u) > 0$  and  $u^{\top} Y Y^{\top} u - \alpha \gamma_{\mathcal{U}}(u)^2 \geq 0$ , hence the claim.

For the general case, we sample  $(u, v)$  from a normal distribution with mean zero and covariance matrix  $\begin{pmatrix} M & Z \\ Z^{\top} & N \end{pmatrix}$ . We then have  $\mathbb{E}[u^{\top} Y v] = \text{tr } Z^{\top} Y$ , and  $\mathbb{E} \gamma_{\mathcal{U} \cup (-\mathcal{U})}(u)^2 \leq \min\{n, d\}$ ,  $\mathbb{E} \gamma_{\mathcal{V} \cup (-\mathcal{V})}(v)^2 \leq \min\{n, d\}$ . We have, with  $\rho$  equal to the relaxed value obtained in Eq. (12):

$$\mathbb{E} \left[ u^{\top} Y v - \frac{\rho}{\min\{n, d\}} \left( \frac{1}{2} \gamma_{\mathcal{U} \cup (-\mathcal{U})}(u)^2 + \frac{1}{2} \gamma_{\mathcal{V} \cup (-\mathcal{V})}(v)^2 \right) \right] \geq 0.$$

Thus, there exists  $u, v$  such that  $\gamma_{\mathcal{U}}(u) > 0$ ,  $\gamma_{\mathcal{V}}(v) > 0$  and  $u^{\top} Y v - \frac{\rho}{\min\{n, d\}} \left( \frac{1}{2} \gamma_{\mathcal{U}}(u)^2 + \frac{1}{2} \gamma_{\mathcal{V}}(v)^2 \right)$ . This implies that Eq. (16) is valid for  $\kappa = \min\{n, d\}$ .

Note that we may also obtain the same results without randomization. When  $\gamma_{\mathcal{V}} = \|\cdot\|_2$ , we have:

$$\Theta^{\circ}(Y)^2 = \max_{\|v\|_2 \leq 1} \gamma_{\mathcal{U}}^{\circ}(Yv)^2 = \max_{\|v\|_2 \leq 1} \max_{M \in \mathcal{C}_{\mathcal{U}}} v^{\top} Y^{\top} M Y v = \max_{M \in \mathcal{C}_{\mathcal{U}}} \lambda_{\max}(Y^{\top} M Y).$$

Thus by taking  $v$  the largest eigenvector of  $Y^{\top} M Y$  and  $u \in \text{argmax}_{u \in \mathcal{U}} u^{\top} Y v$ , we have

$$(u^{\top} Y v)^2 \geq v^{\top} Y^{\top} M Y v = \lambda_{\max}(Y^{\top} M Y) \geq \frac{1}{\min\{n, d\}} \text{tr } Y^{\top} M Y.$$

In the general case, we simply use

$$\begin{aligned} \Theta^{\circ}(Y)^2 &= \max_{v \in \mathcal{V}} \max_{M \in \mathcal{C}_{\mathcal{U}}} v^{\top} Y^{\top} M Y v = \max_{M \in \mathcal{C}_{\mathcal{U}}} \max_{N \in \mathcal{C}_{\mathcal{V}}} \max_{v^{\top} N^{-1} v \leq 1} v^{\top} Y^{\top} M Y v \\ &= \max_{M \in \mathcal{C}_{\mathcal{U}}, N \in \mathcal{C}_{\mathcal{V}}} \lambda_{\max}(N^{1/2} Y^{\top} M Y N^{1/2}) \leq \max_{M \in \mathcal{C}_{\mathcal{U}}, N \in \mathcal{C}_{\mathcal{V}}} \|N^{1/2} Y^{\top} M^{1/2}\|_*^2, \end{aligned}$$

with an approximation ratio of  $\min\{n, p\}$  between the operator norm and the nuclear norm.

■

The bound from the previous proposition is dimension-dependent. However, this is a non-trivial result. For example, with non-negativity constraints, it is not easy to have any guarantee (and our bounds do not). Indeed, For  $\mathcal{V}$  being the  $\ell_2$ -norm and  $\mathcal{U}$  the positive  $\ell_1$ -norm, then computing the polar corresponds to maximizing a quadratic form with positive constraints, which is notoriously difficult [11].

The two previous propositions lead to candidates for vectors  $u$  and  $v$ , through sampling from normal distributions with mean zero and covariance matrix equal to  $M$  (or a normalized version thereof). Note that other possibilities are available, like taking a largest eigenvector (such as done in the proof of Prop. 12) and running the power method from Section 4.5 starting from any of the above candidates.

**Approximation of  $\gamma_{\mathcal{P}_{\mathcal{U}}}$ .** Interestingly, computing  $\gamma_{\mathcal{P}_{\mathcal{U}}}$ , the gauge function considered in Prop. 2 which corresponds to factorizing positive semi-definite matrices with elements  $uu^\top$ ,  $u \in \mathcal{U}$ , is harder than computing the norm  $\Theta$  for  $\mathcal{V}$  equal to the unit  $\ell_2$ -ball. Indeed, the polar of  $\gamma_{\mathcal{P}_{\mathcal{U}}}$  is defined by  $\gamma_{\mathcal{P}_{\mathcal{U}}}^\circ(M) = \sup_{u \in \mathcal{U}} (u^\top M u)_+ = (\sup_{u \in \mathcal{U}} u^\top M u)_+$ , and is finite for *all* symmetric matrices, and in order to have an approximation ratio for  $\gamma_{\mathcal{P}_{\mathcal{U}}}$  we need an approximation ratio for its polar (and not only for its polar restricted to  $\mathcal{S}_n^+$ ). In the two types of guarantees above, the one based on diagonal variational representations may not be easily extended (note that this would notably imply that we would have a polynomial test of complete positivity, which is unlikely). However, if we have a variational formulation of the polar  $\gamma_{\mathcal{U}}^\circ$ , then the bound of Prop. 12 still applies (the two proof techniques may be easily extended). Note that we now impose that  $\mathcal{P}_{\mathcal{U}} \subset \mathcal{C}_{\mathcal{U}}$ .

