Low-Rank Inducing Norms with Optimality Interpretations

Christian Grussler * Pontus Giselsson *

December 9, 2016

Abstract

Optimization problems with rank constraints appear in many diverse fields such as control, machine learning and image analysis. Since the rank constraint is non-convex, these problems are often approximately solved via convex relaxations. Nuclear norm regularization is the prevailing convexifying technique for dealing with these types of problem. This paper introduces a family of low-rank inducing norms and regularizers which includes the nuclear norm as a special case. A posteriori guarantees on solving an underlying rank constrained optimization problem with these convex relaxations are provided. We evaluate the performance of the low-rank inducing norms on three matrix completion problems. In all examples, the nuclear norm heuristic is outperformed by convex relaxations based on other low-rank inducing norms. For two of the problems there exist low-rank inducing norms that succeed in recovering the partially unknown matrix, while the nuclear norm fails. These low-rank inducing norms are shown to be representable as semi-definite programs and to have cheaply computable proximal mappings. The latter makes it possible to also solve problems of large size with the help of scalable first-order methods. Finally, it is proven that our findings extend to the more general class of atomic norms. In particular, this allows us to solve corresponding vector-valued problems, as well as problems with other non-convex constraints.

1 Introduction

Many problems in machine learning, image analysis, model order reduction, multivariate linear regression, etc. (see [1, 6, 7, 10, 25, 30, 31, 39, 40, 45]), can be posed as a low-rank estimation problems based on measurements and prior information about a data matrix. These estimation problems often take the form

$$\begin{aligned} \underset{M}{\text{minimize}} \quad & f_0(M) \\ \text{subject to} \quad & \operatorname{rank}(M) \leq r, \end{aligned} \tag{1}$$

*Department of Automatic Control, Lund University, Sweden {christiang,pontusg}@control.lth.se.

where f_0 is a proper closed convex function and r is a positive integer that specifies the desired or expected rank. Due to non-convexity of the rank constraint a solution to (1) is known only in a few special cases (see e.g. [1, 2, 40]).

A common approach to deal with the rank constraint is to use the nuclear norm heuristic (see [18, 39]). The idea is to convexify the problem by replacing the non-convex rank constraint with a nuclear norm regularization term. For matrix completion problems, this approach is shown to recover the true low-rank matrix with high probability, provided that enough random measurements are available (see [7, 9, 39]). If these assumption are not met, however, the nuclear norm heuristic may fail in producing satisfactory estimates (see [23, 24]).

This paper introduces a family of low-rank inducing norms as alternatives to the nuclear norm. These norms can be interpreted as the largest convex minorizers of non-convex functions f of the form

$$f := \|\cdot\| + \chi_{\text{rank}(\cdot)\leq r}, \tag{2}$$

where $\|\cdot\|$ is an arbitrary unitarily invariant norm, and $\chi_{\text{rank}(\cdot)\leq r}$ is the indicator function for matrices with rank less than or equal to r. This interpretation motivates the use of low-rank inducing norms in convex relaxations to (1). In particular, assume that f_0 in (1) can be split into the sum of a convex function and unitarily invariant norm, and the solution to the corresponding convex relaxation has rank r. Then this solution also solves the non-convex problem, and thus provides an *a posteriori* optimality guarantee. Furthermore, the choice of norms and target ranks r can be considered as regularization parameters when used in convex relaxations of (1). Compared to the nuclear norm approach, it is shown that this gives additional flexibility which can be exploited to improve the quality of the estimate. Specifically, the nuclear norm is the largest convex minorizer of f in (2) with $r = 1$, making it a less natural choice than other low-rank inducing norms, because it convexifies constraints that allow for matrices of rank 1, only.

This work particularly focuses on low-rank inducing norms, where the norm in (2) is the Frobenius norm or the spectral norm. We refer to these norms as *low-rank inducing Frobenius norms* and *low-rank inducing spectral norms*, respectively. The low-rank inducing Frobenius norms, also called $r*$ norms, have been previously discussed in the literature (see [4, 14, 16, 22–24, 36]). In [4, 14, 16, 36], no optimality interpretations are considered, but in previous work we have presented such interpretations for the squared $r*$ norms (see [22–24]). In this paper these findings are shown to extend to any function of low-rank inducing norms that is increasing on the nonnegative real numbers. Most importantly, our results hold for linear increasing functions, i.e. the low-rank inducing norm itself. To the best of our knowledge, no other low-rank inducing norms from the proposed family, including low-rank inducing spectral norms, have been proposed in the literature.

For the family of low-rank inducing norms to be useful in practice, they must be suitable for numerical optimization. We show that low-rank inducing Frobenius norms and spectral norms are representable as semi-definite programs (SDP). This allows us to readily formulate and solve small to medium scale problems using standard SDP-solvers (see [38, 44]). Moreover, it is demonstrated that these norms have cheaply computable proximal mappings, comparable with the computational cost for the proximal

mapping of the nuclear norm. This allows us to solve large-scale problems involving low-rank inducing norms by means of proximal splitting methods (see [12, 37]). To enable formulations with increasing convex functions, the projection onto their epigraphs is computed. This extends the proximal mapping computations of the squared $r*$ norm in [3, 16, 23] to the non-squared case.

The performance of different low-rank inducing norms is evaluated on three matrix completion problems. The evaluation reveals that the choice of low-rank inducing norms has tremendous impact on the ability to complete the covariance matrix. In particular, the nuclear norm is significantly outperformed by the low-rank inducing Frobenius norm, as well as the low-rank inducing spectral norm.

The findings in this work are also valid for the corresponding vector-valued problems by replacing rank with cardinality. This gives rise to optimality interpretations of, e.g., lasso-type and inverse problems (see [25, 42, 45]). More generally, all low-rank inducing norms lie within the class of so-called atomic norms (see [9]). It is shown that our optimality interpretations also hold for atomic norms under very mild assumptions. Therefore, these findings provide optimality interpretations for many other problems, such as those listed in [9, Section 2.2].

The paper is organized as follows. We start by introducing some preliminaries in Section 2. In Section 3, we introduce the class of low-rank inducing norms, and provide optimality interpretations of these in Section 4. In Section 5, computability of low-rank inducing Frobenius and spectral norms is addressed. To support the usefulness of having more low-rank inducing regularizers at our supply, numerical examples are presented in Section 6. The optimality results are extended to the vector case and to atomic norms in Section 7 and conclusions are drawn in Section 8.

2 Preliminaries

The set of reals is denoted by $\mathbb{R}$, the set of real vectors by $\mathbb{R}^n$, and the set of real matrices by $\mathbb{R}^{n\times m}$. Element-wise nonnegative matrices $X \in \mathbb{R}^{n\times m}$ are denoted by $X \in \mathbb{R}^{n\times m}_{\geq 0}$. If symmetric $X \in \mathbb{R}^{n\times n}$ is positive definite (semi-definite), we write $X \succ 0$ ($X \succeq 0$). These notations are also used to describe relations between matrices, e.g., $A \succeq B$ means $A - B \succeq 0$. The non-increasingly ordered singular values of $X \in \mathbb{R}^{n\times m}$, counted with multiplicity, are denoted by $\sigma_1(X) \geq \cdots \geq \sigma_{\min\{m,n\}}(X)$. Furthermore,

$$\langle X, Y \rangle := \sum_{i=1}^{m} \sum_{j=n}^{n} x_{ij} y_{ij} = \operatorname{trace}(X^T Y)$$

defines the Frobennius inner-product for $X, Y \in \mathbb{R}^{n\times m}$. This inner-product gives the *Frobenius norm*

$$\|X\|_F := \sqrt{\operatorname{trace}(X^T X)} = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij}^2} = \sqrt{\sum_{i=1}^{\min\{m,n\}} \sigma_i^2(X)},$$

which is a unitarily invariant norm, i.e., $\|UXV\|_F = \|X\|_F$ for all unitary matrices $U, V \in \mathbb{R}^{n\times m}$. For all $x = (x_1, \ldots, x_q) \in \mathbb{R}^q$, we define

$$\ell_1(x) := \sum_{i=1}^{q} |x_i|, \qquad \ell_2(x) := \sqrt{\sum_{i=1}^{q} x_i^2}, \qquad \ell_\infty(x) := \max_i |x_i|, \tag{3}$$

Then the Frobenius norm satisfies $\|X\|_F = \ell_2(\sigma(X))$, where

$$\sigma(X) := (\sigma_1(X), \ldots, \sigma_q(X)).$$

The functions ℓ_1 and ℓ_∞ define the nuclear norm $\|X\|_{\ell_1} := \ell_1(\sigma(X))$ and the spectral norm $\|X\|_{\ell_\infty} := \ell_\infty(\sigma(X)) = \sigma_1(X)$.

For a set $\mathcal{C} \subset \mathbb{R}^{n\times m}$,

$$\chi_{\mathcal{C}}(X) := \begin{cases} 0, & X \in \mathcal{C} \\ \infty, & X \notin \mathcal{C} \end{cases}$$

denotes the so-called *indicator function*. We also use $\chi_{\text{rank}(\cdot)\le r}$ to denote the indicator function of the set of matrices which have at most rank r.

The following function properties will be used in this paper. The *effective domain* of a function $f : \mathbb{R}^{n\times m} \to \mathbb{R} \cup \{\infty\}$ is defined as

$$\text{dom} f := \{X \in \mathbb{R}^{n\times m} : f(X) < \infty\}$$

and the *epigraph* is defined as

$$\text{epi}(f) := \{(X,t) : f(X) \le t, X \in \text{dom} f,\ t \in \mathbb{R}\}.$$

Further, f is said to be:

- *proper* if $\text{dom} f \neq \emptyset$.
- *closed* if the epigraph is a closed set.
- *positively homogeneous (of degree 1)* if for all $X \in \text{dom}(f)$ and $t > 0$ it holds that $f(tX) = tf(X)$.
- *nonnegative* if $f(X) \ge 0$ for all $X \in \text{dom}(f)$.
- *coercive* if $\lim_{\|X\|_F \to \infty} f(X) = \infty$.

A function $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ is called *increasing* if

$$x \le y \ \Rightarrow\ f(x) \le f(y) \ \text{ for all } \ x, y \in \text{dom}(f)$$

and if there exist $x, y \in \mathbb{R}$ such that $x < y$ and $f(x) < f(y)$.

The *conjugate (dual) function* f^* of f is defined as

$$f^*(Y) := \sup_{X \in \mathbb{R}^{n\times m}} [\langle X, Y\rangle - f(X)]$$

for all $Y \in \mathbb{R}^{n\times m}$. As long as f is proper and minorized by an affine function, the conjugate f^* is proper, closed and convex (see [28]). The function $f^{**} := (f^*)^*$ is called the *biconjugate function* of f and can be shown to be a convex minorizer of f, i.e.

$$f(X) \geq f^{**}(X) \text{ for all } X \in \mathbb{R}^{n\times m}.$$

In fact, f^{**} is the point-wise supremum of all affine functions majorized by f and therefore the largest convex minorizer of f. This can equivalently be stated as follows (see [27, Theorem X.1.3.5,Corollary X.1.3.6]).

Lemma 1 *Let $f : \mathbb{R}^{n\times m} \to \mathbb{R} \cup \{\infty\}$ be such that f^{**} is proper. Then*

$$\operatorname{epi}(f^{**}) = \operatorname{cl}(\operatorname{conv}(\operatorname{epi} f)),$$

where $\operatorname{cl}(\cdot)$ *denotes the topological closure of a set and* $\operatorname{conv}(\cdot)$ *the convex hull. Further, $f^{**} = f$ if and only if f is proper closed and convex.*

Lemma 1 implies that for a closed proper, but possibly non-convex function f, it holds that

$$\inf_{X\in\mathbb{R}^{n\times m}} f(X) = \inf_{X\in\mathbb{R}^{n\times m}} f^{**}(X).$$

However, determining the convex function f^{**} is as difficult as minimizing the non-convex function f. Instead, it is common to convexify the problem by splitting the function into $f = f_1 + f_2$, such that f_1^{**} and f_2^{**} can be easily computed. If f_1 is proper, closed and convex, then $f_1 = f_1^{**}$ and $f_1 + f_2^{**}$ is the largest convex minorizer of f that keeps f_1 as a summand. In particular,

$$\inf_{X\in\mathbb{R}^{n\times m}} [f_1(X) + f_2(X)] \geq \inf_{X\in\mathbb{R}^{n\times m}} [f_1(X) + f_2^{**}(X)], \tag{4}$$

which holds with equality if the solution $X^\star$ to the right-hand side problem satisfies $f_2^{**}(X^\star) = f_2(X^\star)$. Then $X^\star$ also solves the non-convex problem on the left-hand side. This motivates the use of our terminology that $f_1 + f_2^{**}$ is the *optimal convex relaxation* of a given splitting $f_1 + f_2$, when f_1 is proper closed and convex.

Finally, if $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$, then the *monotone conjugate* is defined as

$$f^+(y) := \sup_{x\geq 0} [\langle x, y\rangle - f(x)] \quad \text{for all } y \in \mathbb{R}.$$

3 Low-Rank Inducing Norms

This section introduces the family of *low-rank inducing norms*, which includes the nuclear norm as a special case. These can be used as regularizers in optimization problems to promote low-rank solutions. To define them, we need to characterize the class of unitarily invariant norms in terms of symmetric gauge functions. This characterization can be found in, e.g. [29, Theorem 7.4.7.2].

Definition 1 *A function $g : \mathbb{R}^q \to \mathbb{R}_{\geq 0}$ is a symmetric gauge function if*

i. *g is a norm.*

ii. *$\forall x \in \mathbb{R}^q : g(|x|) = g(x)$, where $|x|$ denotes the element-wise absolute value.*

iii. *$g(Px) = g(x)$ for all permutation matrices $P \in \mathbb{R}^{q\times q}$ and all $x \in \mathbb{R}^q$.*

Proposition 1 *The norm $\|\cdot\| : \mathbb{R}^{n\times m} \to \mathbb{R}$ is unitarily invariant if and only if*

$$\|X\| = g(\sigma_1(X), \ldots, \sigma_{\min\{m,n\}}(X))$$

for all $X \in \mathbb{R}^{n\times m}$, where g is a symmetric gauge function.

As noted in Section 2, the gauge functions for the Frobenius norm, spectral norm, and nuclear norm are $g = \ell_2$, $g = \ell_\infty$, and $g = \ell_1$, respectively, where ℓ_1, ℓ_2, and ℓ_∞, are defined in (3).

The dual norm of a unitarily invariant norm is also unitarily invariant (see [29, Theorem 5.6.39]. Therefore, it has an associated symmetric gauge function. This will be denoted by g^D if the symmetric gauge function of the original norm is denoted by g. More specifically, let $M \in \mathbb{R}^{n\times m}$, $q := \min\{m,n\}$, and $g : \mathbb{R}^q \to \mathbb{R}_{\geq 0}$ be a symmetric gauge function associated with a unitarily invariant norm

$$\|M\|_g := g(\sigma_1(M), \ldots, \sigma_q(M)).$$

The dual of this norm is defined as

$$\|Y\|_{g^D} := \max_{\|M\|_g \leq 1} \langle Y, M\rangle = g^D(\sigma_1(Y), \ldots, \sigma_q(Y)), \tag{5}$$

where the dual gauge function g^D satisfies

$$g^D(\sigma_1(Y), \ldots, \sigma_q(Y)) = \max_{g(\sigma_1(M),\ldots,\sigma_q(M))\leq 1} \sum_{i=1}^{q} \sigma_i(M)\sigma_i(Y). \tag{6}$$

The low-rank inducing norms will be defined as the dual norm of a rank constrained dual norm in (5). This rank constrained dual norm is defined as

$$\|Y\|_{g^D,r} := \max_{\substack{\text{rank}(M)\leq r \\ \|M\|_g \leq 1}} \langle M, Y\rangle \tag{7}$$

and the corresponding *low-rank inducing norm* as

$$\|M\|_{g,r*} := \max_{\|Y\|_{g^D,r}\leq 1} \langle Y, M\rangle. \tag{8}$$

For $q = \min\{m,n\}$, the rank constraint in (7) is redundant and the dual of the dual becomes the norm itself.