**Proposition 13** *Assume that  $\mathcal{P}_{\mathcal{U}} \subset \mathcal{C}_{\mathcal{U}}$ , and that for all  $w \in \mathbb{R}^n$ , we have  $\gamma_{\mathcal{C}_{\mathcal{U}}}^\circ(w w^\top) = \max\{\gamma_{\mathcal{U}}^\circ(w), \gamma_{\mathcal{U}}^\circ(-w)\}^2$ . Then, for all  $M \in \mathcal{S}_n^+$ ,  $\gamma_{\mathcal{P}_{\mathcal{U}}}(M) \geq \gamma_{\mathcal{C}_{\mathcal{U}}}(M) \geq \frac{1}{n} \gamma_{\mathcal{P}_{\mathcal{U}}}(M)$ .*

## 5.2 Random sampling

An interesting alternative to semidefinite relaxations is to consider random sampling of vectors  $u$  and  $v$  and approximate  $\Theta^\circ(Y)$  by the maximum of  $u^\top Y v$  over these samples. This approach was considered by [44] in the context of the nuclear norm. In this section, for simplicity, we consider the case where  $\mathcal{V}$  is the unit  $\ell_2$ -ball and  $\mathcal{U}^\circ$  is compact (i.e.,  $\text{dom}(\gamma_{\mathcal{U}}) = \mathbb{R}^n$ ). In this context, only vectors  $u$  need to be sampled. We provide a positive result that if sufficiently many vectors  $u$  are sampled, then we have provable approximations of  $\Theta^\circ$  and hence  $\Theta$  (though with weak dimension-dependent ratios), as well as a negative result showing that in order to obtain an arbitrarily tight bound, an exponential number of samples is needed.

**Upper bounds on  $\Theta^\circ$ .** The following proposition provides an approximation ratio with high probability.

**Proposition 14** Assume  $\mathcal{U}$  and  $\mathcal{U}^\circ$  are compact, and  $\mathcal{V}$  is the unit  $\ell_2$ -ball. Consider  $r$  independent and identically distributed samples  $w_i \in \mathbb{R}^n$  sampled from a standard normal distribution, and  $u_i = w_i/\gamma_{\mathcal{U}}(w_i)$ . If  $r \geq 4n$ , then, with probability greater than  $1 - e^{-r/50}$ , we have, with  $\kappa = 4\sqrt{n} \left( \max_{\Theta(Z) \leq 1} \|Z\|_F \right) \left( \max_{\|u\|_2 \leq 1} \gamma_{\mathcal{U}}(u) \right)$ :

$$\forall Y \in \mathbb{R}^{n \times d}, \Theta^\circ(Y)^2 \geq \max_{i \in \{1, \dots, r\}} u_i^\top Y Y^\top u_i \geq \frac{1}{\kappa^2} \Theta^\circ(Y)^2. \quad (18)$$

**Proof** A sufficient condition for Eq. (18) is that  $\forall Y \in \mathbb{R}^{n \times d}, \max_{i \in \{1, \dots, r\}} u_i^\top Y Y^\top u_i \geq \left( \frac{1}{\kappa^2} \max_{\Theta(Z) \leq 1} \text{tr } Z^\top Z \right) \text{tr } Y^\top Y$ . A sufficient condition is that there exists  $\eta \in \mathbb{R}_+^r$  such that  $1_r^\top \eta = 1$  and  $\lambda_{\min} \left( \sum_{i=1}^r \eta_i u_i u_i^\top \right) \geq \left( \frac{1}{\kappa^2} \max_{\Theta(Z) \leq 1} \text{tr } Z^\top Z \right)$ . Thus, by choosing  $\eta_i \propto \gamma_{\mathcal{U}}(w_i)^2$ , a sufficient condition is, for  $\rho = 4\sqrt{n}$ :

$$\lambda_{\min} \left( \sum_{i=1}^r w_i w_i^\top \right) \geq \left( \frac{1}{\kappa^2} \max_{\Theta(Z) \leq 1} \text{tr } Z^\top Z \right) \left( \max_{\|u\|_2 \leq 1} \gamma_{\mathcal{U}}(u)^2 \right) \sum_{i=1}^r \|w_i\|_2^2 = \frac{1}{\rho^2} \sum_{i=1}^r \|w_i\|_2^2,$$

i.e., with  $W = [w_1, \dots, w_r] \in \mathbb{R}^{n \times r}$ ,  $f(W) = \sqrt{\lambda_{\min}(WW^\top)} - \frac{1}{\rho} \|W\|_F$ . The function  $f$  is Lipschitz-continuous, i.e., such that  $f(W') - f(W) \leq (1 + \frac{1}{4}) \|W - W'\|_F$ . Moreover, using [20, Theorem II.13],  $\mathbb{E}f(W) \geq \sqrt{r} - \sqrt{n} - \frac{1}{\rho} \sqrt{rn} = \sqrt{r} \left( 1 - \sqrt{n/r} - \frac{\sqrt{n}}{\rho} \right) \geq \sqrt{r}/4$ , if  $r \geq 4n$  and  $\rho = 4\sqrt{n}$ . Thus, by concentration of Lipschitz-continuous functions of standard Gaussian variables, we have  $\mathbb{P}(f(W) \leq 0) \leq \exp \left( - \frac{(\sqrt{r}/4)^2}{2(5/4)^2} \right) = \exp(-r/50)$ , hence the result.  $\blacksquare$

Note that the result above is rather weak as on top of the scaling in  $\sqrt{n}$ , the equivalence constants between  $\Theta$  and  $\|\cdot\|_F$  and between  $\gamma_{\mathcal{U}}$  and  $\|\cdot\|_2$  may be large as well (though finite because  $\mathcal{U}^\circ$  is compact).

**Upper bounds on approximation performance.** We now show that even in the simplest case, ( $\text{rank}(Y) = 1$  and both  $\mathcal{U}$  and  $\mathcal{V}$  equal to unit  $\ell_2$ -ball), the approximation ratio has to be dependent on dimensions. Using the same notation as Prop. 14, we have, using Lemma 2 in Appendix B, for any  $(u, v) \in \mathbb{R}^n \times \mathbb{R}^d$ , and  $Y = uv^\top$   $\mathbb{E} \left[ \max_{i \in \{1, \dots, r\}} u_i^\top Y Y^\top u_i \right] \leq \frac{4 \log r + 16}{n} \Theta^\circ(Y)^2$ . Thus, using random sampling may not give good approximation ratios, even in the simplest case, unless the number  $r$  of samples is exponential in  $n$ .

## 6 Obtaining decompositions from convex relaxations

We have presented above semidefinite relaxations of  $\Theta$  and  $\Theta^\circ$ . Much of earlier work [42, 10] and the previous section has been dedicated to obtaining for a certain  $Y \in \mathbb{R}^{n \times d}$ , pairs  $(u, v) \in \mathcal{U} \times \mathcal{V}$  such that  $u^\top Y v \approx \Theta^\circ(Y)$ . In this section, we focus on obtaining *explicit* decomposition of a certain  $X \in \mathbb{R}^{n \times d}$  as  $X = \sum_{m=1}^r u_m v_m^\top$  such that  $\sum_{m=1}^r \gamma_{\mathcal{U}}(u_m) \gamma_{\mathcal{V}}(v_m)$  is equal or better than the value of the relaxation, and thus getting a good approximation of  $\Theta(X)$ . We consider three main approaches.