For symmetric gauge functions $g : \mathbb{R}^q \to \mathbb{R}_{\geq 0}$, we denote their truncated symmetric gauge functions by $g(\sigma_1, \ldots, \sigma_r) := g(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0)$ for any $r \in \{1, \ldots, q\}$. With this notation in mind, some properties of low-rank inducing norms and their duals are stated in the following lemma. A proof is given in Appendix A.1.1.

Lemma 2 *Let $M, Y \in \mathbb{R}^{n\times m}$, $r \in \mathbb{N}$ be such that $1 \le r \le q := \min\{m,n\}$, and $g : \mathbb{R}^q \to \mathbb{R}_{\ge 0}$ be a symmetric gauge function. Then $\|\cdot\|_{g^D,r}$ is a unitarily invariant norm that satisfies*

$$\|Y\|_{g^D,r} = g^D(\sigma_1(Y),\ldots,\sigma_r(Y)) \tag{9}$$

Its dual norm $\|\cdot\|_{g,r*}$ satisfies

$$\|M\|_{g,r*} = \max_{g^D(\sigma_1(Y),\ldots,\sigma_r(Y))\le 1} \left[\sum_{i=1}^{r}\sigma_i(M)\sigma_i(Y) + \sigma_r(Y)\sum_{i=r+1}^{q}\sigma_i(M)\right], \tag{10}$$

and

$$\|M\|_g = \|M\|_{g,q*} \le \cdots \le \|M\|_{g,1*}, \tag{11}$$

$$\operatorname{rank}(M) \le r \;\Rightarrow\; \|M\|_g = \|M\|_{g,r*}. \tag{12}$$

This paper particularly focuses on low-rank inducing norms originating from the Frobenius norm and the spectral norm. When the original norm is the Frobenius norm, then $g = \ell_2$. Since the norm is self dual, it satisfies $g^D = \ell_2^D = \ell_2$. The truncated version in (9) (which is denote by $\|\cdot\|_r$ to comply with notation used, e.g., in [23]) becomes

$$\|Y\|_r := \|Y\|_{\ell_2^D,r} = \sqrt{\sum_{i=1}^{r}\sigma_i^2(Y)}.$$

The corresponding low-rank inducing norm is referred to as the *low-rank inducing Frobenius norm*, and is denoted by

$$\|M\|_{r*} := \|M\|_{\ell_2,r*} = \max_{\|Y\|_r\le 1}\langle Y, M\rangle.$$

In [23], this norm is referred to as the $r*$ norm.

If the original norm, instead, is the spectral norm, we have $g = \ell_\infty$. The dual norm is the nuclear (trace) norm (see [29, Theorem 5.6.42]), with gauge function $g^D = \ell_1$. The truncated version becomes

$$\|Y\|_{\ell_1,r} := \sum_{i=1}^{r}\sigma_i(Y),$$

and its dual, which we refer to as the *low-rank inducing spectral norm*, is denoted by

$$\|M\|_{\ell_\infty,r*} := \max_{\|Y\|_{\ell_1,r}\le 1}\langle Y, M\rangle.$$

The nuclear norm is a special case of these low-rank inducing norms, corresponding to $r = 1$.

Proposition 2 *The nuclear norm satisfies $\|\cdot\|_{\ell_1} = \|\cdot\|_{g,1*}$, where $\|\cdot\|_g$ is any unitarily invariant norm with $g(\sigma_1) = |\sigma_1|$.*

A proof to this proposition is found in Appendix A.1.2.

Next, we state a result that is the key to our optimality interpretations for low-rank inducing norms in the next section.

Lemma 3 *Let* $B^1_{g,r*} := \{X \in \mathbb{R}^{n\times m} : \|X\|_{g,r*} \leq 1\}$ *be the unit low-rank inducing norm ball and let*

$$E_{g,r} := \{X \in \mathbb{R}^{n\times m} : \|X\|_g = 1,\ \text{rank}(X) \leq r\}. \tag{13}$$

Then $B^1_{g,r*} = \text{conv}(E_{g,r})$*, i.e. all* $M \in \mathbb{R}^{n\times m}$ *can be decomposed as*

$$M = \textstyle\sum_i \alpha_i M_i \quad \textit{with} \quad \sum_i \alpha_i = 1,\ \alpha_i \geq 0,$$

where M_i *satisfies* $\text{rank}(M_i) \leq r$ *and*

$$\|M_i\|_g = \|M_i\|_{g,r*} = \|M\|_{g,r*}.$$

A proof to this lemma is given in Appendix A.1.3. The result is a direct consequence of Lemma 2 and extends what is known about the nuclear norm, and the results on low-rank inducing Frobenius norms in [23].

In many cases, the set $E_{g,r}$ is the set of extreme points to the unit ball $B^1_{g,r*}$. The following result is proven in Appendix A.1.4.

Proposition 3 *Suppose that* $\|\cdot\|_g$ *satisfies*

$$\|\textstyle\sum_i \alpha_i M_i\|_g < \sum_i \alpha_i \|M_i\|_g$$

for all $\alpha_i \in (0,1)$ *such that* $\sum_i \alpha_i = 1$*, and all* $M_i \in \mathbb{R}^{n\times m}$ *with* $\|M_i\|_g = 1$*. Then* $E_{g,r}$ *in* (13) *is the set of extreme points to* $B^1_{g,r*}$*.*

All ℓ_p norms with $1 < p < \infty$ satisfy these assumptions, and therefore the unit balls of their low-rank inducing norms have $E_{g,r}$ as their extreme point sets.

The extreme point sets for the unit balls of the low-rank inducing spectral norms are characterized next.

Corollary 1 *The extreme point set of the unit ball to the low-rank inducing spectral norm* $B^1_{\ell_\infty,r*}$ *is given by*

$$\mathcal{E}_r := \{X \in \mathbb{R}^{n\times m} : \sigma_1(X) = \cdots = \sigma_r(X) = 1 \ \textit{ and } \ \text{rank}(X) = r\}.$$

This result is proven in Appendix A.1.5.

We could also use the nuclear norm as a basis for the low-rank inducing norm. By Proposition 2, we know that $\|\cdot\|_{\ell_1,1*} = \|\cdot\|_{\ell_1}$. Therefore (11) implies that any low-rank inducing nuclear norm is just the nuclear norm, i.e.,

$$\|\cdot\|_{\ell_1} = \|\cdot\|_{\ell_1,q*} = \cdots = \|\cdot\|_{\ell_1,1*}.$$

Compared to using the low-rank inducing Frobenius and spectral norms, this does not provide us with a richer family of low-rank inducing norms.

4 Optimality Interpretations

In this section, we shown that low-rank inducing norms can be interpreted as the largest convex minorizers, i.e., the biconjugates of non-convex functions of the form (2), where the norm is arbitrary but unitarily invariant. Using this interpretation, we show how to create optimal convex relaxations of rank constrained optimization problems. This yields a posteriori guarantees on when a convex relaxation involving a low-rank inducing norm solves the corresponding rank constrained problem.

The interpretation of low-rank inducing norms follows as a special case of the following more general result.

Theorem 1 *Assume $f : \mathbb{R}_{\geq 0} \to \mathbb{R} \cup \{\infty\}$ is an increasing closed convex function, and let $f_{\text{reg}} := f(\|\cdot\|_g) + \chi_{\text{rank}(\cdot)\leq r}$ with $r \in \mathbb{N}$ such that $1 \leq r \leq \min\{m,n\}$. Then,*

$$f_{\text{reg}}^* = f^+(\|\cdot\|_{g^D,r}), \tag{14}$$

$$f_{\text{reg}}^{**} = f(\|\cdot\|_{g,r*}). \tag{15}$$

Proof. Since $\text{epi}(f(\|\cdot\|_{g,r*}))$ is closed by [28, Proposition IV.2.1.8], it follows by Lemma 1 that if

$$\text{epi}(f(\|\cdot\|_{g,r*})) = \text{conv}(\text{epi}(f_{\text{reg}})),$$

then (15) follows.

Let us start by showing that $\text{epi}(f(\|\cdot\|_{g,r*})) \subset \text{conv}(\text{epi}(f_{\text{reg}}))$. Assume that $(M,t) \in \text{epi}(f(\|\cdot\|_{g,r*}))$. By Lemma 3,

$$M = \textstyle\sum_i \alpha_i M_i \quad \text{with} \quad \sum_i \alpha_i = 1,\ \alpha_i \geq 0$$

where M_i satisfies

$$\text{rank}(M_i) \leq r, \quad \text{and} \quad \|M_i\|_{g,r*} = \|M\|_{g,r*}.$$

Hence, $(M,t) = \sum_i \alpha_i (M_i,t)$, where

$$t \geq f(\|M\|_{g,r*}) = f(\|M_i\|_{g,r*}) \quad \text{and} \quad \text{rank}(M_i) \leq r.$$

This shows that $(M_i,t) \in \text{epi}(f_{\text{reg}})$, and therefore $(M,t) \in \text{conv}(\text{epi}(f_{\text{reg}}))$.

Conversely, if $(M,t) \in \text{conv}(\text{epi}(f_{\text{reg}}))$, then

$$(M,t) = \textstyle\sum_i \alpha_i (M_i,t_i) \quad \text{with} \quad \sum_i \alpha_i = 1, \alpha_i \geq 0,$$

where M_i satisfies

$$\text{rank}(M_i) \leq r, \quad \text{and} \quad t_i \geq f(\|M_i\|_g) = f(\|M_i\|_{g,r*}),$$

where the equality is due to (12) in Lemma 2. Since f is convex and increasing, it holds that the composition $f(\|\cdot\|_{g,r*})$ is convex (see [28, Proposition IV.2.1.8]). Thus,

$$t := \textstyle\sum_i \alpha_i t_i \geq \sum_i \alpha_i f(\|M_i\|_{g,r*}) \geq f\left(\left\|\sum_i \alpha_i M_i\right\|_{g,r*}\right) = f\left(\|M\|_{g,r*}\right),$$

which implies that $(M,t) \in \text{epi}(f(\|\cdot\|_{g,r*}))$, and (15) follows. Applying [41, Theorem 15.3] to $f(\|\cdot\|_{g,r*})$ shows (14). □

This result generalizes the corresponding result in [23], in which the special case $f(x) \equiv x^2$ and $\|\cdot\|_g$ being the Frobenius norm is shown. For linear $f(x) \equiv x$, the biconjugate in (15) reduces to the low-rank inducing norms of Section 3. Therefore, they can be characterized as follows.

Corollary 2 *Let $r \in \mathbb{N}$ be such that $1 \leq r \leq q := \min\{m,n\}$. Then*

$$\|\cdot\|_{r*} = \left(\|\cdot\|_F + \chi_{\text{rank}(\cdot)\leq r}\right)^{**},$$
$$\|\cdot\|_{\ell_\infty,r*} = \left(\|\cdot\|_{\ell_\infty} + \chi_{\text{rank}(\cdot)\leq r}\right)^{**},$$

and the nuclear norm satisfies

$$\|\cdot\|_{\ell_1} = \left(\|\cdot\|_g + \chi_{\text{rank}(\cdot)\leq 1}\right)^{**},$$

where $\|\cdot\|_g$ is an arbitrary unitarily invariant norm that satisfies $\|M\|_g = \sigma_1(M)$ for all rank-1 matrices M.

Proof. This follows immediately from Theorem 1, since $\|\cdot\|_{r*} = \|\cdot\|_{\ell_2,r*}$, where $\|\cdot\|_{\ell_2} = \|\cdot\|_F$ is the Frobenius norm, and from Proposition 2. □

Remark 1 *This nuclear norm representation differs from the one in [17, 18], where it is shown that $\|\cdot\|_{\ell_1} = (\text{rank} + \chi_{B^1_{\ell_\infty}})^{**}$, i.e., it is the convex hull of the rank function restricted to the unit spectral norm ball.*

Using Theorem 1, optimal convex relaxations of rank constrained problems

$$\begin{aligned} \underset{M}{\text{minimize}} \quad & f_0(M) + f(\|M\|_g) \\ \text{subject to} \quad & \text{rank}(M) \leq r, \end{aligned} \tag{16}$$

can be provided, where $f_0 : \mathbb{R}^{n\times m} \to \mathbb{R}\cup\{\infty\}$ is a proper and closed convex function and $f : \mathbb{R}_{\geq 0} \to \mathbb{R}\cup\{\infty\}$ is an increasing and closed convex function. The problem in (16) is equivalent to minimizing $f_0 + f_{\text{reg}}$ with the non-convex f_{reg} defined in Theorem 1. Therefore, the optimal convex relaxation of (16) is given by

$$\underset{M}{\text{minimize}} \quad f_0(M) + f(\|M\|_{g,r*}). \tag{17}$$

Including an additional regularization parameter $\theta \geq 0$ (that can be included in f) yields the following proposition.

Proposition 4 *Assume that $f_0 : \mathbb{R}^{n\times m} \to \mathbb{R}\cup\{\infty\}$ is a proper closed convex function, and that $r \in \mathbb{N}$ is such that $1 \leq r \leq \min\{m,n\}$. Let $f : \mathbb{R}_{\geq 0} \to \mathbb{R}\cup\{\infty\}$ be an increasing, proper closed convex function, and let $\theta \geq 0$. Then*

$$\inf_{\substack{M\in\mathbb{R}^{n\times m} \\ \text{rank}(M)\leq r}} \left[f_0(M) + \theta f(\|M\|_g)\right] \geq \inf_{M\in\mathbb{R}^{n\times m}} \left[f_0(M) + \theta f(\|M\|_{g,r*})\right]. \tag{18}$$

If $M^\star$ solves the problem on the right such that $\text{rank}(M^\star) \leq r$, then equality holds, and $M^\star$ is also a solution to the problem on the left.

Proof. The inequality holds since $f(\|\cdot\|_{g,r*}) = f_{\text{reg}}^{**} \leq f_{\text{reg}}$. From Lemma 2 it follows that if $\text{rank}(M^\star) \leq r$ then

$$f_{\text{reg}}^{**}(M^\star) = f(\|M^\star\|_{g,r*}) = f(\|M^\star\|_g) = f_{\text{reg}}(M^\star),$$

which implies that the lower bound is attained with $M^\star$ and equality holds. □

Since the nuclear norm is obtained by creating a low-rank inducing norm with $r = 1$, it follows that any nuclear norm regularized problem can be interpreted as an optimal convex relaxation to a non-convex problem of the form (16), with the constraint $\text{rank}(M) \leq 1$.

Proposition 4 also covers the results in our previous work [23], where the matrix approximation problem

$$\begin{aligned}&\min_{\substack{M\in\mathbb{R}^{n\times m}\\ \text{rank}(M)\leq r}} \left[\tfrac{1}{2}\|N-M\|_F^2 + h(M)\right] \\ &\qquad = \min_{\substack{M\in\mathbb{R}^{n\times m}\\ \text{rank}(M)\leq r}} \left[\tfrac{1}{2}\|N\|_F^2 - \langle N, M\rangle + \tfrac{1}{2}\|M\|_F^2 + h(M)\right],\end{aligned}$$

is considered. Letting

$$f_0(\cdot) = \tfrac{1}{2}\|N\|_F^2 - \langle N, \cdot\rangle + h(\cdot), \quad f(x) = \tfrac{1}{2}x^2, \quad \text{and} \quad \|\cdot\|_g = \|\cdot\|_F,$$

the results in [23] are a special cases of Theorem 1.

5 Computability

This section addresses the computability of convex optimization problems involving low-rank inducing regularizers of the form $f(\|\cdot\|_{g,r*})$. We restrict ourselves to low-rank inducing Frobenius and spectral norm regularizers. A requirement for the optimal convex relaxation problem in (17) to be solved efficiently, is that these regularizers are suitable for numerical optimization.

Assuming that f_0 and f are SDP representable, it is shown that (17) can be solved via semi-definite programming. To be able to solve larger problems using first-order proximal splitting methods (see [12, 37] and references therein), we show how to efficiently compute the proximal mappings of the considered regularizers. The computational cost of computing these proximal mappings is comparable to the cost of computing the proximal mapping for the nuclear norm, since the cost in all cases is dominated by the singular value decomposition.

In order to deal with increasing convex functions f in (17), the problem is rewritten into the equivalent epigraph form

$$\underset{M,v}{\text{minimize}}\ f_0(M) + f(v) + \chi_{\text{epi}(\|\cdot\|_{g,r*})}(M, v). \tag{19}$$

5.1 SDP representation

The low-rank inducing Frobenius norm and spectral norm

$$\|M\|_{r*} := \max_{\|Y\|_r \le 1} \langle M, Y\rangle = \max_{\|Y\|_r^2 \le 1} \langle M, Y\rangle, \tag{20}$$

$$\|M\|_{\ell_\infty, r*} := \max_{\|Y\|_{\ell_1, r} \le 1} \langle M, Y\rangle, \tag{21}$$

are SDP representable via $\|Y\|_r^2$ and $\|Y\|_{\ell_1,r}$. From [22, 23], it is known that

$$\begin{aligned} \|Y\|_r^2 = \min_{T,\gamma} \quad & \mathrm{trace}(T) - \gamma(n-r) \\ \text{s.t.} \quad & \begin{pmatrix} T & Y \\ Y^T & I \end{pmatrix} \succeq 0,\ T \succeq \gamma I. \end{aligned}$$

Similarly, one can verify that

$$\begin{aligned} \|Y\|_{\ell_1,r} = \min_{T_1,T_2,\gamma} \quad & \frac{1}{2}[\mathrm{trace}(T_1) + \mathrm{trace}(T_2) - (n+m-2r)\gamma] \\ \text{s.t.} \quad & \begin{pmatrix} T_1 & Y \\ Y^T & T_2 \end{pmatrix} \succeq 0,\ T_1, T_2 \succeq \gamma I, \end{aligned}$$

which generalizes the SDP representation of $\|Y\|_{\ell_1,\min\{m,n\}}$ in [39]. This implies that

$$\begin{aligned} \|M\|_{r*} = \max_{Y,T,\gamma} \quad & \langle M, Y\rangle \\ \text{s.t.} \quad & \begin{pmatrix} T & Y \\ Y^T & I \end{pmatrix} \succeq 0,\ T \succeq \gamma I, \\ & \mathrm{trace}(T) - \gamma(n-r) \le 1, \end{aligned}$$

$$\begin{aligned} \|M\|_{\ell_\infty, r*} = \max_{Y,T_1,T_2,\gamma} \quad & \langle M, Y\rangle \\ \text{s.t.} \quad & \begin{pmatrix} T_1 & Y \\ Y^T & T_2 \end{pmatrix} \succeq 0,\ T_1, T_2 \succeq \gamma I, \\ & \frac{1}{2}[\mathrm{trace}(T_1) + \mathrm{trace}(T_2) - (n+m-2r)\gamma] \le 1, \end{aligned}$$

However, these formulations cannot be used in convex optimization problems with M as a decision variable due to the inner product $\langle M, Y\rangle$. Therefore, we use duality to

arrive at

$$\begin{aligned}\|M\|_{r*} = \min_{W_1,W_2,k} \quad & \frac{1}{2}(\operatorname{trace}(W_2)+k) \\ \text{s.t.} \quad & \begin{pmatrix} kI-W_1 & M \\ M^T & W_2 \end{pmatrix} \succeq 0,\ W_1 \succeq 0, \\ & \operatorname{trace}(W_1) = (n-r)k;\end{aligned}$$

$$\begin{aligned}\|M\|_{\ell_\infty,r*} = \min_{W_1,W_2,k} \quad & k \\ \text{s.t.} \quad & \begin{pmatrix} kI-W_1 & M \\ M^T & kI-W_2 \end{pmatrix} \succeq 0,\ W_1,W_2 \succeq 0, \\ & \operatorname{trace}(W_1)+\operatorname{trace}(W_2) = [(n-r)+(m-r)]k.\end{aligned}$$

These formulations can be used to, e.g. solve problems on the epigraph form (19) by enforcing the respective costs to be smaller than or equal to $v \in \mathbb{R}$. This gives constraints of the form $\|M\|_{g,r*} \leq v$, i.e., $(M,v) \in \operatorname{epi}(\|\cdot\|_{g,r*})$. If f and f_0 are SDP representable, then (19) can be solved via semi-definite programming.

5.2 Splitting algorithms

Conventional SDP solvers are often based on interior point methods (see [38, 43]). These have good convergence properties, but the iteration complexity typically grows unfavorably with the problem dimension. This limits their application to small or medium scale problems. First order proximal splitting methods (see e.g. [12, 37]) typically have a lower complexity per iteration, and are thus more suitable for large problems.

These methods require the proximal mapping for all non-smooth parts of the problem to be available. The *proximal mapping* for a proper closed and convex functions $h : \mathbb{R}^{n\times m} \to \mathbb{R} \cup \{\infty\}$ is defined as

$$\operatorname{prox}_{\gamma h}(Z) := \operatorname*{argmin}_X \left(h(X) + \frac{1}{2\gamma}\|X-Z\|_F^2 \right). \tag{22}$$

Applying proximal splitting methods to (19) therefore requires that the proximal mapping of $\chi_{\operatorname{epi}(\|\cdot\|_{g,r*})}$ is readily computable. Since $\chi_{\operatorname{epi}(\|\cdot\|_{g,r*})}$ is an indicator function of the epigraph set, the proximal mapping becomes a projection, which is denoted by $\Pi_{\operatorname{epi}(\|\cdot\|_{g,r*})}$.

The epigraph of a norm is a cone (see [5, Proposition 10.2]). Appealing to the Moreau-decomposition (see [5, Theorem 6.29]), we compute the projection $\Pi_{\operatorname{epi}(\|\cdot\|_{g,r*})}$ via

$$\Pi_{\operatorname{epi}(\|\cdot\|_{g,r*})}(Z,z_v) = (Z,z_v) - \Pi_{(\operatorname{epi}(\|\cdot\|_{g,r*})^\circ}(Z,z_v), \tag{23}$$

where $Z \in \mathbb{R}^{n\times m}$, $z_v \in \mathbb{R}$, and $\Pi_{(\operatorname{epi}(\|\cdot\|_{g,r*})^\circ}$ is projection onto the polar cone (which is the negative dual cone of $\operatorname{epi}(\|\cdot\|_{g,r*})$ by definition).

Algorithms for projecting onto the polar cones of the low-rank inducing Frobenius and spectral norms are derived in Appendix A.2. In these algorithms, the first step is to perform a singular value decomposition of the prox argument $Z \in \mathbb{R}^{n\times m}$. Then a vector optimization problem of dimension $q := \min\{m,n\}$ needs to be solved. To this end, a nested binary search is applied that only requires the solutions to simple optimization problems with at most $r+1$ decision variables.

In case of the low-rank inducing Frobenius norm, these problems can be solved explicitly, and results in an overall worst-case complexity of $\mathcal{O}(\log(r)\log(q-r))$ with an additional $\mathcal{O}(q)$ to set up the inner problems and to return the full solution. The cost of the prox computation is therefore dominated by the cost of computing the SVD. For large q one may consider sparse SVD algorithms such as [34].

The projection onto the epigraph of the low-rank inducing spectral norm is performed via the projection onto the epigraph of the truncated nuclear norm (modulo a sign flip). Since this requires a third layer in the nested binary search, the worst-case complexity is given by $\mathcal{O}(\log^2(r)\log(q-r)+q)$. In [47], another algorithm to project onto the truncated nuclear norm is presented. It uses similar techniques, but performs a linear search for finding the parameters and thus has a higher worst case computational cost.

Finally, note that the detour over the epigraph projection is not needed for all increasing functions. The proximal mapping for the low-rank inducing Frobenius and spectral norms can be derived very similarly to the epigraph case in Appendix A.2. The proximal mapping for the squared low-rank inducing Frobenius norm is derived in [16, 23]. Details are omitted for brevity.

6 Examples: Matrix Completion

The matrix completion problem seeks to complete a low-rank matrix based on limited knowledge about its entries. The problem is often posed as

$$\begin{aligned} \text{minimize} \quad & \operatorname{rank}(X) \\ \text{subject to} \quad & \hat{x}_{ij} = x_{ij},\ (i,j) \in \mathcal{I}, \end{aligned} \tag{24}$$

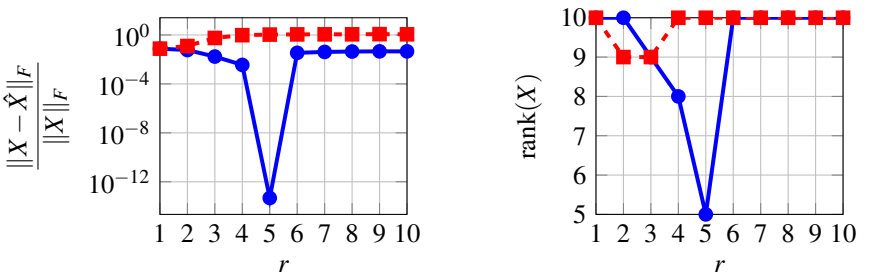
where $\mathcal{I}$ denotes the index set of the known entries. Another formulation that fits with the low-rank inducing norms proposed in this paper is

$$\begin{aligned} \text{minimize} \quad & \|X\|_g \\ \text{subject to} \quad & \operatorname{rank}(X) \leq r \\ & \hat{x}_{ij} = x_{ij},\ (i,j) \in \mathcal{I}, \end{aligned} \tag{25}$$

where r is the target rank of the matrix to be completed. In the following, two examples of this form will be convexified using different low-rank inducing norms. That is,

$$\begin{aligned} \text{minimize} \quad & \|X\|_{g,r*} \\ \text{subject to} \quad & \hat{x}_{ij} = x_{ij},\ (i,j) \in \mathcal{I}, \end{aligned} \tag{26}$$

is solved for different low-rank inducing norms $\|\cdot\|_{g,r*}$.



(a) Relative completion errors of:
(26) with $\|\cdot\|_{g,r*} = \|\cdot\|_{r*}$
(26) with $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_\infty,r*}$

(b) Rank of the solutions to:
(26) with $\|\cdot\|_{g,r*} = \|\cdot\|_{r*}$
(26) with $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_\infty,r*}$

Figure 1: Example 1: Relative completion error and ranks of the solution to (26) with $\|\cdot\|_{g,r*} = \|\cdot\|_{r*}$ and $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_\infty,r*}$.

Further, we discuss a covariance completion problem which is a generalization of the problem above. In all problems it will be observed that there are convex relaxations with low-rank inducing norms whose solutions give better completion than the nuclear norm approach, without increasing the rank.

6.1 Example 1

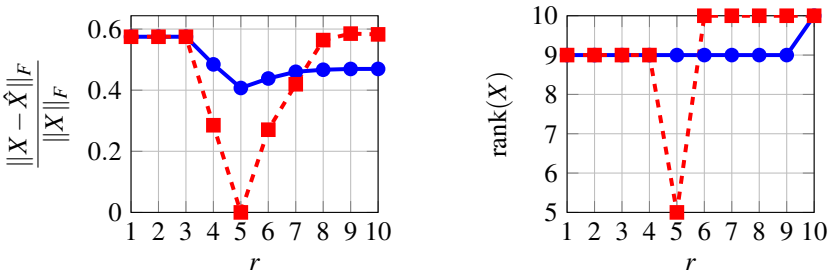
In the first problem, which is taken from [23], the matrix $\hat{X}$ to be completed is a low-rank approximation of the Hankel matrix

$$H = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & & \iddots & & 0 \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ 1 & & \iddots & & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix} \in \mathbb{R}^{10\times 10}. \tag{27}$$

Let the singular value decomposition of H be given by $H = \sum_{i=1}^{10} \sigma_i(H) u_i u_i^T$ and

$$\hat{X} := \sum_{i=1}^{5} \sigma_i(H) u_i u_i^T \quad \text{and } \mathcal{I} := \{(i,j) : \hat{x}_{ij} > 0\},$$

where $\mathcal{I}$ is the index set of known entries. The cardinality of $\mathcal{I}$ is 78, i.e. 22 out of 100 entries are unknown. Fig. 1 shows the completion errors and ranks of the completed matrices for different value of r. The nuclear norm ($r = 1$) returns a full rank matrix and gives a worse completion error than all other low-rank inducing Frobenius norms. For $r = 5$, the solution with the low-rank inducing Frobenius norm has rank 5. Given the known entries, this is the matrix of smallest Frobenius norm which has at most rank 5, by Proposition 4. As indicated by the small relative error, this matrix coincides with $\hat{X}$. In fact, this is also verified analytically in [23, Theorem 3].



(a) Relative completion errors of:
(26) with $\|\cdot\|_{g,r*} = \|\cdot\|_{r*}$
(26) with $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_\infty,r*}$

(b) Rank of the solutions to:
(26) with $\|\cdot\|_{g,r*} = \|\cdot\|_{r*}$
(26) with $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_\infty,r*}$

Figure 2: Example 2: Relative completion error and ranks of the solution to (26) with $\|\cdot\|_{g,r*} = \|\cdot\|_{r*}$ and $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_\infty,r*}$.

Notice that

$$10^{1.2}\text{rank}(\hat{X})\log(10) \gg \text{card}(\mathcal{I}) = 78,$$

which is why exact completion results for the nuclear norm (see [7]) do not apply. Furthermore, the low-rank inducing spectral norm shows no improvement in comparison with the nuclear norm.

6.2 Example 2

In the this second example, it as assumed that

$$\hat{X} := \sum_{j=1}^{5} \sigma_j \sum_{i=1}^{5} (H) u_i v_i^T \quad \text{and } \mathcal{I} := \{(i,j) : \hat{x}_{ij} > 0\},$$

where H is given in (27) with the singular value decomposition $H = \sum_{i=1}^{10} \sigma_i(H) u_i v_i^T$. The cardinality of $\mathcal{I}$ is 67, that is, 33 out of 100 entries are unknown. Fig. 2 shows the completion errors and ranks of the completed matrices with different value of r. The nuclear norm ($r = 1$) returns a close to full rank matrix with a relative completion error that is among the largest for all r. In this example, the low-rank inducing spectral norms perform significantly better than the low-rank inducing Frobenius norms. In particular, for $r = 5$, the low-rank inducing spectral norm returns a rank 5 solution. Given the known entries, this solution is the matrix of smallest spectral norm of rank at most 5 (see Proposition 4). As indicated by the zero completion error, this matrix coincides with $\hat{X}$. Just as in the exact recovery result for the low-rank inducing Frobenius norm in [23, Theorem 3], it can be analytically guaranteed that the low-rank inducing spectral norm with $r = 5$ recovers the true matrix. Analogous to the previous example,

$$10^{1.2}\text{rank}(\hat{X})\log(10) \gg \text{card}(\mathcal{I}) = 67,$$

which is why exact completion with the nuclear norm cannot be expected.

In both examples, the nuclear norm neither produces the lowest rank solution, nor recovers the true matrix. In contrast, other low-rank inducing norms succeed in both aspects. This indicates that the richness in the family of low-rank inducing norms should be exploited to achieve satisfactory performance in rank constrained problems. In practical applications, the 'true' matrix is not known, and this comparison cannot be made. However, cross validation techniques can often be used to assess the performance.

6.3 Covariance Completion

In this section, the performance of the low-rank inducing Frobenius and spectral norms is evaluated by means of a covariance completion problem. This is a variation of the matrix completion problems above.

Consider the linear state-space system

$$\dot{x}(t) = Ax(t) + Bu(t),$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $m \leq n$ and $u(t)$ is a zero-mean stationary stochastic process. For Hurwitz A and reachable (A,B), it has been shown (see [19, 20]) that the following are equivalent:

i. $X := \lim_{t\to\infty} \mathbf{E}\left(x(t)x^T(t)\right) \succeq 0$ is the steady-state covariance matrix of $x(t)$, where $\mathbf{E}(\cdot)$ denotes the expected value.

ii. $\exists H \in \mathbb{R}^{m\times n} : AX + XA^T = -(BH + H^T B^T)$.

iii. $\operatorname{rank}\begin{pmatrix} AX + XA^T & B \\ B^T & 0 \end{pmatrix} = \operatorname{rank}\begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}$.