## 6.1 Singular value decomposition

The first possibility is to follow Section 4.4 and use solutions  $Q, S$  of

$$\sup_{Q \in \mathcal{C}_U^\circ \cap \mathcal{S}_n^+, S \in \mathcal{C}_V^\circ \cap \mathcal{S}_d^+} \|Q^{1/2} X S^{1/2}\|_*$$

and decompose  $Q = AA^\top$  with  $A$  full rank (i.e.,  $A^\top A$  invertible), and  $S = BB^\top$  with  $B$  full rank. Then  $U = (AA^\top)^{-1} A G \text{Diag}(s)^{1/2}$  and  $V = (BB^\top)^{-1} B H \text{Diag}(s)^{1/2}$ , are such that  $X = UV^\top$  and  $\frac{1}{2}\gamma_{\mathcal{C}_U}(UU^\top) + \frac{1}{2}\gamma_{\mathcal{C}_V}(VV^\top)$  is minimal (and equal to the value of the relaxation). However, for any orthogonal matrix  $R$ ,  $(UR, VR)$  is also such a pair, and it is not possible to obtain a decomposition of  $X$  such that the gauge functions  $\gamma_U$  and  $\gamma_V$  of all columns of  $U$  and  $V$  are small.

## 6.2 Conditional gradient algorithms

We may also find decompositions by approximately solving the following convex optimization problem (which is a generalized basis pursuit [18] problem):

$$\min_{Z \in \mathbb{R}^{n \times d}} \frac{1}{2} \|X - Z\|_F^2 + \lambda \Theta(Z), \quad (19)$$

for  $\lambda$  small enough, using convex optimization techniques that only access  $\Theta$  through computing  $\Theta^\circ(Y) = \sup_{(u,v) \in \mathcal{U} \times \mathcal{V}} u^\top Y v$ , and the associated minimizers. This is exactly what generalized conditional gradient algorithms can do [5, 27, 52]. However, we need an algorithm which is robust to obtaining only approximate maximizers  $(u, v)$ , with potentially *multiplicative* approximation guarantees for the computation of  $\Theta^\circ$ .

We consider the following algorithm started from  $Z_0 = 0$ , which iterates the following recursion, for  $\rho_t = 2/(t+1)$ ,  $t \geq 1$ :

- (a)  $(u_{t-1}, v_{t-1}) \in \arg \max_{u \in \mathcal{U}, v \in \mathcal{V}} u^\top (X - Z_{t-1}) v$
- (b)  $\alpha_t = \arg \min_{\alpha \geq 0} \frac{1}{2} \|X - (1 - \rho_t) Z_{t-1} - \rho_t \alpha u_{t-1} v_{t-1}^\top\|_F^2 + \rho_t \lambda \alpha$
- (c)  $Z_t = (1 - \rho_t) Z_{t-1} - \rho_t \alpha_t u_{t-1} v_{t-1}^\top$ .

In Appendix A, we show that if we can find only approximate maximizers  $(u, v)$  with approximation ratio  $\kappa \geq 1$ , then, if  $X_\lambda$  is the unique solution of Eq. (19), then we have

$$\begin{aligned} \frac{1}{2} \|Z_t - X_\lambda\|_F^2 &\leq \frac{1}{2} \|X - Z_t\|_F^2 + \lambda \Theta(Z_t) - \frac{1}{2} \|X - X_\lambda\|_F^2 - \lambda \Theta(X_\lambda) \\ &\leq \frac{2}{(t+1)} \max\{4, \kappa^2\} \Theta(X)^2 \max_{u \in \mathcal{U}} \|u\|_2^2 \max_{v \in \mathcal{V}} \|v\|_2^2 + \lambda(\kappa - 1) \Theta(X), \end{aligned}$$

and  $Z_t$  is a positive linear combination of matrices  $u_{s-1} v_{s-1}^\top$ ,  $s \leq t$ , with a sum of coefficients which is less than  $\Theta(Z_t)$ . The previous inequality implies that  $\Theta(Z_t) \leq \kappa \Theta(X) + O(1/(\lambda t))$ . Thus, when  $\lambda$  is small enough and  $t$  is large enough, we obtain an approximation of  $\Theta(X)$  with approximation ratio which converges to a value less than  $\kappa$ .

**Approaching  $X$  from finite combinations of rank-one factors.** When  $\lambda = 0$  in the algorithm above, then, every  $X \in \mathbb{R}^{n \times d}$  may be approximated up to distance  $\varepsilon$  with a positive linear combination of  $O(1/\sqrt{\varepsilon})$  rank-one factors, even if maximizing  $u^\top Y v$  may only be done approximately for all  $Y \in \mathbb{R}^{n \times d}$ . Moreover, the sum of coefficients is bounded by  $\kappa\Theta(X) + O(\varepsilon)$ .

Finding such decomposition by greedily and iteratively adding factors has been studied thoroughly in signal processing [39] and statistics [8]. In particular, if we assume that  $\Theta$  is a norm, with our set of assumptions, the matching pursuit algorithm of [39] may obtain an  $\varepsilon$ -approximation of  $X$  with  $O(\log \frac{1}{\varepsilon})$  rank-one factors, however, while the norm of the coefficients is bounded, it is not related to the decomposition gauge function  $\Theta(X)$ . In Appendix A.1, we show how optimizing over the scalar  $\rho$  in the algorithm above leads to a similar result, while the sum of coefficients converge to a value which is less than  $\kappa\Theta(X)$ .

### 6.3 Simplicial methods

Conditional gradient algorithms to solve Eq. (19) may be extended by simply replacing step (b), which is the minimization over a half-line, by the minimization with respect to the cone generated by the already obtained rank-one matrices, i.e.,

$$\beta^t = \arg \min_{\beta \in \mathbb{R}^{t-1}} \frac{1}{2} \left\| X - \sum_{s=1}^t \beta_s u_{s-1} v_{s-1}^\top \right\|_F^2 + \lambda \sum_{s=1}^t \beta_s. \quad (20)$$

This approach is sometimes referred to as fully corrective [29] and is an instance of a simplicial method (see, e.g., [12]); typically, it requires much fewer iterations while the cost of each iteration is higher. Note that when the algorithm stops (i.e., there is no further progress in reducing the cost function), then we have solved Eq. (19) up to  $\lambda(\kappa-1)\Omega(x_*)$ . An inbetween alternative is to optimize only over  $\alpha$  and  $\rho$ . Note that the bound derived above also applies to these two extensions, which typically converge much quicker (see examples in Section 7).

Note that when  $(u, v)$  is obtained from randomized rounding, the algorithm is related to what is proposed by [44], which uses *non-adaptive* random sampling for  $u_{s-1} v_{s-1}^\top$ . Moreover, the algorithm may be accelerated by only storing only vectors  $u_1, \dots, u_{t-1}$ , and replacing the subproblem in Eq. (20) by

$$(v_0^t, \dots, v_{t-1}^t) = \arg \min_{(v_0, \dots, v_{t-1}) \in \mathbb{R}^{d \times t}} \frac{1}{2} \left\| X - \sum_{s=1}^t u_{s-1} v_{s-1}^\top \right\|_F^2 + \lambda \sum_{s=1}^t \gamma_{\mathcal{V}}(v_{s-1}). \quad (21)$$

## 7 Simulations

In this section, we provide illustrations of the convex relaxations presented in the paper. We consider  $\mathcal{V}$  the unit  $\ell_2$ -ball and  $\mathcal{U} = \{0, 1\}^d$ . We use  $\mathcal{C}_{\mathcal{U}} = \{M \in \mathcal{S}_n, M \succcurlyeq \text{diag}(M) \text{diag}(M)^\top, M \succcurlyeq 0, \text{diag}(M) \leq 1\}$ , for which we may find an approximation guarantee of  $\sqrt{\pi/2}$  as follows: for  $M \succcurlyeq 0$ , maximizing  $x^\top M x$  with respect to  $x \in \{0, 1\}^n$  may be done by maximizing  $y^\top \begin{pmatrix} M & M \mathbf{1}_n \\ \mathbf{1}_n^\top M & 1 \end{pmatrix} y$  with respect to  $y \in \{-1, 1\}^{n+1}$  such that



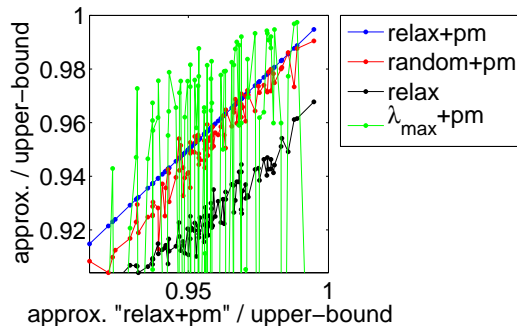


Figure 1: Approximating the polar gauge function  $\Theta^\circ(Y)$  for  $Y$  standard random Gaussian matrix with  $n = 64$  and  $d = 32$ , for 100 samples. For each sample, we compute the relaxed value (upper-bound) based on the semidefinite program and compute several approximations. The samples are ordered so that the performance of the randomized rounding followed by the power method (“relax+pm”) is increasing.

$y_{n+1} = 1$ , which can be done using the usual semi-definite relaxation [42, 25], with an approximation ratio of  $\pi/2$ , and thus the gauge function  $\Theta$  and its polar may be computed with an approximation ratio of  $\sqrt{\pi/2}$ .

**Computation of polar gauge function.** We first compared several approaches to estimating  $\Theta^\circ(Y) = \max_{u \in \mathcal{U}, v \in \mathcal{V}} u^\top Y v = \max_{u \in \{0,1\}^n} \|Y^\top u\|_2$ , for  $Y$  a random matrix with independent and identically distributed components from a normal distribution with mean zero and variance one. We consider several strategies: (a) random sampling of  $u$ , then running the power method to convergence, (b) sampling from the solution of the relaxed semi-definite program, with and without running the power method, and (c) taking the non-randomized approach described in the proof of Prop. 12. In Figure 1, we can see that (a) the performance of the semidefinite-relaxation, even without the power method, is typically much better than the guarantee, (b) that sampling from the relaxed solution and then running the power method outperforms random initializations, and (c) the non-randomized rounding based on eigenvectors has a less stable behavior and sometimes performs better.

In Figure 1, we report averaged value of the randomized rounding procedures. If we take the best values over more than a thousand samples, the obtained values of  $u^\top Y v$  of all three schemes happens to be very close, with a slight advantage to the initializations of the power methods from the convex relaxation.

**Finding decomposition.** We aim to solve the problem in Eq. (19) for  $\lambda = 10^{-4}$ , and consider four approaches: (a) using the semidefinite relaxation to obtain a lower-bound (with no explicit decomposition), (b) using the conditional gradient algorithm, (c) using a simplicial method (the regular version based on Eq. (20) or the one adapted to storing only values of  $u$  in Eq. (21)), (d) using alternating optimization and (e) random sampling (Section 5.2). In Figure 3, we compare these algorithms in two situations, one where the relaxation is tight (left:  $n = 32$ ,  $d = 1$ ) and one where it is not tight (right:  $n = 32$ ,

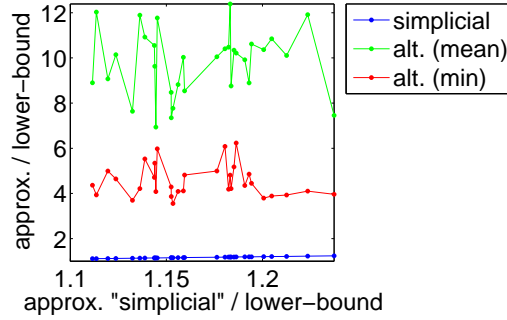


Figure 2: Approximating the gauge function  $\Theta(X)$  for  $X$  standard random Gaussian matrix with  $n = 32$  and  $d = 16$ , for 32 samples. For each sample, we compute the relaxed value (lower-bound) based on the semidefinite program and compute several approximations.

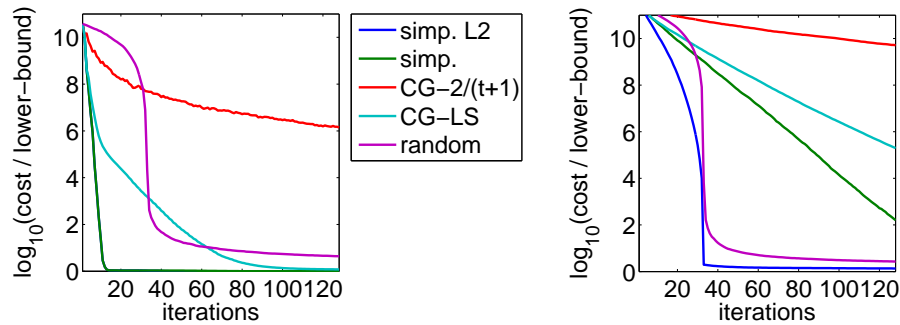


Figure 3: Solving the proximal problem for two random matrices  $Y$ , with  $n = 32$  and  $d = 1$  (left) and  $d = 16$  (right). Comparison of several algorithms. See text for details.

$d = 16$ ). The simplicial algorithms are the faster to converge, with a clear advantage to the one that stores only the vectors  $u$ . The random selection procedure starts slow but eventually catches up, but never reaches the objective function of adaptive methods.

Finally, in Figure 2, we compare the result of using the simplicial method to obtain a decomposition to a simple alternating optimization method. We see that the convex relaxation outperforms significantly the non-convex approach.

## 8 Conclusion

We have presented a general framework for structured matrix decompositions based on gauge functions and semi-definite programming. Emphasis was put on situations where the rank-one factors belong to potentially non-convex and non-centrally symmetric sets. A series of algorithms and relaxations have been presented for a variety of structures.