In particular, $H = \frac{1}{2}\mathbf{E}\left(u(t)u^T(t)\right)B^T$ if u is white noise. The problem considered in [11, 33, 48–50] is to reconstruct the partially known covariance matrix X and the input matrix B, via $M = -(BH + H^T B^T)$, where the rank of M sets an upper bound on the rank of B, i.e., the number of inputs. The objective is to keep the rank of M low, while achieving satisfactory completion of X. In [11, 33, 48–50] the problem is addressed by searching for the lowest rank solution:

$$\begin{aligned} \text{minimize} \quad & \operatorname{rank}(M) \\ \text{subject to} \quad & \hat{x}_{ij} = x_{ij},\ (i,j) \in \mathcal{I} \\ & A\hat{X} + \hat{X}A^T = -M \\ & \hat{X} \succeq 0, \end{aligned} \tag{28}$$

where $\mathcal{I}$ denotes set of pairs of indices of known entries. Another option is to search for a low-rank solution, while minimizing the norm of M measured by some unitarily invariant norm. This helps to avoid overfitting, and gives

$$\begin{aligned} \text{minimize} \quad & \|M\|_g \\ \text{subject to} \quad & \operatorname{rank}(M) \leq r \\ & \hat{x}_{ij} = x_{ij},\ (i,j) \in \mathcal{I} \\ & A\hat{X} + \hat{X}A^T = -M \\ & \hat{X} \succeq 0. \end{aligned} \tag{29}$$

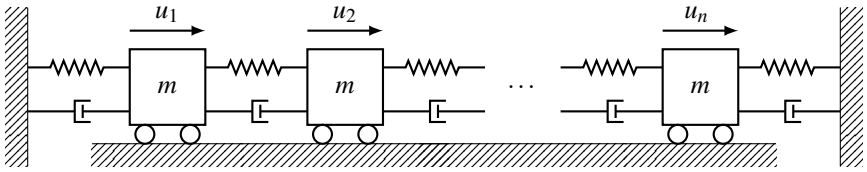


Figure 3: Mass-spring-damper system with n masses and input forces $u_1, \ldots, u_n$.

The authors in [11, 33, 48–50] convexify the problem by using the nuclear norm. In [24], a similar problem is instead convexified with the low-rank inducing Frobenius norm. We will also make a comparison with convex relaxations based on low-rank inducing spectral norms. All these convex relaxations are of the form

$$\begin{aligned} \text{minimize} \quad & \|M\|_{g,r*} \\ \text{subject to} \quad & \hat{x}_{ij} = x_{ij},\ (i,j) \in \mathcal{I} \\ & A\hat{X} + \hat{X}A^T = -M \\ & \hat{X} \succeq 0, \end{aligned} \tag{30}$$

with the appropriate low-rank inducing norm in the cost.

6.3.1 Mass-spring-damper system

The system considered in our example is the so-called *mass-spring-damper system (MSD)* (see [24, 49]) with n masses (see Fig. 3).

Assuming that the stochastic forcing affects all masses, this yields the following state-space representation

$$\dot{x}(t) = Ax(t) + B\xi(t)$$

with

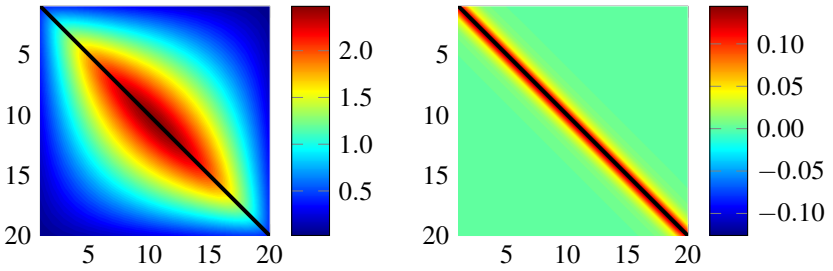
$$A = \begin{pmatrix} 0 & I \\ -S & -I \end{pmatrix} \in \mathbb{R}^{2n\times 2n}, \quad B = \begin{pmatrix} 0 \\ I \end{pmatrix} \in \mathbb{R}^{2n\times n}.$$

Here, S is a symmetric tridiagonal Toeplitz matrix with 2 on the main diagonal, -1 on the first upper and lower sub-diagonals, and I and 0 stand for the identity and zero matrices of appropriate size. The state vector x consists of the positions and velocities of the masses, $x = (p, v)$. Furthermore, $\xi(t)$ is generated via a low-pass filtered white noise signal $w(t)$ with unit covariance $\mathbf{E}\left(w(t)w(t)^T\right) = I$ as

$$\dot{\xi}(t) = -\xi(t) + w(t).$$

The extended covariance matrix

$$X_e := \mathbf{E}\left(x_e x_e^T\right) = \begin{pmatrix} X & X_{x\xi} \\ X_{\xi x} & X_\xi \end{pmatrix} \text{ with } x_e := \begin{pmatrix} x(t) \\ \xi(t) \end{pmatrix}$$



(a) X_{pp} (b) X_{vv}

Figure 4: Interpolated colormap of the steady-state covariance matrices X_{pp} and X_{vv} of the positions and the velocities in the MSD system with $n = 20$. ━━ indicates the available one-point correlations.

is then determined by

$$A_e X_e + X_e A_e^T = -B_e B_e^T,$$

where X is the steady-state covariance matrix of $x(t)$ and

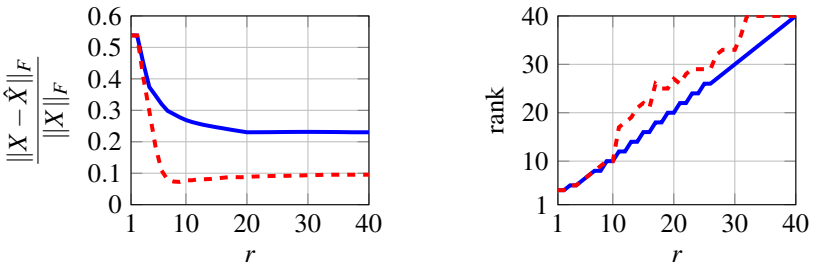
$$A_e := \begin{pmatrix} A & B \\ 0 & -I \end{pmatrix}, \quad B_e := \begin{pmatrix} 0 \\ I \end{pmatrix}.$$

In our numerical experiments, we choose $n = 20$ masses and assume that only one-point correlations are available, i.e. the known entries are given by the diagonal of X. The steady-state covariance matrix can be partitioned as

$$X = \begin{pmatrix} X_{pp} & X_{pv} \\ X_{vp} & X_{vv} \end{pmatrix},$$

where X_{pp} and X_{vv} are the covariance matrices of the positions and the velocities, respectively. To visualize the effects of using different low-rank inducing norms in (30), an interpolated colormap of the reconstructed $\hat{X}_{pp}$ and $\hat{X}_{vv}$ is used (see Fig. 6). The interpolated colormap of the true covariance matrices is shown in Fig. 4, where the black lines indicate the known measured entries.

Fig. 5 displays the relative errors and the ranks of the estimates obtained by (30) for different low-rank inducing norms as functions of r. The nuclear norm minimization ($r = 1$), as shown in Figures 6a and 6b, gives the same rank as both the low-rank inducing Frobenius and spectral norms for $r = 2$. However, the latter approaches give better completions. The low-rank inducing spectral norm outperforms the low-rank inducing Frobenius norm for all $r \geq 2$. In particular, $r = 9$ gives the best completion, with a solution of rank 10 (see Figures 6e and 6f). It is interesting that the solutions to (30) with $r = 10$ for both the low-rank inducing Frobenius and spectral norms are of rank 10. By Proposition 4, there are no better feasible rank-10 solutions that minimize



(a) Relative completion errors of:
(30) with $\|\cdot\|_{g,r*} = \|\cdot\|_{r*}$
(30) with $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_\infty,r*}$

(b) Rank of the solutions to:
(30) with $\|\cdot\|_{g,r*} = \|\cdot\|_{r*}$
(30) with $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_\infty,r*}$

Figure 5: Relative errors and ranks of solutions to (30) with $\|\cdot\|_{g,r*} = \|\cdot\|_{r*}$ and $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_\infty,r*}$.

the Frobenius and spectral norms respectively. The solution to (30) with the low-rank inducing Frobenius norm and $r = 10$, is shown in Figure 6c and 6d. The solution to the low-rank inducing spectral norm with $r = 10$ looks identical to Figures 6e and 6f.

7 Extensions

7.1 The Vector Case

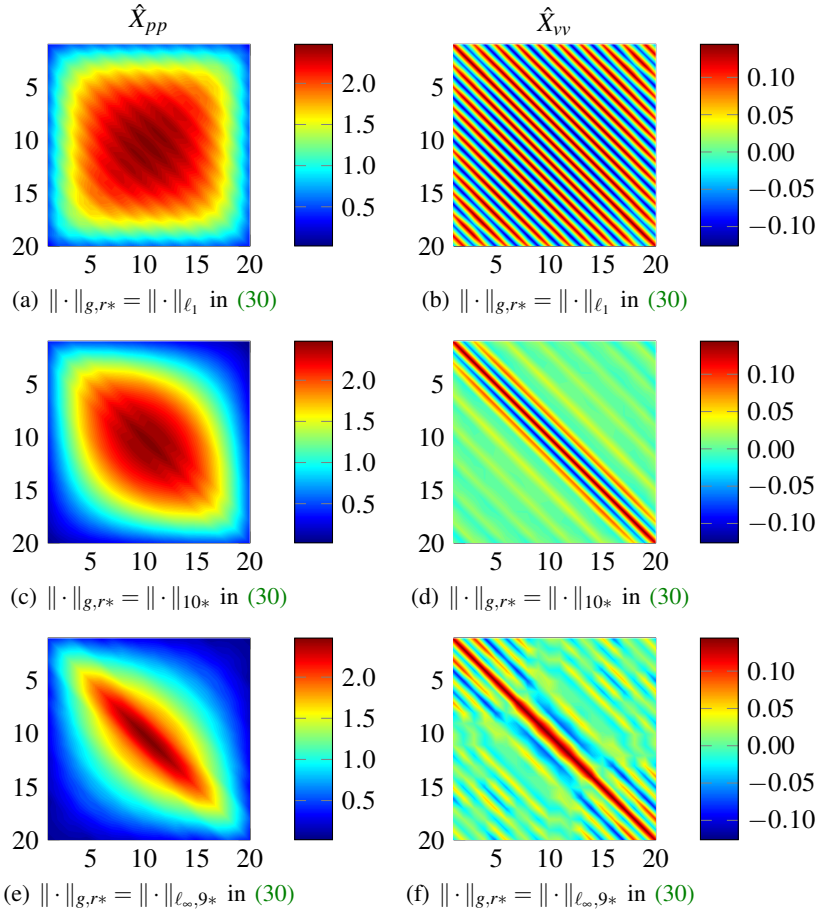
The results in Section 4 translate to the corresponding vector-valued problem, by replacing rank with cardinality, and $\|M\|_g$ with $\|x\|_g := \|diag(x)\|_g$. Therefore, our optimality interpretations, as well as the variety of regularizers, can be applied to problems such as sparse linear regression (see [3, 8, 42]). The SDP representation and the proximal mapping computations in Section 5 carry over, though here they have lower computational cost. For instance, the required SVD in the prox computations turns into a sorting, which reduces the total complexity.

7.2 Atomic Norms

In [9], the concept of an atomic norm is introduced. An atomic norm is defined as the gauge function or the Minkowski functional of the convex hull of a set of atoms $\mathcal{A}$ (see [9])

$$\|x\|_{\mathcal{A}} := \inf\{t > 0 : t^{-1}x \in \text{conv}(\mathcal{A})\}. \tag{31}$$

Despite its name, the atomic norm is not necessarily a norm, but always defines a distance measure. The atoms are used to model properties of a quantity that is to be estimated. The atomic norm is a way of imposing these properties on the solution of an optimization problem. In [9], examples of atomic sets that naturally appear in different applications are listed. For instance, if $\mathcal{A}$ is the set of rank 1 matrices with unit



(a) $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_1}$ in (30)
(b) $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_1}$ in (30)
(c) $\|\cdot\|_{g,r*} = \|\cdot\|_{10*}$ in (30)
(d) $\|\cdot\|_{g,r*} = \|\cdot\|_{10*}$ in (30)
(e) $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_\infty,9*}$ in (30)
(f) $\|\cdot\|_{g,r*} = \|\cdot\|_{\ell_\infty,9*}$ in (30)

Figure 6: Recovered covariance matrices of positions ($\hat{X}_{pp}$ to the left), and velocities ($\hat{X}_{vv}$ to the right), in the MSD system with $n = 20$ masses resulting from problem (30), with different low-rank inducing norms.

Frobenius norm, then the resulting atomic norm is the nuclear norm. More generally, all low-rank inducing norms in Section 3 can be considered as atomic norms, because Lemma 3 implies that

$$\|X\|_{g,r*} = \inf\left\{t > 0 : t^{-1}X \in \operatorname{conv}(E_{g,r})\right\},$$

with $E_{g,r} := \{X \in \mathbb{R}^{n\times m} : \|X\|_g = 1, \operatorname{rank}(X) \leq r\}$.

As presented for the low-rank inducing norms and regularizers in Section 4, this section provides similar optimality interpretations for general atomic norms. It is assumed that the atoms lie within a finite-dimensional real Hilbert space $\mathcal{H}$ with inner product $\langle\cdot,\cdot\rangle$, i.e. $\mathcal{A} \subset \mathcal{H}$. In the following, the definitions of the *conic hull* of $\mathcal{A} \subset \mathcal{H}$

$$\operatorname{cone}(\mathcal{A}) := \{\alpha x : x \in \mathcal{A},\ \alpha \geq 0\},$$

and the *polar gauge function* to (31)

$$\|y\|_{\mathcal{A}}^{\circ} := \inf\{\mu \geq 0 : \langle x, y\rangle \leq \mu\|x\|_{\mathcal{A}} \text{ for all } x \in \mathcal{H}\},$$

are needed. Note that, if the atomic norm in (31) is a norm, then the polar gauge function is equal to the corresponding dual norm Our optimality interpretations will hold if the atomic set denoted by $\mathcal{A}_G$ can be represented as

$$\mathcal{A}_G := \{a \in \operatorname{cone}(\mathcal{A}) : G(a) = 1\}, \tag{32}$$

where $\mathcal{A} \subset \mathcal{H}$, and $G : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$, satisfy the following assumptions.

Assumption 1 *The set $\mathcal{A} \subset \mathcal{H}$ is nonempty such that* $\operatorname{cone}(\mathcal{A})$ *is closed. The function $G : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ is positively homogeneous (of degree 1), proper, closed, convex and nonnegative with $G(a) > 0$ for all $a \in \mathcal{A} \setminus \{0\}$.*

Many atomic sets from [9] satisfy these assumptions. For example, if $\mathcal{A}$ is the set of all permutation matrices, then

$$\|\cdot\|_{\mathcal{A}} = \|\cdot\|_{\mathcal{A}_G} \quad \text{with} \quad G(\cdot) = \|\cdot\|_{\ell_\infty}.$$

Similar constructions apply to the atomic norms that are induced, e.g. by *binary vectors*, *sparse vectors*, *low-rank matrices*, *vectors from lists*, and many more (see [9])

Using the definition of atomic norms in (31), an explicit expression of the atomic norm associated with $\mathcal{A}_G$ is

$$\|x\|_{\mathcal{A}_G} = \inf\{t > 0 : t^{-1}x \in \operatorname{conv}(\{a \in \operatorname{cone}(\mathcal{A}) : G(a) = 1\})\}. \tag{33}$$

The next theorem gives optimality interpretations of these atomic norms, and generalizes Theorem 1 in the following two aspects:

I. The rank-constraint is generalized to other non-convex constraints.

II. The norms are replaced by more general functions G.

To prove the result, the following lemma is needed.