Our limited experiments have focused on  $\{0, 1\}$ -valued factors. It would be worth studying more precisely recovery guarantees and potentially more scalable algorithms for such cases, in particular for additional constraints such as cardinality or submodularity. Simi-

larly, our general framework leads to alternatives to non-convex algorithms for non-negative matrix factorization, for which it may be possible to show tightness under certain assumptions similar to [3].

Moreover, it would be interesting to see if techniques designed to adaptively reduce the rank for nuclear-norm penalized problems [51] may be extended to our setting as well in order to enforce a smaller number of factors. Finally, the non-convex approaches (power method and alternating minimization) tend to work well in practice, and finding sufficient conditions or random instances (see, e.g., [21] for such work for sparse principal component analysis) where they provably behave well would provide valuable additional insights into the problem of structured matrix factorizations.

## Acknowledgments

This work was supported by the European Research Council (SIERRA project 239993). The author thanks Guillaume Obozinski and Alexandre d’Aspremont for fruitful discussions related to this work.

## A Generalized conditional gradient with approximate oracle

We consider a convex function  $f$  on  $\mathbb{R}^d$ , with  $L$ -Lipschitz continuous gradients, a gauge function  $\gamma_{\mathcal{C}}$  on  $\mathbb{R}^d$ , with  $\mathcal{C}$  a (potentially non convex) closed set that contains the origin, and the following optimization problem  $\min_{x \in \mathbb{R}^d} f(x) + \lambda \gamma_{\mathcal{C}}(x)$ . We assume that (a) the global minimum  $x_*$  is attained, (b) that we know a bound on  $\gamma_{\mathcal{C}}(x_*)$ , i.e.,  $\gamma_{\mathcal{C}}(x_*) \leq \omega$  and (c) we may maximize  $x^\top y$  with respect to  $x \in \mathcal{C}$  approximately, i.e., obtain  $\tilde{x} \in \mathcal{C}$  such that  $\gamma_{\mathcal{C}}^\circ(y) \geq \tilde{x}^\top y \geq \frac{1}{\kappa} \gamma_{\mathcal{C}}^\circ(y)$ , for a fixed  $\kappa \geq 1$ . We consider the following algorithm, started from any  $x_0$  such that  $\gamma_{\mathcal{C}}(x_0) \leq \kappa \omega$ , for  $t \geq 1$ , and  $\rho_t = 2/(t+1)$ :

- (a)  $\tilde{x}_{t-1}$  approximate maximizer of  $-f'(x_{t-1})^\top x$  with respect to  $x \in \mathcal{C}$ ,
  - (b)  $\alpha_t \in \arg \min_{\alpha \in [0, \kappa \omega]} f((1 - \rho_t)x_{t-1} + \rho_t \alpha \tilde{x}_{t-1}) + \lambda \rho_t \alpha$ ,
- or (b')  $\alpha_t \in \arg \min_{\alpha \in [0, \kappa \omega]} -f'(x_{t-1})^\top (x_{t-1} - \alpha \tilde{x}_{t-1}) + \frac{L \rho_t}{2} \|x_{t-1} - \alpha \tilde{x}_{t-1}\|_2^2 + \lambda \alpha$ ,
- (c)  $x_t = (1 - \rho_t)x_{t-1} + \rho_t \alpha_t \tilde{x}_{t-1}$ .

The algorithm generates a sequence of vectors  $\tilde{x}_{t-1} \in \mathcal{C}$  and a vector

$$x_t = \sum_{u=1}^t \left[ \prod_{s=u+1}^t (1 - \rho_s) \right] \rho_u \alpha_u \tilde{x}_{u-1} = \sum_{u=1}^t \frac{2u}{t(t+1)} \alpha_u \tilde{x}_{u-1}.$$

Note that since  $\tilde{x}_{t-1} \in [0, \kappa \omega]$  for all  $t$ , we always have  $\gamma_{\mathcal{C}}(x_t) \leq \kappa \omega$ . We now show that

$$\begin{aligned} f(x_t) + \lambda \gamma_{\mathcal{C}}(x_t) &\leq f(x_t) + \lambda \sum_{u=1}^t \frac{2u}{t(t+1)} \alpha_u \\ &\leq f(x_*) + \lambda \gamma_{\mathcal{C}}(x_*) + \frac{2L\omega^2 \kappa^2 \max_{u \in \mathcal{U}} \|u\|^2}{t+1} + \lambda(\kappa - 1) \gamma_{\mathcal{C}}(x_*). \end{aligned}$$

Let  $g_t = f(x_t) + \lambda \sum_{u=1}^t \frac{2u}{t(t+1)} \alpha_u$ . We have, for any  $\alpha \in [0, \kappa\omega]$ , and  $\tilde{x} \in \text{hull}(\mathcal{C})$ :

$$\begin{aligned}
g_t &= f(x_t) + (1 - \rho_t)[g_{t-1} - f(x_{t-1})] + \lambda\alpha_t\rho_t \\
&= f((1 - \rho_t)x_{t-1} + \rho_t\alpha_t\tilde{x}_{t-1}) + (1 - \rho_t)[g_{t-1} - f(x_{t-1})] + \lambda\alpha_t\rho_t \\
&\leq f(x_{t-1}) - \rho_t f'(x_{t-1})^\top (x_{t-1} - \alpha\tilde{x}_{t-1}) + \frac{L\rho_t^2}{2} \|x_{t-1} - \alpha\tilde{x}_{t-1}\|_2^2 + (1 - \rho_t)[g_{t-1} - f(x_{t-1})] + \lambda\alpha_t\rho_t \\
&\quad \text{by definition of } \alpha_t, \\
&\leq f(x_{t-1}) - \rho_t f'(x_{t-1})^\top (x_{t-1} - \alpha\tilde{x}_{t-1}) + \frac{L\rho_t^2}{2} \kappa^2 \omega^2 \max_{u \in \mathcal{U}} \|u\|^2 + (1 - \rho_t)[g_{t-1} - f(x_{t-1})] + \lambda\alpha_t\rho_t \\
&\leq f(x_{t-1}) - \rho_t f'(x_{t-1})^\top (x_{t-1} - \frac{\alpha}{\kappa}\tilde{x}) + \frac{L\rho_t^2}{2} \kappa^2 \omega^2 \max_{u \in \mathcal{U}} \|u\|^2 + (1 - \rho_t)[g_{t-1} - f(x_{t-1})] + \lambda\alpha_t\rho_t \\
&\quad \text{using the approximate optimality of } \tilde{x}_{t-1}, \\
&\leq f(x_{t-1}) - \rho_t f'(x_{t-1})^\top (x_{t-1} - x_*) + \frac{L\rho_t^2}{2} \kappa^2 \omega^2 \max_{u \in \mathcal{U}} \|u\|^2 + (1 - \rho_t)[g_{t-1} - f(x_{t-1})] + \lambda\kappa\gamma_{\mathcal{C}}(x_*)\rho_t \\
&\quad \text{using } \alpha = \kappa\gamma_{\mathcal{C}}(x_*) \text{ and } \tilde{x} = x_*/\gamma_{\mathcal{C}}(x_*) \in \text{hull}(\mathcal{C}), \\
&\leq f(x_{t-1}) - \rho_t[f(x_{t-1}) - f(x_*)] + \frac{L\rho_t^2}{2} \kappa^2 \omega^2 \max_{u \in \mathcal{U}} \|u\|^2 + (1 - \rho_t)[g_{t-1} - f(x_{t-1})] + \lambda\kappa\gamma_{\mathcal{C}}(x_*)\rho_t \\
&= (1 - \rho_t)g_{t-1} + \rho_t f(x_*) + \frac{L\rho_t^2}{2} \kappa^2 \omega^2 \max_{u \in \mathcal{U}} \|u\|^2 + \lambda\kappa\gamma_{\mathcal{C}}(x_*)\rho_t,
\end{aligned}$$

leading to

$$g_t - f(x_*) - \lambda\gamma_{\mathcal{C}}(x_*) \leq (1 - \rho_t)[g_{t-1} - f(x_*) - \lambda\gamma_{\mathcal{C}}(x_*)] + \frac{L\rho_t^2}{2} \kappa^2 \omega^2 \max_{u \in \mathcal{U}} \|u\|^2 + \lambda(\kappa - 1)\gamma_{\mathcal{C}}(x_*)\rho_t.$$

This implies by recursion (see [5] for details) that

$$g_t - f(x_*) - \lambda\gamma_{\mathcal{C}}(x_*) \leq \frac{2L}{(t+1)} \kappa^2 \omega^2 \max_{u \in \mathcal{U}} \|u\|^2 + \lambda(\kappa - 1)\gamma_{\mathcal{C}}(x_*).$$

If moreover,  $f$  is  $\mu$ -strongly convex and the line search is performed exactly (and without a bound on  $\alpha$ ), then we may show a different bound. Up to the oracle with multiplicative approximation guarantees, the algorithm is then the same as what is proposed by [27, 52], but our results use a slightly different set of assumptions.

First, we show that  $f(x_t)$  remains bounded. Indeed, we have  $f(x_t) \leq f((1 - \rho_t)x_{t-1}) \leq (1 - \rho_t)f(x_{t-1}) + \rho_t f(0)$ , which leads to  $f(x_t) \leq f(0)$  for all  $t \geq 1$ . This implies that  $f(0) \geq f(0) + x_t^\top f'(0) + \frac{\mu}{2} \|x_t\|_2^2 \geq -\|x_t\|_2 \|f'(0)\|_2 + \frac{\mu}{2} \|x_t\|_2^2$ , leading to  $\|x_t\|_2 \leq 2\|f'(0)\|_2/\mu$ .

We may then derive a different recursion, leading to

$$g_t - f(x_*) - \lambda\gamma_{\mathcal{C}}(x_*) \leq \frac{2L}{(t+1)} \max \left\{ \frac{2\|f'(0)\|_2}{\mu}, \kappa\omega \max_{u \in \mathcal{C}} \|u\| \right\}^2 + \lambda(\kappa - 1)\gamma_{\mathcal{C}}(x_*).$$

When applied to  $f(x) = \frac{1}{2}\|x - y\|_2^2$ , we obtain a decaying factor of

$$\frac{2}{(t+1)} \max\{4, \kappa^2\} \gamma_{\mathcal{C}}(y)^2 \max_{u \in \mathcal{C}} \|u\|^2.$$

In this case, our algorithm is strongly related to the relaxed greedy algorithm of [8] and our analysis provides an explicit link between these algorithms and basis pursuit [18].

### A.1 Linearly convergent conditional gradient for strongly convex functions with inexact oracle

We now assume that  $f$  is  $\mu$ -strongly convex and that  $f$  has a global minimum attained at  $x_*$  such that  $\gamma_{\mathcal{C}}(x_*) \leq \omega$ , and that we have a  $\kappa$ -approximate oracle for maximizing linear functions on  $\mathcal{C}$ . We consider the algorithm:

- (a)  $\tilde{x}_{t-1}$  approximate maximizer of  $-f'(x_{t-1})^\top x$  with respect to  $x \in \mathcal{C}$ ,
  - (b)  $(\alpha_t, \rho_t) \in \arg \min_{\alpha \in [0, \kappa\omega + \kappa\varepsilon], \rho \in [0, 1]} f((1 - \rho)x_{t-1} + \rho\alpha\tilde{x}_{t-1})$ ,
- or (b')
- (b')  $(\alpha_t, \rho_t) \in \arg \min_{\alpha \in [0, \kappa\omega + \kappa\varepsilon], \rho \in [0, 1]} -f'(x_{t-1})^\top (x_{t-1} - \alpha\tilde{x}_{t-1}) + \frac{L\rho}{2} \|x_{t-1} - \alpha\tilde{x}_{t-1}\|_2^2$ ,
  - (c)  $x_t = (1 - \rho_t)x_{t-1} + \rho_t\alpha_t\tilde{x}_{t-1}$ .

The algorithm generates a sequence of vectors  $\tilde{x}_{t-1} \in \mathcal{C}$  and a vector

$$x_t = \sum_{u=1}^t \left[ \prod_{s=u+1}^t (1 - \rho_s) \right] \rho_u \alpha_u \tilde{x}_{u-1}$$

such that  $\sum_{u=1}^t \left[ \prod_{s=u+1}^t (1 - \rho_s) \right] \rho_u \alpha_u \leq \kappa\omega + \kappa\varepsilon$ . We have, for any  $\alpha \in [0, \kappa\omega + \kappa\varepsilon]$ ,  $\rho \in [0, 1]$ , and  $\tilde{x} \in \text{hull}(\mathcal{C})$ :