Lemma 4 *Let $\mathcal{A} \subset \mathcal{H}$ and $G : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ satisfy Assumption 1, and let $\mathcal{A}_G$ and $\|\cdot\|_{\mathcal{A}_G}$ be defined as in (32) and (33). Then,*

i. $\operatorname{conv}(\mathcal{A}_G)$ *is closed and bounded.*

ii. $\|x\|_{\mathcal{A}_G} = 0$ *if and only if* $x = 0$.

iii. $\|x\|_{\mathcal{A}_G} \geq G(x)$ *for all* $x \in \mathcal{H}$, *and* $\|x\|_{\mathcal{A}_G} = G(x)$ *for all* $x \in \operatorname{cone}(\mathcal{A})$.

iv. For all $x \in \operatorname{dom}(\|\cdot\|_{\mathcal{A}_G})$ *there exist* $x_i \in \operatorname{cone}(\mathcal{A})$ *such that*

$$x = \textstyle\sum_i \alpha_i x_i, \quad \sum_i \alpha_i = 1,\ \alpha_i \geq 0, \quad \text{and} \quad G(x_i) = \|x\|_{\mathcal{A}_G}.$$

Proof. Item i: Since $G + \chi_{\operatorname{cone}(\mathcal{A})}$ is coercive, it follows from [5, Proposition 11.11] that the sub-level set

$$\{a \in \operatorname{cone}(\mathcal{A}) : G(a) \leq 1\}$$

is bounded. Thus the same applies to $\mathcal{A}_G$. Further, convexity of G implies that

$$\{x \in \mathcal{H} : G(x) = 1\}$$

is closed, because, by [28, Proposition VI.1.3.3], it is the boundary of

$$\{x \in \mathcal{H} : G(x) \leq 1\}.$$

Thus, as the intersection of two closed sets is closed,

$$\mathcal{A}_G = \operatorname{cone}(\mathcal{A}) \cap \{x \in \mathcal{H} : G(x) = 1\}$$

is closed. Applying [28, Theorem III.1.4.3] shows that $\operatorname{conv}(\mathcal{A}_G)$ is closed and bounded.

Item ii: This claim follows by [28, Corollary V.1.2.6].

Item iii: Let us introduce the sub-levelset

$$S_G^s := \{x \in \mathcal{H} : G(x) \leq s\},$$

which by the positive homogeneity of G satisfies

$$S_G^s = \{sx \in \mathcal{H} : G(x) \leq 1\}$$

for all $s \geq 0$. By the definition of $\mathcal{A}_G$, it holds that

$$\begin{aligned}
\operatorname{conv}(\mathcal{A}_G) &= \operatorname{conv}(\{a \in \operatorname{cone}(\mathcal{A}) : G(a) = 1\}) \\
&\subset \operatorname{conv}(\{a \in \operatorname{cone}(\mathcal{A}) : G(a) \leq 1\}) \\
&\subset \operatorname{conv}(\{a \in \mathcal{H} : G(a) \leq 1\}) \\
&= \{a \in \mathcal{H} : G(a) \leq 1\}) = S_G^1.
\end{aligned}$$

This yields that

$$\begin{aligned}\|x\|_{\mathcal{A}_G} &= \inf\{t>0 : x \in t\operatorname{conv}(\mathcal{A}_G)\}\\ &\geq \inf\{t>0 : x \in tS_G^1\}\\ &= \inf\{t>0 : G(x) \leq t\} = G(x)\end{aligned}$$

for all $x \in \mathcal{H}$, and the first claim of this item is proven.

To prove the second claim, let $x \in \operatorname{cone}(\mathcal{A})$. If $x \notin \operatorname{dom}(G)$, the above implies that

$$\|x\|_{\mathcal{A}_G} = G(x) = \infty.$$

Further, Item ii shows that

$$x = 0 \;\Rightarrow\; \|0\|_{\mathcal{A}_G} = G(0) = 0.$$

It remains to show the claim for $x \in \operatorname{dom}(G) \setminus \{0\}$. In this case, we can define $\bar{x} := G(x)^{-1}x$, which satisfies

$$\bar{x} \in \operatorname{cone}(\mathcal{A}) \quad \text{and} \quad G(\bar{x}) = 1,$$

i.e. $\bar{x} \in \mathcal{A}_G \subset \operatorname{conv}(\mathcal{A}_G)$, and therefore

$$\|\bar{x}\|_{\mathcal{A}_G} = \inf\{t>0 : \bar{x} \in t\operatorname{conv}(\mathcal{A}_G)\} \leq \inf\{t>0 : \bar{x} \in t\bar{x}\} = 1.$$

That is, $\|x\|_{\mathcal{A}_G} \leq G(x)$, which in conjunction with $\|x\|_{\mathcal{A}_G} \geq G(x)$ proves that

$$\|x\|_{\mathcal{A}_G} = G(x) \text{ for all } x \in \operatorname{cone}(\mathcal{A}).$$

Item iv: Since the claim holds trivially if $x = 0$, it is enough to assume that

$$x \in \operatorname{dom}(\|\cdot\|_{\mathcal{A}_G}) \setminus \{0\}.$$

By Item ii it follows that $\infty > \|x\|_{\mathcal{A}_G} > 0$. Further, Item i and the definition of $\|x\|_{\mathcal{A}_G}$ in (31) imply that

$$\|x\|_{\mathcal{A}_G}^{-1} x \in \operatorname{conv}(\mathcal{A}_G).$$

Thus,

$$\|x\|_{\mathcal{A}_G}^{-1} x = \textstyle\sum_i \alpha_i \bar{x}_i \quad \text{with} \quad \sum_i \alpha_i = 1,\ \alpha_i \geq 0,$$

where $\bar{x}_i$ satisfies

$$\bar{x}_i \in \operatorname{cone}(\mathcal{A}) \quad \text{and} \quad G(\bar{x}_i) = 1.$$

Defining $x_i := \bar{x}_i \|x\|_{\mathcal{A}_G}$, it follows that

$$x = \sum_i \alpha_i x_i \quad \text{with} \quad x_i \in \operatorname{cone}(\mathcal{A}).$$

Finally, the positive homogeneity of G ensures that

$$G(x_i) = G(\|x\|_{\mathcal{A}_G} \bar{x}_i) = \|x\|_{\mathcal{A}_G} G(\bar{x}_i) = \|x\|_{\mathcal{A}_G}.$$

□

Theorem 2 *Assume $\mathcal{A} \subset \mathcal{H}$ and $G : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ satisfy Assumption 1, and let $\mathcal{A}_G$ and $\|\cdot\|_{\mathcal{A}_G}$ be defined as in (32) and (33). Further, let $f_{\text{reg}} := f(G(\cdot)) + \chi_{\text{cone}(\mathcal{A})}$, where $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ is an increasing, proper closed convex function. Then,*

$$f^*_{\text{reg}} = f^+(\|\cdot\|^\circ_{\mathcal{A}_G}), \tag{34}$$

$$f^{**}_{\text{reg}} = f(\|\cdot\|_{\mathcal{A}_G}). \tag{35}$$

Proof. Since $\|\cdot\|_{\mathcal{A}_G}$ is a Minkowski functional, it is closed function (see [35, Lemma 1 in 5.12]). Thus, $\text{epi}(f(\|\cdot\|_{\mathcal{A}_G})$ is a closed set, and by Lemma 1,

$$\text{epi}(f(\|\cdot\|_{\mathcal{A}_G})) = \text{conv}(\text{epi}(f_{\text{reg}})) \tag{36}$$

implies (35).

We start with $\text{conv}(\text{epi}(f_{\text{reg}})) \subset \text{epi}(f(\|\cdot\|_{\mathcal{A}_G}))$. If $(x,t) \in \text{conv}(\text{epi}(f_{\text{reg}}))$, then

$$(x,t) = \sum_i \alpha_i (x_i, t_i) \quad \text{with} \quad \sum_i \alpha_i = 1,\ \alpha_i \geq 0,$$

where x_i satisfies

$$x_i \in \text{cone}(\mathcal{A}), \quad \text{and} \quad t_i \geq f(G(x_i)) = f(\|x_i\|_{\mathcal{A}_G}),$$

and the equality follows by Lemma 4(Item iii). Since f is convex and increasing, it holds that the composition $f(\|\cdot\|_{\mathcal{A}_G})$ is convex (see [28, Proposition IV.2.1.8]). Therefore,

$$t := \textstyle\sum_i \alpha_i t_i \geq \sum_i \alpha_i f(\|x_i\|_{\mathcal{A}_G}) \geq f\left(\|\sum_i \alpha_i x_i\|_{\mathcal{A}_G}\right) = f(\|x\|_{\mathcal{A}_G}),$$

and $(x,t) \in \text{epi}(f(\|\cdot\|_{\mathcal{A}_G}))$.

Conversely, let $(x,t) \in \text{epi}(f(\|\cdot\|_{\mathcal{A}_G}))$ with $\|x\|_{\mathcal{A}_G} \neq 0$. Lemma 4(Item iv) implies that

$$x = \textstyle\sum_i \alpha_i x_i \quad \text{with} \quad \sum_i \alpha_i = 1, \alpha_i \geq 0,$$

where x_i satisfies

$$x_i \in \text{cone}(\mathcal{A}) \quad \text{and} \quad G(x_i) = \|x\|_{\mathcal{A}_G}.$$

Thus, $(x,t) = \sum_i \alpha_i (x_i, t)$ such that

$$t \geq f(\|x\|_{\mathcal{A}_G}) = f(G(x_i)), \quad \text{and} \quad x_i \in \text{cone}(\mathcal{A}).$$

Consequently,

$$(x_i, t) \in \text{epi}(f_{\text{reg}}), \text{ and therefore } (x,t) \in \text{conv}(\text{epi}(f_{\text{reg}})).$$

Lemma 4 (Item ii) shows that $(x,t) \in \text{conv}(\text{epi}(f_{\text{reg}}))$ is trivially fulfilled if $\|x\|_{\mathcal{A}_G} = 0$. Finally, (34) can be proven by applying [41, Theorem 15.3] to $f(\|\cdot\|_{\mathcal{A}_G})$. □

Similarly to Section 4, this result gives rise to optimal convex relaxations for atomic norms.

Proposition 5 *Assume $\mathcal{A} \subset \mathcal{H}$ and $G : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ satisfy Assumption 1, and let $\mathcal{A}_G$ and $\|\cdot\|_{\mathcal{A}_G}$ be defined as in (32) and (33). Further, let $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ be an increasing, closed convex function, and let $f_0 : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ be closed, proper, and convex. For $\theta \geq 0$, it holds that*

$$\inf_{x \in \mathcal{A}} \left[f_0(x) + \theta f(G(x)) \right] \geq \inf_{x \in \operatorname{conv}(\mathcal{A})} \left[f_0(x) + \theta f(\|x\|_{\mathcal{A}_G}) \right]. \tag{37}$$

If the right-hand side of the inequality is solved by $x^\star \in \mathcal{A}$, then $x^\star$ is a solution to the left-hand side.

Proof. By (4) and Theorem 2 it follows that

$$\inf_{x \in \operatorname{cone}(\mathcal{A})} \left[\tilde{f}_0(x) + \theta f(G(x)) \right] \geq \inf_{x \in \mathcal{H}} \left[\tilde{f}_0(x) + \theta f(\|x\|_{\mathcal{A}_G}) \right], \tag{38}$$

for any closed and proper convex function $\tilde{f}_0 : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$. In particular, let $\tilde{f}_0 = f_0 + \chi_{\operatorname{conv}(\mathcal{A})}$, which is closed by Assumption 1. Then the left-hand side of (38) satisfies

$$\begin{aligned} \inf_{x \in \operatorname{cone}(\mathcal{A})} \left[\tilde{f}_0(x) + \theta f(G(x)) \right] &= \inf_{\substack{x \in \operatorname{cone}(\mathcal{A}) \\ x \in \operatorname{conv}(\mathcal{A})}} \left[f_0(x) + \theta f(G(x)) \right] \\ &\leq \inf_{x \in \mathcal{A}} \left[f_0(x) + \theta f(G(x)) \right], \end{aligned}$$

because $\mathcal{A} \subset \operatorname{cone}(\mathcal{A}) \cap \operatorname{conv}(\mathcal{A})$. The right-hand side of (38) satisfies

$$\inf_{x \in \mathcal{H}} \left[\tilde{f}_0(x) + \theta f(\|x\|_{\mathcal{A}_G}) \right] = \inf_{x \in \operatorname{conv}(\mathcal{A})} \left[f_0(x) + \theta f(\|x\|_{\mathcal{A}_G}) \right],$$

and (37) is proven. The last claim follows by Lemma 4 (Item iii). □

In [9] exact recovery results are presented for the cases when f_0 is an indicator of an affine set that contains measurement of an observed quantity $x_0 \in \mathcal{H}$. Let $\mathcal{Q} := \{x \in \mathcal{H} : Ax = Ax_0\}$ denote that affine set and let $f_0 = \chi_{\mathcal{Q}}$. Then the recovery problem becomes

$$\operatorname*{minimize}_{x \in \mathcal{Q}} \|x\|_{\mathcal{A}_G}.$$

Assume that this problem has a unique solution $x^\star$. In [9], conditions on the measurement set $\mathcal{Q}$ are stated under which exact recovery $x^\star = x_0$ is guaranteed. The underlying assumption in [9], is that for small k it holds that

$$x_0 = \sum_{i=1}^{k} c_i a_i \text{ with } c_i \geq 0 \text{ and } a_i \in \mathcal{A}_G.$$

That is, the observed quantity is assumed to be a conic combination of a few atoms. For many examples in [9, Section 2.2], the assumption holds with $k = 1$ and $c_1 = 1$, i.e., $x_0 = a$ for some $a \in \mathcal{A}$. A notable exception is the case of low rank matrix recovery. In [9], rank one matrices of unit norm are used as atoms, which yields the nuclear norm as the corresponding atomic norm. Therefore, a conic combination of r atoms is needed

to recover a rank-r matrix x_0. By using a low-rank inducing norm $\|\cdot\|_{g,r*}$ instead, the matrix x_0 satisfies $x_0 = a$ for some $a \in \mathcal{A}$, where $\mathcal{A}$ is the set of matrices with rank less than or equal to r. With this atomic set, the problem in [9] reduces to recover $x_0 = a$, where $a \in \mathcal{A}$. Upon successful recovery, the convex atomic norm minimization problem on the right-hand side of (37) solves the corresponding non-convex problem on its left-hand side.

8 Conclusion

We have proposed a family of low-rank inducing norms and regularizers. These norms are interpreted as the largest convex minorizers of a unitarily invariant norm that is restricted to matrices of at most rank r. One feature of these norms is that optimality interpretations in the form of a posteriori guarantees can be provided. In particular, it can be checked if the solutions to a convex relaxation involving low-rank inducing norms, also solve an underlying rank constrained problem. Our numerical examples indicate that this is useful for, e.g. the so-called matrix completion problem. A suitably chosen low-rank inducing norm yields significantly better completion and/or lower rank than the commonly used nuclear norm approach. This has been demonstrated on the basis of what we called low-rank inducing Frobenius and spectral norms. Both norms have been shown to have cheaply computable proximal mappings, as well as simple SDP representations. As a result, this extends proximal mapping computations that are found, in e.g. [16, 23, 47]. Moreover, The class of low-rank inducing norms can be further broadened by using continuous r as discussed in [23] for the low-rank inducing Frobenius norm. Finally, it has been highlighted that our findings also generalize to atomic norms, and to other non-convex problems.