$$\begin{aligned} f(x_t) &\leq f(x_{t-1}) - \rho f'(x_{t-1})^\top (x_{t-1} - \alpha\tilde{x}_{t-1}) + \frac{L\rho^2}{2} \|x_{t-1} - \alpha\tilde{x}_{t-1}\|_2^2 \\ &\quad \text{by definition of } \alpha_t \text{ and } \rho_t, \\ &\leq f(x_{t-1}) - \rho f'(x_{t-1})^\top (x_{t-1} - \alpha\tilde{x}_{t-1}) + \frac{L\rho^2}{2} \kappa^2 (\omega + \varepsilon)^2 \max_{u \in \mathcal{C}} \|u\|^2 \\ &\leq f(x_{t-1}) - \rho f'(x_{t-1})^\top \left( x_{t-1} - \frac{\alpha}{\kappa} \tilde{x} \right) + \frac{L\rho^2}{2} \kappa^2 (\omega + \varepsilon)^2 \max_{u \in \mathcal{C}} \|u\|^2 \\ &\quad \text{using the approximate optimality of } \tilde{x}_{t-1}, \\ &\leq f(x_{t-1}) - \rho f'(x_{t-1})^\top \left( x_{t-1} - x_* + \frac{\varepsilon}{\gamma_{\mathcal{C}}(x_{t-1} - x_*)} (x_{t-1} - x_*) \right) + \frac{L\rho^2}{2} \kappa^2 (\omega + \varepsilon)^2 \max_{u \in \mathcal{C}} \|u\|^2 \\ &\quad \text{using } \alpha = \kappa\gamma_{\mathcal{C}}(\tilde{x}) \text{ and } \tilde{x} \propto x_* - \frac{\varepsilon}{\gamma_{\mathcal{C}}(x_{t-1} - x_*)} (x_{t-1} - x_*), \\ &\leq f(x_{t-1}) - \rho [f(x_{t-1}) - f(x_*)] \left( 1 + \frac{\varepsilon}{\gamma_{\mathcal{C}}(x_{t-1} - x_*)} \right) + \frac{L\rho^2}{2} \kappa^2 \omega^2 \max_{u \in \mathcal{C}} \|u\|^2. \end{aligned}$$

Moreover, we have  $f(x_{t-1}) - f(x_*) \geq \frac{\mu}{2} \|x_{t-1} - x_*\|_2^2 \geq \frac{\mu}{2} \frac{\gamma_{\mathcal{C}}(x_{t-1} - x_*)^2}{\max_{\|u\|_2=1} \gamma_{\mathcal{C}}(u)^2}$ , leading to, with  $\Delta_t = f(x_{t-1}) - f(x_*)$ ,

$$\Delta_t \leq \Delta_{t-1} - \rho \left[ \Delta_{t-1} + \frac{\varepsilon \sqrt{\mu} \Delta_{t-1}^{1/2}}{\sqrt{2} \max_{\|u\|_2=1} \gamma_{\mathcal{C}}(u)} \right] + \frac{L\rho^2}{2} \kappa^2 \omega^2 \max_{u \in \mathcal{C}} \|u\|^2.$$

If  $\left[ \Delta_{t-1} + \frac{\varepsilon\sqrt{\mu}\Delta_{t-1}^{1/2}}{\sqrt{2} \max_{\|u\|_2=1} \gamma_{\mathcal{C}}(u)} \right] \frac{1}{L\kappa^2\omega^2 \max_{u \in \mathcal{C}} \|u\|^2} < 1$ , we have a minimizer  $\rho \in [0, 1)$ , and

$$\begin{aligned} \Delta_t &\leq \Delta_{t-1} - \frac{1}{2} \left[ \frac{\varepsilon\sqrt{\mu}\Delta_{t-1}^{1/2}}{\sqrt{2} \max_{\|u\|_2=1} \gamma_{\mathcal{C}}(u)} \right]^2 \frac{1}{L\kappa^2\omega^2 \max_{u \in \mathcal{C}} \|u\|^2} \\ &\leq \Delta_{t-1} \left( 1 - \frac{\varepsilon^2\mu}{4L\kappa^2\omega^2 \max_{\|u\|_2=1} \gamma_{\mathcal{C}}(u)^2 \times \max_{u \in \mathcal{C}} \|u\|^2} \right). \end{aligned}$$

Otherwise, we have

$$\Delta_t \leq \Delta_{t-1} - \frac{1}{2} \left[ \Delta_{t-1} + \frac{\varepsilon\sqrt{\mu}\Delta_{t-1}^{1/2}}{\sqrt{2} \max_{\|u\|_2=1} \Omega(u)} \right] \leq \frac{1}{2} \Delta_{t-1}$$

Thus, with  $\tau = \min \left\{ \frac{1}{2}, \frac{\varepsilon^2}{4\kappa^2\omega^2} \frac{\mu}{L \max_{\|u\|_2=1} \gamma_{\mathcal{C}}(u) \times \max_{u \in \mathcal{C}} \|u\|^2} \right\}$ , we have  $\Delta_t \leq (1 - \tau)\Delta_{t-1}$ , and hence a linear convergence rate.

## B Maximum of beta random variables

Given  $u \in \mathbb{R}^n$  such that  $\|u\|_2 = 1$ , our goal is to upper-bound  $\mathbb{E} \left[ \max_{i \in \{1, \dots, r\}} \frac{(w_i^\top u)^2}{w_i^\top w_i} \right]$ , for  $w_i$  sampled i.i.d from a standard normal distribution. Using the representation of Beta random variables as ratios of independent Gamma variables, each  $\frac{(w_i^\top u)^2}{w_i^\top w_i}$  is Beta-distributed with parameters  $(\frac{1}{2}, \frac{n-1}{2})$ . The following lemma provides a bound on the expectation of maxima of independent Beta variables.

**Lemma 2** *Let  $X_i, i = 1, \dots, r$ , be  $r$  i.i.d. Beta random variables with parameter  $(\frac{1}{2}, \frac{n-1}{2})$ ,  $n \geq 2$ . Then  $\mathbb{E} \left[ \max_{i \in \{1, \dots, r\}} X_i \right] \leq \frac{4 \log r + 16}{n}$ .*

**Proof** We have  $\mathbb{E}X^k = \prod_{j=0}^{k-1} \frac{1+2j}{n+2j}$ . For all  $k > 0$ , we have  $\mathbb{E}X^k \leq \prod_{j=0}^{k-1} \frac{2+2j}{n} = \frac{2^k k!}{n^k}$ , while for  $k \geq \lfloor n/2 \rfloor$ , we have  $\mathbb{E}X^k \leq \prod_{j=0}^{\lfloor n/2 \rfloor - 1} \frac{1+2j}{n+2j} \leq 2^{-\lfloor n/2 \rfloor}$ . This leads to, for  $t \leq n/2$ ,

$$\mathbb{E}e^{tX} \leq \sum_{k=0}^{\lfloor n/2 \rfloor} \left( \frac{2t}{n} \right)^k + 2^{-\lfloor n/2 \rfloor} \sum_{k=\lfloor n/2 \rfloor + 1}^{\infty} \frac{t^k}{k!} \leq \frac{1}{1 - 2t/n} + 2^{1/2 - n/2} e^t.$$

Using standard results from probability (see, e.g., [14]), we get, with  $t = n/4$ :

$$\mathbb{E} \left[ \max_{i \in \{1, \dots, r\}} X_i \right] \leq \frac{\log r + \mathbb{E}e^{tX}}{t} \leq \frac{\log r + 2 + 2^{1/2 - n/2} e^{n/4}}{n/4} \leq \frac{4 \log r + 16}{n}.$$