A Appendix

A.1 Proofs to Results in Section 3

A.1.1 Proof to Lemma 2

Proof. Let $1 \leq r \leq q := \min\{m,n\}$, $g : \mathbb{R}^q \to \mathbb{R}_{\geq 0}$ be a symmetric gauge function, $\Sigma_j(M) := \operatorname{diag}(\sigma_1(M),\ldots,\sigma_j(M))$ for $M \in \mathbb{R}^{n\times m}$, and $1 \leq j \leq q$. Then for all $Y \in \mathbb{R}^{n\times m}$,

$$\begin{aligned}\|Y\|_{g^D,r} &= \max_{\substack{\operatorname{rank}(M)\leq r\\ \|M\|_g\leq 1}} \langle M, Y\rangle = \max_{\substack{\operatorname{rank}(\Sigma_q(M))\leq r\\ \|\Sigma_q(M)\|_g\leq 1}} \langle \Sigma_q(Y), \Sigma_q(M)\rangle \\ &= \max_{\|\Sigma_r(M)\|_g\leq 1} \langle \Sigma_r(Y), \Sigma_r(M)\rangle = \|\Sigma_r(Y)\|_{g^D},\end{aligned}$$

where the second equality follows by [29, Corollary 7.4.1.3(c)]. Further, $\|\cdot\|_{g^D,r}$ is unitarily invariant, since

$$\|\Sigma_r(Y)\|_{g^D} = g^D(\sigma_1(Y),\ldots,\sigma_r(Y))$$

defines a symmetric gauge function (see Proposition 1). Similarly to the above, this implies that

$$\begin{aligned}\|M\|_{g,r*} &= \max_{\|Y\|_{g^D,r}\leq 1} \langle M,Y\rangle = \max_{g^D(\sigma_1(Y),\dots,\sigma_r(Y))\leq 1} \sum_{i=1}^{q} \sigma_i(M)\sigma_i(Y) \\ &= \max_{g^D(\sigma_1(Y),\dots,\sigma_r(Y))\leq 1} \left[\sum_{i=1}^{r} \sigma_i(M)\sigma_i(Y) + \sigma_r(Y) \sum_{i=r+1}^{q} \sigma_i(M)\right].\end{aligned}$$

It remains to prove (11) and (12). The constraint set for $r+1$ is a superset of the constraint set for r and by the definition of $\|\cdot\|_{g^D,r}$ in (9) it follows that $\|Y\|_{g^D,r} \leq \|Y\|_{g^D,r+1}$. Therefore,

$$\|M\|_{g,r*} = \max_{\|Y\|_{g^D,r}\leq 1} \langle M,Y\rangle \geq \max_{\|Y\|_{g^D,r+1}\leq 1} \langle M,Y\rangle = \|M\|_{g,(r+1)*}.$$

Note that $\|\cdot\|_{g^D} = \|\cdot\|_{g^D,q}$, which implies that $\|\cdot\|_{g,q*} = \|\cdot\|_g$ and thus (11) is proven. The implication in (12) follows from the derived expression for $\|\cdot\|_{g,r*}$, since for rank-r matrices M, $\sigma_i(M) = 0$ for all $i \in \{r+1,\dots,q\}$. □

A.1.2 Proof to Proposition 2

By [29, Corollary 7.4.1.3(c)] it holds that $g^D(\sigma_1) = \sigma_1$ if and only if $g(\sigma_1) = \sigma_1$. Thus, (10) yields for all $M \in \mathbb{R}^{n\times m}$ that

$$\|M\|_{g,1*} = \max_{\sigma_1(Y)\leq 1} \sigma_r(Y) \sum_{i=1}^{\min\{m,n\}} \sigma_i(M) = \|M\|_{\ell_\infty^D} = \|M\|_{\ell_1},$$

where we use the fact that the dual norm of the spectral norm is the nuclear norm (see [29, Theorem 5.6.42]).

A.1.3 Proof to Lemma 3

Proof. By definition of $\|\cdot\|_{g^D,r}$ in (9) in Lemma 2, it holds that for all $Y \in \mathbb{R}^{n\times m}$,

$$\max_{X\in\operatorname{conv}(E_{g,r})} \langle X,Y\rangle = \max_{\substack{\operatorname{rank}(X)\leq r \\ \|X\|_{g^D}\leq 1}} \langle X,Y\rangle = \|Y\|_{g^D,r} = \max_{\|X\|_{g,r*}\leq 1} \langle X,Y\rangle = \max_{X\in B^1_{g,r*}} \langle X,Y\rangle.$$

Since $\operatorname{conv}(E_{g,r})$ and $B^1_{g,r*}$ are closed convex sets, this equality can only be fulfilled if the sets are equal (see [28, Theorem V.3.3.1]).

Next, we prove the decomposition. Since the decomposition trivially holds for $M = 0$, we assume that $M \in \mathbb{R}^{n\times m} \setminus \{0\}$ and define $\bar{M} := \|M\|^{-1}_{g,r*} M$. Then $\bar{M} \in B^1_{g,r*} = \operatorname{conv}(E)$ and therefore be decomposed as

$$\bar{M} = \textstyle\sum_i \alpha_i \tilde{M}_i \quad \text{with} \quad \sum_i \alpha_i = 1,\ \alpha_i \geq 0$$

where all $\tilde{M}_i$ satisfy

$$\|\tilde{M}_i\|_g = \|\tilde{M}_i\|_{g,r*} = 1 \quad \text{and} \quad \operatorname{rank}(\tilde{M}_i) \leq r,$$

where the first equality is from (12) in Lemma 2. Defining $M_i := \tilde{M}_i \|M\|_{g,r*}$ gives

$$M = \textstyle\sum_i \alpha_i M_i \quad \text{with} \quad \operatorname{rank}(M_i) \leq r$$

and

$$\|M_i\|_g = \|M_i\|_{g,r*} = \|\|M\|_{g,r*}\tilde{M}_i\|_{g,r*} = \|M\|_{g,r*}\|\tilde{M}_i\|_{g,r*} = \|M\|_{g,r*}.$$

This concludes the proof. □

A.1.4 Proof to Proposition 3

Proof. Let $\bar{M} = \sum_i \alpha_i M_i$ with $M_i \in E_{g,r}$ and $\alpha_i \in (0,1)$, $\sum_i \alpha_i = 1$ be a convex combination of points in $E_{g,r}$. Then, by assumption,

$$\|\bar{M}\|_g = \|\textstyle\sum_i \alpha_i M_i\|_g < \sum_i \alpha_i \|M_i\|_g = \sum_i \alpha_i = 1$$

and thus $\bar{M} \notin E_{g,r}$. Since $\operatorname{conv}(E_{g,r}) = B^1_{g,r*}$, this implies that $E_{g,r}$ is the set of extreme points of $B^1_{g,r*}$. □

A.1.5 Proof to Corollary 1

Proof. Let us start by showing that $\operatorname{conv}(\mathcal{E}_r) = B^1_{\ell_\infty,r*}$. Since $\|\cdot\|_{\ell_1,r}$ and $\|\cdot\|_{\ell_\infty,r}$ are dual norms to each other, it follows by Lemma 3 that

$$\|Y\|_{\ell_1,r} = \max_{X \in B^1_{\ell_\infty,r*}} \langle X, Y \rangle = \max_{\substack{\operatorname{rank}(X)=r \\ 1=\sigma_1(X)=\ldots=\sigma_r(X)}} \sum_{i=1}^{r} \sigma_i(X)\sigma_i(Y) = \max_{X \in \operatorname{conv}(\mathcal{E}_r)} \langle X, Y \rangle,$$

where the last two equalities are a result of [29, Corollary 7.4.1.3(c)]. However, $\operatorname{conv}(\mathcal{E}_r)$ and $B^1_{\ell_\infty,r*}$ are closed convex sets and therefore this equation can only hold if the sets are identical (see [28, Proposition V.3.3.1]).

It remains to show that no point in $\mathcal{E}_r$ can be constructed as a convex combination of other points in $\mathcal{E}_r$. To this end, note that a necessary condition for $M \in \mathcal{E}_r$ is that

$$\|M\|_F^2 = \sum_{i=1}^{\min\{m,n\}} \sigma_i^2(M) = \sum_{i=1}^{r} \sigma_i^2(M) = r.$$

Let $\bar{M} = \sum_i \alpha_i M_i$ be an arbitrary convex combination with $\alpha_i > 0$ and $\sum_i \alpha_i = 1$, of distinct points $M_i \in \mathcal{E}_r$. By the strict convexity of $\|\cdot\|_F^2$, it holds that

$$\|\bar{M}\|_F^2 = \|\textstyle\sum_i \alpha_i M_i\|_F^2 < \sum_i \alpha_i \|M_i\|_F^2 = r \sum_i \alpha_i = r.$$

Hence, $\bar{M} \notin \mathcal{E}_r$ and this concludes the proof. □

A.2 Derivations to $\Pi_{\text{epi}(\|\cdot\|_{g,r*})}$

Utilizing the Moreau decomposition in (23), we determine the projection onto $\text{epi}(\|\cdot\|_{g,r*})$, by computing a projecting onto the polar cone $(\text{epi}(\|\cdot\|_{g,r*}))^\circ$. The latter is by definition (see [5, Definition 6.21]) the negative of the dual cone to $\text{epi}(\|\cdot\|_{g,r*})$, i.e.

$$\begin{aligned}(\text{epi}(\|\cdot\|_{g,r*}))^\circ &= -\text{epi}(\|\cdot\|_{g^D,r}) \\ &= \{(-Y,-w) : \|Y\|_{g^D,r} \le w\} = \{(Y,w) : \|Y\|_{g^D,r} \le -w\}.\end{aligned}$$

Thus, the projection onto the polar cone becomes

$$\Pi_{(\text{epi}(\|\cdot\|_{g,r*}))^\circ}(Z,z_v) = \underset{\substack{w\in\mathbb{R},\ Y\in\mathbb{R}^{n\times m} \\ w+\|Y\|_{g^D,r}\le 0}}{\text{argmin}} \frac{1}{2}\left[(w-z_v)^2 + \|Y-Z\|_F^2\right]$$

and we need to solve

$$\begin{aligned}\underset{Y,w}{\text{minimize}} \quad & \frac{1}{2}\left[(w-z_v)^2 + \|Y-Z\|_F^2\right] \\ \text{subject to} \quad & -w \ge \|Y\|_{g^D,r},\ Y\in\mathbb{R}^{n\times m}.\end{aligned} \tag{39}$$

Since the cost and the constraint in (39) are unitarily invariant, it can be shown (see [32, 46]) that $Y^\star$ and Z have a simultaneous SVD, i.e. if $Z = \sum_{i=1}^q \sigma_i(Z) u_i v_i^T$ is an SVD of Z then $Y^\star = \sum_{i=1}^q \sigma_i(Y^\star) u_i v_i^T$ where we define $q := \min\{m,n\}$. Consequently, it is equivalent to consider the vector-valued problem

$$\begin{aligned}\underset{y,w}{\text{minimize}} \quad & \frac{1}{2}\left[(w-z_v)^2 + \sum_{i=1}^q (y_i - z_i)^2\right] \\ \text{subject to} \quad & -w \ge \|\text{diag}(y)\|_{g^D,r},\ y\in\mathbb{R}^q, \\ & y_1 \ge \cdots \ge y_q,\end{aligned} \tag{40}$$

with $z_1 \ge \cdots \ge z_q \ge 0$, $z_i = \sigma_i(Z)$ and $y_i = \sigma_i(Y)$ for $1 \le i \le q$.

Remark 2 *The unique solution $(y^\star, w^\star)$ fulfills $0 \le y_i^\star \le z_i$ for $1 \le i \le q$. The upper bound holds, because otherwise $\bar{y}^\star$ with $\bar{y}_i^\star = \min\{z_i, y_i^\star\}$ is a feasible solution with smaller cost. Similarly, the lower bound holds, because otherwise $\bar{y}^\star$ with $\bar{y}_i^\star = \max\{0, y_i^\star\}$ is a feasible solution with smaller cost. Thus, it is not necessary to explicitly restrict y to be nonnegative.*

To solve (40), note that there exists a $t^\star \in \{1,\dots,r\}$ such that

$$y_{r-t^\star}^\star > y_{r-t^\star+1}^\star = \cdots = y_r^\star, \tag{41}$$

where $t^\star = r$ if $y_1^\star = y_r^\star$. This assumption implies that $y_{r-t^\star} \ge y_{r-t^\star+1}$ is assumed to be inactive and therefore can be removed from (40). Then also the constraints $y_1 \ge \cdots \ge$

$y_{r-t^\star}$ can be removed, because the cost function and the sorting of z ensures that the solution will always fulfill them. This yields the following problem

$$\begin{aligned} \underset{y,w}{\text{minimize}} \quad & \frac{1}{2}\left[(w-z_v)^2+\sum_{i=1}^{q}(y_i-z_i)^2\right] \\ \text{subject to} \quad & -w \geq \|\operatorname{diag}(y)\|_{g^D,r},\ y\in\mathbb{R}^q, \\ & y_{r-t+1}=\cdots=y_r\geq\cdots\geq y_q. \end{aligned} \tag{42}$$

Thus, solving (40) reduces to finding $t^\star$ such that (42) solves (40). As it is shown later, solving (42) can be done efficiently for the low-rank inducing norms that are considered in this paper. The following lemma shows that $t^\star$ can be found by a binary search over t, where the decision to increase or decrease t is based on the solution of (42).

Lemma A.1 *Let* $\left(y^{(t)},w^{(t)}\right)$ *denote the solution to* (42) *depending on* t *such that* $1\leq t\leq r$. *Further let* $\left(y^{(t^\star)},w^{(t^\star)}\right)$ *be the solution to* (40) *such that* $y^{(t^\star)}_{r-t^\star}>y^{(t^\star)}_{r-t^\star+1}$ *and* $y^{(t^\star)}_{r-t^\star}=y^{(t^\star)}_{r-t^\star+1}$ *if* $t^\star=r$. *Then,*

i. $t^\star=\min\{\{t:y^{(t)}_{r-t}>y^{(t)}_{r-t+1}\}\cup\{r\}\}$.

ii. *If* $y^{(t')}_{r-t'}\geq y^{(t')}_{r-t'+1}$ *then* $y^{(t)}_{r-t}\geq y^{(t)}_{r-t+1}$ *for all* $t\geq t'$.

iii. *If* $y^{(t')}_{r-t'}< y^{(t')}_{r-t'+1}$ *then* $y^{(t)}_{r-t}< y^{(t)}_{r-t+1}$ *for all* $t\leq t'$.

In particular,

I. $y^{(t)}_{r-t}\geq y^{(t)}_{r-t+1}$ *for all* $t\geq t^\star$.

II. $y^{(t)}_{r-t}\leq y^{(t)}_{r-t+1}$ *for all* $t< t^\star$

III. *If* $t<t^\star$ *and* $y^{(t)}_{r-t}\leq y^{(t)}_{r-t+1}$ *then* $\left(y^{(t)},w^{(t)}\right)=\left(y^{(t^\star)},w^{(t^\star)}\right)$.

Proof. Throughout this proof, we let $p(t)$ denote the optimal cost of (42) as a function of t. Since adding constraints cannot reduce the optimal cost, p is a nondecreasing function.

Item i: By the same reasoning that led to (42), it holds that

$$y^{(t)}_1\geq\cdots\geq y^{(t)}_{r-t} \text{ for } 1\leq t\leq r. \tag{43}$$

Using (43), the set $\min\{\{t:y^{(t)}_{r-t}>y^{(t)}_{r-t+1}\}\cup\{r\}\}$ contains all t for which the solution of (42) is feasible for (40). Since p is nondecreasing and $\left(y^{(t^\star)},w^{(t^\star)}\right)$ is unique, the first claim follows.

Item ii: The second claim is proven by contradiction. Let $(y^{(t')},w^{(t')})$ be such that

$y^{(t')}_{r-t'} \geq y^{(t')}_{r-t'+1}$. Further assume that $y^{(t'+1)}_{r-t'-1} < y^{(t'+1)}_{r-t'}$. In the following, we construct another solution $(\tilde{y}, \tilde{w}) \in \mathbb{R}^{q+1}$ to (42) with $t = t'+1$, which has a cost that is no larger than $p(t'+1)$. However, (42) has a unique solution due to strong convexity of the cost function. This yields the desired contradiction.

The contradicting solution is constructed as a convex combination $\tilde{w} = (1-\alpha)w^{(t'+1)} + \alpha w^{(t')}$ with $\alpha \in (0,1]$ and a partially sorted convex combination of $y^{(t')}$ and $y^{(t'+1)}$ with the same α. Let $\hat{y} := (1-\alpha)y^{(t'+1)} + \alpha y^{(t')}$ and let

$$\tilde{y} := (\mathrm{sort}(\hat{y}_1, \ldots, \hat{y}_{r-t'-2}, \hat{y}_{r-t'}), \hat{y}_{r-t'-1}, \hat{y}_{r-t'+1}, \ldots, \hat{y}_q),$$

be the partially sorted convex combination, where $\mathrm{sort}(\cdot)$ denotes sorting in descending order.

To select α, we note that by assumption,

$$y^{(t')}_{r-t'-1} \geq y^{(t')}_{r-t'} \geq y^{(t')}_{r-t'+1} \quad \text{and} \quad y^{(t'+1)}_{r-t'-1} < y^{(t'+1)}_{r-t'} = y^{(t'+1)}_{r-t'+1}.$$

Therefore, there exists an $\alpha \in (0,1]$ such that

$$\begin{aligned}\tilde{y}_{r-t'} = \hat{y}_{r-t'-1} &= (1-\alpha)y^{(t'+1)}_{r-t'-1} + \alpha y^{(t')}_{r-t'-1} \\ &= (1-\alpha)y^{(t'+1)}_{r-t'+1} + \alpha y^{(t')}_{r-t'+1} = \hat{y}_{r-t'+1} = \tilde{y}_{r-t'+1}.\end{aligned}$$

Since

$$y^{(t')}_{r-t'+1} = \cdots = y^{(t')}_r \quad \text{and} \quad y^{(t'+1)}_{r-t'-1} = \cdots = y^{(t'+1)}_r,$$

it follows that

$$\tilde{y}_{r-t'} = \cdots = \tilde{y}_r.$$

Furthermore, the construction of $\tilde{y}$ as well as the sorting give that

$$\tilde{y}_r \geq \cdots \geq \tilde{y}_q \quad \text{and} \quad \tilde{y}_1 \geq \cdots \geq \tilde{y}_{r-t'-1}.$$

Hence, $\tilde{y}$ satisfies the chain of inequalities in (42) for $t = t'+1$.