■

## References

- [1] M. Aharon, M. Elad, and A. M. Bruckstein. The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representations. *IEEE Transactions on Signal Processing*, 54(11):4311–4322, 2006.
- [2] N. Alon and A. Naor. Approximating the cut-norm via Grothendieck’s inequality. *SIAM Journal on Computing*, 35(4):787–803, 2006.
- [3] S. Arora, S. Ge, R. Kannan, and A. Moitra. Computing a nonnegative matrix factorization—provably. In *Proceedings of the Symposium on Theory of Computing (STOC)*, pages 145–162, 2012.
- [4] F. Bach. Learning with submodular functions: A convex optimization perspective. *Arxiv preprint arXiv:1111.6453*, 2011.
- [5] F. Bach. Duality between subgradient and conditional gradient methods. Technical Report 00757696, HAL, 2013.
- [6] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski. Optimization with sparsity-inducing penalties. *Foundations and Trends® in Machine Learning*, 4(1):1–106, 2011.
- [7] F. Bach, J. Mairal, and J. Ponce. Convex sparse matrix factorizations. Technical Report 00345747, HAL, 2008.
- [8] A. R. Barron, A. Cohen, W. Dahmen, and R. A. DeVore. Approximation and learning by greedy algorithms. *The annals of Statistics*, 36(1):64–94, 2008.
- [9] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton University Press, 2009.
- [10] A. Ben-Tal and A. Nemirovski. On approximating matrix norms. Technical report, Technion, 2011.
- [11] A. Berman and N. Shaked-Monderer. *Completely positive matrices*. World Scientific, 2003.
- [12] D. P. Bertsekas and H. Yu. A unifying polyhedral approximation framework for convex optimization. *SIAM Journal on Optimization*, 21(1):333–360, 2011.
- [13] J. M. Borwein and A. S. Lewis. *Convex Analysis and Nonlinear Optimization: Theory and Examples*. Springer, 2006.
- [14] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.
- [15] D. W. Boyd. The power method for  $\ell_p$ -norms. *Linear Algebra and its Applications*, 9:95–101, 1974.

- [16] S. Burer and R. D. C. Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, 95(2):329–357, 2003.
- [17] V. Chandrasekaran, B. Recht, P. A., and A. S. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational Mathematics*, 12(6):805–849, 2012.
- [18] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20:33–61, 1999.
- [19] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk. Iteratively reweighted least squares minimization for sparse recovery. *Communications on Pure and Applied Mathematics*, 63(1):1–38, 2010.
- [20] K. R. Davidson and S. J. Szarek. Local operator theory, random matrices and Banach spaces. *Handbook of the geometry of Banach spaces*, 1:317–366, 2001.
- [21] Y. Deshpande and A. Montanari. Finding hidden cliques of size  $\sqrt{N/e}$  in nearly linear time. Technical Report 1304.7047, arXiv, 2013.
- [22] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern classification*. John Wiley & Sons, 2012.
- [23] C. Févotte, N. Bertin, and J.-L. Durrieu. Nonnegative matrix factorization with the itakura-saito divergence: With application to music analysis. *Neural computation*, 21(3):793–830, 2009.
- [24] S. Fujishige. *Submodular Functions and Optimization*. Elsevier, 2005.
- [25] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.
- [26] G. H. Golub and C. F. Van Loan. *Matrix computations*. John Hopkins University Press, 1996.
- [27] Z. Harchaoui, A. Juditsky, and A. Nemirovski. Conditional gradient algorithms for norm-regularized smooth convex optimization. Technical Report 1302.2325, arXiv, 2013.
- [28] A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634, 1999.
- [29] M. Jaggi. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2013.
- [30] G. J. O. Jameson. *Summing and nuclear norms in Banach space theory*. Cambridge University Press, 1987.



- [31] M. Journée, F. Bach, P.-A. Absil, and R. Sepulchre. Low-rank optimization on the cone of positive semidefinite matrices. *SIAM Journal on Optimization*, 20(5):2327–2351, 2010.
- [32] M. Journée, Y. Nesterov, P. Richtárik, and R. Sepulchre. Generalized power method for sparse principal component analysis. *The Journal of Machine Learning Research*, 11:517–553, 2010.
- [33] R. Krauthgamer, B. Nadler, and D. Vilenchik. Do semidefinite relaxations really solve sparse PCA? Technical Report 1306.3690, ArXiv, 2013.
- [34] D. D. Lee and S. H. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.
- [35] J. Lee, B. Recht, N. Srebro, J. Tropp, and R. Salakhutdinov. Practical large-scale optimization for max-norm regularization. In *Advances in Neural Information Processing Systems (NIPS)*, 2010.
- [36] A. S. Lewis. The mathematics of eigenvalue optimization. *Mathematical Programming*, 97(1-2):155–176, 2003.
- [37] N. Linial, S. Mendelson, G. Schechtman, and A. Shraibman. Complexity measures of sign matrices. *Combinatorica*, 27(4):439–463, 2007.
- [38] J. Mairal, F. Bach, J. Ponce, and G. Sapiro. Online learning for matrix factorization and sparse coding. *Journal of Machine Learning Research*, 11(1):19–60, 2010.
- [39] S. G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, 1993.
- [40] C. A. Micchelli, J. M. Morales, and M. Pontil. Regularizers for structured sparsity. *Advances in Computational Mathematics*, 38(3):455–489, 2013.
- [41] K. P. Murphy. *Machine learning: a probabilistic perspective*. The MIT Press, 2012.
- [42] Y. Nesterov. Semidefinite relaxation and nonconvex quadratic optimization. *Optimization Methods and Software*, 9(1-3):141–160, 1998.
- [43] G. Obozinski and F. Bach. Convex relaxation of combinatorial penalties. Technical Report 00694765, HAL, 2012.
- [44] G. Obozinski, B. Taskar, and M. I. Jordan. Joint covariate selection and joint subspace selection for multiple classification problems. *Statistics and Computing*, 20(2):231–252, 2009.
- [45] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review*, 52(3):471–501, 2010.
- [46] B. Recht, C. Re, J. Tropp, and V. Bittorf. Factoring nonnegative matrices with linear programs. In *Advances in Neural Information Processing Systems (NIPS)*, 2012.

- [47] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1997.
- [48] N. Srebro and A. Shraibman. Rank, trace-norm and max-norm. In *Learning Theory*, pages 545–560. Springer, 2005.
- [49] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Academic Press, 1990.
- [50] H. Xu, C. Caramanis, and S. Sanghavi. Robust PCA via outlier pursuit. *IEEE Transactions on Information Theory*, 58(5):3047–3064, 2012.
- [51] T. Zhang. Analysis of multi-stage convex relaxation for sparse regularization. *The Journal of Machine Learning Research*, 11:1081–1107, 2010.
- [52] X. Zhang, D. Schuurmans, and Y. Yu. Accelerated training for matrix-norm regularization: A boosting approach. In *Advances in Neural Information Processing Systems (NIPS)*, 2012.