It remains to show that $\tilde{y}$ satisfies the epigraph constraint and that the cost is not higher than $p(t'+1)$. These properties are already fulfilled for $\hat{y}$ being a convex combination of two feasible points with costs $p(t')$ and $p(t'+1)$, respectively, where $p(t') \leq p(t'+1)$. Therefore, it is left to show that the sorting involved in $\tilde{y}$ maintains these properties. First, we show that sorting of any sub-vector in y does not increase the cost. Suppose that $z_i \geq z_j$, $y_i \leq y_j$, i.e., y is not sorted the same way as z. Then

$$\begin{aligned}\tfrac{1}{2}\left((z_i - y_i)^2 + (z_j - y_j)^2\right) &= (z_i - z_j)(y_j - y_i) + \tfrac{1}{2}\left((z_i - y_j)^2 + (z_j - y_i)^2\right) \\ &\geq \left((z_i - y_j)^2 + (z_j - y_i)^2\right),\end{aligned}$$

and thus the cost is not increased by sorting y or any sub-vector of it. Further, note that the permutation caused by the sorting of the first r elements of y does not influence the epigraph constraint, because $\|\mathrm{diag}(y)\|_{g^D,r}$ is permutation invariant by definition.

Next notice that $\tilde{y}$ is obtained from $\hat{y}$ by first swapping $\hat{y}_{r-t'-1}$ and $\hat{y}_{r-t'}$. From the choice of α, we conclude that

$$\hat{y}_{r-t'} = (1-\alpha)y^{(t'+1)}_{r-t'} + \alpha y^{(t')}_{r-t'} \geq (1-\alpha)y^{(t'+1)}_{r-t'+1} + \alpha y^{(t')}_{r-t'+1} = \hat{y}_{r-t'+1} = \hat{y}_{r-t'-1}.$$

Thus, this swap is a sorting which does neither increase the cost, nor does it violate the epigraph constraint. Analogously, sorting the first $r-t'$ elements of the resulting vector to obtain $\tilde{y}$ has the same effect and therefore we receive the desired contradiction.

Item iii: Suppose that there exist t and t' with $t' > t$ such that $y^{(t')}_{r-t'} < y^{(t')}_{r-t'+1}$ and $y^{(t)}_{r-t} \geq y^{(t)}_{r-t+1}$. Then Item ii shows that $y^{(t')}_{r-t'} \geq y^{(t')}_{r-t'+1}$, which is a contradiction.

Items I to III: The statements follow immediately from Items i to iii. □

In order to solve (42), one can proceed similarly to solving (40). There always exists $s^\star \geq 0$ such that the solution $(y^{(t)}, w^{(t)})$ of (42) satisfies

$$y^{(t)}_{r-t+1} = \cdots = y^{(t)}_{r+s^\star} > y^{(t)}_{r+s^\star+1},$$

where $s^\star = q-r$ if $y^{(t)}_r = y^{(t)}_q$. As before, this allows us to remove the inactive constraint $y_{r+s} \geq y_{r+s+1}$. Then the constraints $y_{r+s+1} \geq \cdots \geq y_q$ become redundant, because any solution fulfills $y_j = z_j,\ j \geq r+s+1$. Finally, we are left with the following reduced optimization problem

$$\begin{aligned} \underset{y,w}{\text{minimize}} \quad & \frac{1}{2}\left[(w-z_v)^2 + \sum_{i=1}^{r+s}(y_i - z_i)^2\right] \\ \text{subject to} \quad & -w \geq \|\operatorname{diag}(y)\|_{g^D,r},\ y \in \mathbb{R}^q, \\ & y_{r-t+1} = \cdots = y_{r+s}. \end{aligned} \tag{44}$$

For given t, one can perform a binary search on s in (44) in order find $s^\star$. This can be done with the help of the following lemma.

Lemma A.2 *For fixed t with $1 \leq t \leq r$, let $\left(y^{(t,s)}, w^{(t,s)}\right)$ denote the solution to (44) for different s satisfying $0 \leq s \leq r-q$. Further let $\left(y^{(t,s^\star)}, w^{(t,s^\star)}\right)$ be the solution to (42) such that $y^{(t,s^\star)}_{r+s^\star} > y^{(t,s^\star)}_{r+s^\star+1}$ and $y^{(t,s^\star)}_{r+s^\star} = y^{(t)}_{r+s^\star+1}$ if $s^\star = q-r$. Then,*

i. $s^\star = \min\{\{s : y^{(t,s^\star)}_{r+s^\star} > y^{(t,s^\star)}_{r+s^\star}\} \cup \{q-r\}\}$.

ii. *If $y^{(t,s')}_{r+s'} \geq y^{(t,s')}_{r+s'+1}$ then $y^{(t,s)}_{r+s} \geq y^{(t,s)}_{r+s+1}$ for all $s \geq s'$.*

iii. *If $y^{(t,s')}_{r+s'} < y^{(t,s')}_{r+s'+1}$ then $y^{(t,s)}_{r+s} < y^{(t,s)}_{r+s+1}$ for all $s \leq s'$.*

In particular,

I. $y^{(t,s)}_{r+s} \geq y^{(t,s)}_{r+s+1}$ *for all* $s \geq s^\star$.

II. $y^{(t,s)}_{r+s} \leq y^{(t,s)}_{r+s+1}$ *for all* $s < s^\star$.

III. *If $s < s^\star$ and $y^{(t,s)}_{r+s} \geq y^{(t,s)}_{r+s+1}$ then $\left(y^{(t,s)}, w^{(t)}\right) = \left(y^{(t,s^\star)}, w^{(t,s^\star)}\right)$.*

Algorithm A.1 Determine $(Y^\star, w^\star) = \Pi_{(\mathrm{epi}(\|\cdot\|_{g,r*}))^\circ}(Z, z_v)$, i.e., solve (39)

Input: Let $Z \in \mathbb{R}^{n\times m}$, $z_v \in \mathbb{R}$ and $r \in \mathbb{N}$ such that $1 \le r \le q := \min\{m,n\}$ be given.

Let $Z = \sum_{i=1}^{q} \sigma_i(Z) u_i v_i^T$ be an SVD of Z.
// Solve (39) via the vector problem (40) with data $z = (\sigma_1(Z), \ldots, \sigma_q(Z))$ and z_v
Set $t_{\min} = 1$, $t_{\max} = r$, and $t = \lceil \frac{t_{\min}+t_{\max}}{2} \rceil$
// Solve (40) via (42) and binary search over t
while $t_{\min} \neq t_{\max}$ **do**
  Set $s_{\min} = 0$, $s_{\max} = q - r$, and $s = \lceil \frac{s_{\min}+s_{\max}}{2} \rceil$
  // Solve (42) via (44) and binary search over s
  while $s_{\min} \neq s_{\max}$ **do**
    Solve (44)
    Update $s_{\min}$, $s_{\max}$, and s using the binary search rules in Lemma A.2
  end while
  Update $t_{\min}$, $t_{\max}$, and t using the binary search rules in Lemma A.1
end while
Output: $(Y^\star, w^\star) = (\sum_{i=1}^{q} y_i u_i v_i^T, w)$ with (y, w) being the last solution to (44).

Proof. The proof goes analogously to the proof of Lemma A.1 and is therefore omitted. □

The nested binary search algorithm to solve (39) via (40) is summarized in Algorithm A.1. The problem that decides how to update the parameters in the nested binary search is (44). In order to solve (44) explicitly, we introduce new variables $\tilde{y}, \tilde{z} \in \mathbb{R}^{r-t+1}$ as

$$\tilde{y}_i = \begin{cases} y_i, \text{ if } 1 \le i \le r-t \\ \sqrt{t+s}\, y_r, \text{ if } i = r-t+1 \end{cases} \qquad \tilde{z}_i = \begin{cases} z_i, \text{ if } 1 \le i \le r-t \\ \frac{1}{\sqrt{t+s}} \sum_{i=r-t+1}^{r+s} z_i, \text{ if } i = r-t+1 \end{cases} \tag{45}$$

This gives

$$\sum_{i=r-t+1}^{r+s} (y_r - z_i)^2 = (\tilde{y}_{r-t+1} - \tilde{z}_{r-t+1})^2 + \sum_{i=r-t+1}^{r+s} z_i^2 - \left(\frac{1}{\sqrt{t+s}} \sum_{i=r-t+1}^{r+s} z_i \right)^2.$$

Since we can ignore the constant terms, we are left with the following projection problem of reduced dimension

$$\begin{aligned} \underset{\tilde{y}, w}{\text{minimize}} \quad & \frac{1}{2} \left[(w - z_v)^2 + \sum_{i=1}^{r-t+1} (\tilde{y}_i - \tilde{z}_i)^2 \right] \\ \text{subject to} \quad & -w \ge \|\mathrm{diag}(\tilde{y}_1, \ldots, \tilde{y}_{r-t}, \underbrace{\tfrac{\tilde{y}_{r-t+1}}{\sqrt{s+t}}, \ldots, \tfrac{\tilde{y}_{r-t+1}}{\sqrt{s+t}}}_{t \text{ times}})\|_{g^D,r}, \ \tilde{y} \in \mathbb{R}^{r-t+1}. \end{aligned}$$

Below, it is shown how to explicitly solve this projection problem for $g^D = \ell_2$ and $g^D = \ell_1$ in order to arrive at the epigraph projections of the low-rank inducing Frobenius and spectral norms.

A.2.1 The case $\|\cdot\|_{g^D,r} = \|\cdot\|_r$

In this case, $g^D = \ell_2$ and the projection problem becomes

$$\begin{aligned} \underset{\tilde{y},w}{\text{minimize}} \quad & \frac{1}{2}\left[(w-z_v)^2 + \sum_{i=1}^{r-t+1}(\tilde{y}_i - \tilde{z}_i)^2\right] \\ \text{subject to} \quad & -w \geq \sqrt{\sum_{i=1}^{r-t}\tilde{y}_i^2 + \frac{t}{s+t}\tilde{y}_{r-t+1}^2},\ y \in \mathbb{R}^{r-t+1}. \end{aligned}$$

Consequently, the solution $(\tilde{y}^\star, w^\star)$ is the orthogonal projection of $(\tilde{z}, z_v)$ onto the second-order cone

$$K := \left\{(\tilde{y}, w) \in \mathbb{R}^{r-t+2} : \sqrt{\sum_{i=1}^{r-t}\tilde{y}_i^2 + \frac{t}{s+t}\tilde{y}_{r-t+1}^2} \leq -w\right\}. \tag{46}$$

The associated polar cone $K^\circ := \{y : \langle y, x\rangle \leq 0 \text{ for all } x \in K\}$ is then given by (see e.g. [21])

$$K^\circ := \left\{(y, p) \in \mathbb{R}^{r-t+2} : \sqrt{\sum_{i=1}^{r-t}\tilde{y}_i^2 + \frac{s+t}{t}\tilde{y}_{r-t+1}^2} \leq p\right\}.$$

This allows us to summarize the following two simple cases:

i. $(\tilde{y}^\star, w^\star) = (\tilde{z}, z_v)$ if and only if $(\tilde{z}, z_v) \in K$, i.e.

$$\sqrt{\sum_{i=1}^{r-t}\tilde{z}_i^2 + \frac{t}{s+t}\tilde{z}_{r-t+1}^2} \leq -z_v,$$

ii. $(\tilde{y}^\star, w^\star) = (0, 0)$ if and only if $(\tilde{z}, z_v) \in K^\circ$, i.e.

$$\sqrt{\sum_{i=1}^{r-t}\tilde{z}_i^2 + \frac{s+t}{t}\tilde{z}_{r-t+1}^2} \leq z_v,$$

where the last statement follows by [28, Proposition III.3.2.3].

Next, it is shown how to compute the projection if $(\tilde{z}, z_v)$ does not belong to either of these cones. By [5, Proposition 6.46] it holds that $(\tilde{z} - \tilde{y}^\star, z_v - w^\star)$ is an element of the normal cone to the cone K at $(\tilde{y}^\star, w^\star)$. Using the normal cone description in [28, Theorem VI.1.3.5], this implies that

$$(\tilde{z} - \tilde{y}^\star, z_v - w^\star) = \mu\, \nabla_{(\tilde{y},w)}\sqrt{\sum_{i=1}^{r-t}\tilde{y}_i^2 + \frac{t}{s+t}\tilde{y}_{r-t+1}^2} + w\Bigg|_{(\tilde{y},w)=(\tilde{y}^\star,w^\star)} \tag{47}$$

for some $\mu \geq 0$. Since $(\tilde{z}, z_v) \notin K$ we conclude that the optimal point is on the boundary of the cone K, i.e.

$$-w^\star = \sqrt{\sum_{i=1}^{r-t}\tilde{y}_i^{\star 2} + \frac{t}{s+t}\tilde{y}_{r-t+1}^{\star 2}}. \tag{48}$$

Solving the equations in (47) and using (48) give

$$\begin{aligned}\tilde{y}_i^\star &= \frac{\tilde{z}_i}{1-\frac{\mu}{w^\star}},\ 1\le i\le r-t,\\ \tilde{y}_{r-t+1}^\star &= \frac{\tilde{z}_{r-t+1}}{1-\frac{\mu t}{w^\star(s+t)}},\\ w^\star &= z_v-\mu.\end{aligned}$$

To characterize the solution, it is left to compute μ. By plugging the solution into (48), diving by $w^\star$ and taking the square, we arrive at

$$1=\frac{\sum_{i=1}^{r-t}\tilde{z}_i^2}{(2\mu-z_v)^2}+\frac{t}{s+t}\frac{\tilde{z}_{r-t+1}^2}{\left(\mu-z_v+\frac{\mu t}{s+t}\right)^2}.$$

Defining $c_1:=\sum_{i=1}^{r-t}\tilde{z}_i^2=\sum_{i=1}^{r-t}z_i^2$ and $c_2:=\sqrt{t+s}\tilde{z}_{r-t+1}=\sum_{i=r-t+1}^{r+s}z_i$ this can rewritten as the fourth order polynomial equation

$$[(2\mu-z_v)^2-c_1][(t+s)(\mu-z_v)+\mu t]^2-tc_2^2(2\mu-z_v)^2=0, \tag{49}$$

which can be solved explicitly for $\mu\ge 0$. Resubstitution in (45) gives that the solution $\left(y^{(t,s)},w^{(t,s)}\right)$ to (44) can be expressed as

i. $1\le j\le r-t: y_j^{(t,s)}=\dfrac{z_i(\mu-z_v)}{2\mu-z_v}$,

ii. $r-t+1\le j\le r+s: y_j^{(t,s)}=\dfrac{(\mu-z_v)\sum_{i=r-t+1}^{r+s}z_i}{(s+t)(\mu-z_v)+\mu t}$,

iii. $r+s+1\le j\le q: y_j^{(t,s)}=z_j$,

iv. $w^{(t,s)}=z_v-\mu$,

if $(z,z_v)\notin K\cup K^\circ$.

A.2.2 The case $\|\cdot\|_{g^D,r}=\|\cdot\|_{\ell_1,r}$

The second case is analog to the first case. We would like to solve

$$\begin{aligned}\underset{\tilde{y},w}{\text{minimize}}\quad & \frac{1}{2}\left[(w-z_v)^2+\sum_{i=1}^{r-t+1}(\tilde{y}_i-\tilde{z}_i)^2\right]\\ \text{subject to}\quad & 0\ge\sum_{i=1}^{r-t}|\tilde{y}_i|+\frac{t}{\sqrt{t+s}}|\tilde{y}_{r-t+1}|+w,\ y\in\mathbb{R}^{r-t+1}.\end{aligned} \tag{50}$$

Consequently, the solution $(\tilde{y}^\star,w^\star)$ is the orthogonal projection of $(\tilde{z},z_v)$ onto

$$K:=\left\{(\tilde{y},w)\in\mathbb{R}^{r-t+2}:\sum_{i=1}^{r-t}|\tilde{y}_i|+\frac{t}{\sqrt{t+s}}|\tilde{y}_{r-t+1}|\le -w\right\}. \tag{51}$$

The polar cone $K^\circ := \{y : \langle y, x\rangle \leq 0 \text{ for all } x \in K\}$ is then given by

$$K^\circ := \left\{ (y,p) \in \mathbb{R}^{r-t+2} : \max\left(|y_1|, \ldots, |y_{r-t-2}|, \frac{\sqrt{t+s}}{t}|y_{r-t+1}| \right) \leq p \right\}.$$

Similarly to before, we get the following two simple cases:

i. $(\tilde{y}^\star, w^\star) = (\tilde{z}, z_v)$ if and only if $(\tilde{z}, z_v) \in K$, i.e.

$$\sum_{i=1}^{r-t} \tilde{z}_i + \frac{t}{\sqrt{t+s}} \tilde{z}_{r-t+1} \leq -z_v,$$

ii. $(\tilde{y}^\star, w^\star) = (0,0)$ if and only if $(\tilde{z}, z_v) \in K^\circ$, i.e.

$$\max\left(\tilde{z}_1, \frac{\sqrt{t+s}}{t} \tilde{z}_{r-t+1} \right) \leq z_v,$$

where it is used that the $\tilde{z}_i$ are nonnegative and decreasingly sorted.

It remains to show how to compute the projection if $(\tilde{z}, z_v)$ does not belong to either of these cones. By [5, Proposition 6.46] it holds that $(\tilde{z} - \tilde{y}^\star, z_v - w^\star)$ is an element of the normal cone to the cone K at $(\tilde{y}^\star, w^\star)$. Using the normal cone description in [28, Theorem VI.1.3.5], we get

$$(\tilde{z} - \tilde{y}^\star, z_v - w^\star) \in \mu \, \partial_{(\tilde{y},w)} \left(\sum_{i=1}^{r-t} |\tilde{y}_i| + \frac{t}{\sqrt{s+t}} |\tilde{y}_{r-t+1}| + w \right) \Bigg|_{(\tilde{y},w)=(\tilde{y}^\star, w^\star)} \tag{52}$$

for some $\mu \geq 0$. First note that any solution to (50) satisfies $\tilde{y}^\star \geq 0$. The optimality conditions for $y_i^\star = 0$ and $y_i^\star > 0$ become

$$\tilde{y}_i^\star = 0 \Leftrightarrow \tilde{z}_i \in [0, \mu], \qquad \tilde{y}_i^\star > 0 \Leftrightarrow \tilde{y}_i^\star = \tilde{z}_i - \mu$$

for all $i \in \{1, \ldots, r-t\}$. These equivalences also hold for $\tilde{y}_{r-t+1}$ with μ multiplied by $t/\sqrt{s+t}$. Therefore,

$$\begin{aligned} \tilde{y}_i^\star &= \max(\tilde{z}_i - \mu, 0), \ 1 \leq i \leq r-t, \\ \tilde{y}_{r-t+1}^\star &= \max\left(\tilde{z}_{r-t+1} - \frac{t\mu}{\sqrt{t+s}}, 0 \right), \\ w^\star &= z_v - \mu. \end{aligned}$$

In order to determine μ, notice that $(\tilde{y}^\star, w^\star)$ lies on the boundary of the cone K in (51), which implies

$$\begin{aligned} 0 &= \sum_{i=1}^{r-t} |\tilde{y}_i^\star| + \frac{t}{\sqrt{t+s}} |\tilde{y}_{r-t+1}^\star| + w^\star \\ &= \sum_{i=1}^{r-t} \max(\tilde{z}_i - \mu, 0) + \frac{t}{\sqrt{t+s}} \max(\tilde{z}_{r-t+1} - \tfrac{t\mu}{\sqrt{t+s}}, 0) + z_v - \mu. \end{aligned}$$

We denote the solution to this equation by $\mu^\star$, and solve it using a so-called *break point searching algorithm*, as it has been done for similar problems in [13, 15, 26]. To this end, let

$$\hat{z} = \left(\tilde{z}_1, \dots, \tilde{z}_j, \frac{t}{\sqrt{t+s}} \tilde{z}_{r-t+1}, \tilde{z}_{j+1}, \dots, \tilde{z}_{r-t} \right),$$

be the vector that sorts $\tilde{z}$ according to the break points of the max expressions, i.e., the index j satisfies $\tilde{z}_j > \frac{\sqrt{t+s}}{t} \tilde{z}_{r-t+1} \geq \tilde{z}_{j+1}$. Defining

$$\alpha = \left(1, \dots, 1, \frac{t^2}{t+s}, 1, \dots, 1 \right)$$

gives that $\mu^\star$ can be found by solving

$$\sum_{i=1}^{r-t+1} \max(\hat{z}_i - \alpha_i \mu, 0) + z_v - \mu = 0. \tag{53}$$

Assuming that we know an index $k = k^\star$ such that

$$\hat{z}_{k^\star+1} - \alpha_{k^\star+1} \mu^\star \leq 0 \qquad \text{and} \qquad \hat{z}_{k^\star} - \alpha_{k^\star} \mu^\star \geq 0, \tag{54}$$

then $\mu^\star$ can be determined from (53) as

$$\mu^\star = \frac{\sum_{i=1}^{k^\star} \hat{z}_i + z_v}{1 + \sum_{i=1}^{k^\star} \alpha_i}. \tag{55}$$

Thus, computing $\mu^\star$ reduces to searching for $k^\star \in \{1, \dots, r-t\}$ for which (55) satisfies (54). This can be done using a binary search, with rules from the following proposition.

Lemma A.3 *Let $\mu^\star$ be the solution to (53), let μ_k be the solution to*

$$\left(\sum_{i=1}^{r-t+1} \hat{z}_i - \alpha_i \mu \right) + z_v - \mu = 0, \qquad \text{i.e.,} \qquad \hat{\mu}_k = \frac{\sum_{i=1}^{k} \hat{z}_i + z_v}{1 + \sum_{i=1}^{k} \alpha_i}, \tag{56}$$

and let $k^\star$ be such that $\hat{\mu}_{k^\star} = \mu^\star$. Then,

i. $k^\star = \max(\{k : \hat{z}_k - \alpha_k \hat{\mu}_k \geq 0\})$.

ii. If $\hat{z}_k - \alpha_k \hat{\mu}_k \geq 0$, then $\hat{z}_i - \alpha_i \hat{\mu}_i \geq 0$ for all $i \in \{1, \dots, k\}$.

iii. If $\hat{z}_k - \alpha_k \hat{\mu}_k < 0$, then $\hat{z}_i - \alpha_i \hat{\mu}_i < 0$ for all $i \in \{k, \dots, r-t\}$.

In particular,

I. $\hat{z}_k - \alpha_k \hat{\mu}_k \geq 0$ *for all* $k \in \{1, \dots, k^\star\}$.

II. $\hat{z}_k - \alpha_k \hat{\mu}_k < 0$, *for all* $k \in \{k^\star + 1, \dots, r-t\}$.

Proof. We first show some results needed to prove Items i and ii. Let

$$g_k(\mu) := \sum_{i=1}^{k} \max(\hat{z}_i - \alpha_i \mu, 0) + z_v - \mu,$$

which is strictly decreasing in μ. Let μ_k be the unique solution to the equation

$$g_k(\mu) = 0.$$

For all $\mu \in \mathbb{R}$, we have

$$g_{k-1}(\mu) = g_k(\mu) - \max(\hat{z}_k - \alpha_k \mu, 0) \leq g_k(\mu).$$

Since all g_i are strictly decreasing in μ, we conclude the following facts:

a. $\mu_{k-1} \leq \mu_k$.

b. If $\hat{z}_k - \alpha_k \mu_k \leq 0$, then $g_{k-1}(\mu_k) = g_k(\mu_k) = 0$, hence $\mu_{k-1} = \mu_k$.

Because $\hat{z}$ is sorted according to break points, we conclude that if l and μ are such that $\hat{z}_l - \alpha_l \mu \geq 0$, then also $\hat{z}_i - \alpha_i \mu \geq 0$ for all $i \in \{1, \ldots, l\}$. Therefore, if μ is such that $\hat{z}_k - \alpha_k \mu \geq 0$, we get

$$\sum_{i=1}^{k} \max(\hat{z}_k - \alpha_k \mu, 0) + z_v - \mu = \left(\sum_{i=1}^{k} \hat{z}_k - \alpha_k \mu \right) + z_v - \mu.$$

Hence,

c. If $\hat{z}_k - \alpha_k \mu_k \geq 0$ or $\hat{z}_k - \alpha_k \hat{\mu}_k \geq 0$, then $\hat{\mu}_k = \mu_k$.

Item i: Using Items b and c, we conclude that

$$\hat{\mu}_{k^\star} = \mu_{k^\star} = \mu_{k^\star+1} = \mu_{r-t+1} = \mu^\star.$$

Item ii: Now, assume that $\hat{z}_k - \alpha_k \hat{\mu}_k \geq 0$. Then, by break point sorting, it holds that $\hat{z}_{k-1} - \alpha_{k-1} \hat{\mu}_k \geq 0$. Using Items a and c, we conclude that

$$0 \leq \hat{z}_{k-1} - \alpha_{k-1} \hat{\mu}_k = \hat{z}_{k-1} - \alpha_{k-1} \mu_k \leq \hat{z}_{k-1} - \alpha_{k-1} \mu_{k-1} = \hat{z}_{k-1} - \alpha_{k-1} \hat{\mu}_{k-1}.$$

Using induction proves the result.
Item iii: Assume, on the contrary, that k is such that $\hat{z}_k - \alpha_k \hat{\mu}_k < 0$ but that there exists $i \in \{k, \ldots, r-t\}$ such that $\hat{z}_i - \alpha_i \hat{\mu}_i \geq 0$. Then, by Item ii, $\hat{z}_k - \alpha_k \hat{\mu}_k \geq 0$ and we have reached the desired contradition.
Items I and II: Follow immediately from Items i to iii. □

Now, that we know how to compute the dual variable $\mu = \mu^\star$, we go back to the original variables in (45), to conclude that the solution $\left(y^{(t,s)}, w^{(t,s)} \right)$ to (44) can be expressed as

i. $1 \le j \le r-t : y_j^{(t,s)} = \max(z_j - \mu, 0)$,

ii. $r-t+1 \le j \le r+s : y_j^{(t,s)} = \dfrac{1}{\sqrt{t+s}} \max(\sum_{i=r-t+1}^{r+s} z_i - t\mu, 0)$,

iii. $r+s+1 \le j \le q : y_j^{(t,s)} = z_j$,

iv. $w^{(t,s)} = z_v - \mu$.

if $(z, z_v) \notin K \cup K^\circ$.

References

[1] A. ANTOULAS, *Approximation of Large-Scale Dynamical Systems*, SIAM, 2005.

[2] A. C. ANTOULAS, *On the Approximation of Hankel Matrices*, Vieweg+Teubner Verlag, 1997, pp. 17–22.

[3] A. ARGYRIOU, R. FOYGEL, AND N. SREBRO, *Sparse prediction with the k-support norm*, in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds., Curran Associates, Inc., 2012, pp. 1457–1465.

[4] F. BACH, R. JENATTON, J. MAIRAL, AND G. OBOZINSKI, *Optimization with sparsity-inducing penalties*, Foundations and Trends in Machine Learning, 4 (2012), pp. 1–106.

[5] H. H. BAUSCHKE AND P. L. COMBETTES, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer New York, 2011.

[6] E. J. CANDÈS AND Y. PLAN, *Matrix completion with noise*, Proceedings of the IEEE, 98 (2010), pp. 925–936.

[7] E. J. CANDÈS AND B. RECHT, *Exact matrix completion via convex optimization*, Foundations of Computational Mathematics, 9 (2009), p. 717.

[8] E. J. CANDÈS, J. ROMBERG, AND T. TAO, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Transactions on Information Theory, 52 (2006), pp. 489–509.

[9] V. CHANDRASEKARAN, B. RECHT, P. A. PARRILO, AND A. S. WILLSKY, *The convex geometry of linear inverse problems*, Foundations of Computational Mathematics, 12 (2012), pp. 805–849.

[10] V. CHANDRASEKARAN, S. SANGHAVI, P. A. PARRILO, AND A. S. WILLSKY, *Rank-sparsity incoherence for matrix decomposition*, SIAM Journal on Optimization, 21 (2011), pp. 572–596.

[11] Y. CHEN, M. R. JOVANOVIĆ, AND T. T. GEORGIOU, *State covariances and the matrix completion problem*, in 52nd IEEE Conference on Decision and Control, 2013, pp. 1702–1707.

[12] P. L. COMBETTES AND J.-C. PESQUET, *Proximal Splitting Methods in Signal Processing*, Springer New York, 2011, pp. 185–212.

[13] L. CONDAT, *Fast projection onto the simplex and the ℓ_1 ball*, Mathematical Programming, 158 (2016), pp. 575–585.

[14] X. V. DOAN AND S. VAVASIS, *Finding the largest low-rank clusters with Ky Fan 2-k-norm and ℓ_1-norm*, SIAM Journal on Optimization, 26 (2016), pp. 274–312.

[15] J. DUCHI, S. SHALEV-SHWARTZ, Y. SINGER, AND T. CHANDRA, *Efficient projections onto the L1-ball for learning in high dimensions*, in Proceedings of the 25th International Conference on Machine Learning, 25th International Conference on Machine Learning (ICML), 2008, pp. 272–279.

[16] A. ERIKSSON, T. T. PHAM, T.-J. CHIN, AND I. REID, *The k-support norm and convex envelopes of cardinality and rank*, in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3349–3357.

[17] M. FAZEL, *Matrix Rank Minimization with Applications*, PhD thesis, Stanford University, 2002.

[18] M. FAZEL, H. HINDI, AND S. P. BOYD, *A rank minimization heuristic with application to minimum order system approximation*, in Proceedings of the 2001 American Control Conference., vol. 6, 2001, pp. 4734–4739.

[19] T. T. GEORGIOU, *Spectral analysis based on the state covariance: the maximum entropy spectrum and linear fractional parametrization*, IEEE Transactions on Automatic Control, 47 (2002), pp. 1811–1823.

[20] T. T. GEORGIOU, *The structure of state covariances and its relation to the power spectrum of the input*, IEEE Transactions on Automatic Control, 47 (2002), pp. 1056–1066.

[21] C. GRUSSLER AND A. RANTZER, *Modified balanced truncation preserving ellipsoidal cone-invariance*, in 53rd IEEE Conference on Decision and Control (CDC), 2014, pp. 2365–2370.

[22] C. GRUSSLER AND A. RANTZER, *On optimal low-rank approximation of non-negative matrices*, in 54th IEEE Conference on Decision and Control (CDC), 2015, pp. 5278–5283.

[23] C. GRUSSLER, A. RANTZER, AND P. GISELSSON, *Low-rank optimization with convex constraints*, (2016), arXiv:1606.01793.

[24] C. GRUSSLER, A. ZARE, M. R. JOVANOVIC, AND A. RANTZER, *The use of the r∗ heuristic in covariance completion problems*, in 55th IEEE Conference on Decision and Control (CDC), 2016.

[25] T. HASTIE, R. TIBSHIRANI, AND M. WAINWRIGHT, *Statistical Learning with Sparsity: The Lasso and Generalizations*, CRC Press, 2015.

[26] M. HELD, P. WOLFE, AND H. P. CROWDER, *Validation of subgradient optimization*, Mathematical Programming, 6 (1974), pp. 62–88.

[27] J.-B. HIRIART-URRUTY AND C. LEMARÉCHAL, *Convex analysis and minimization algorithms II*, Grundlehren der mathematischen Wissenschaften, Springer Berlin Heidelberg, 1996.

[28] J.-B. HIRIART-URRUTY AND C. LEMARÉCHAL, *Convex analysis and minimization algorithms I: Fundamentals*, vol. 305 of Grundlehren der mathematischen Wissenschaften, Springer Berlin Heidelberg, 2013.

[29] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, 2 ed., 2012.

[30] A. J. IZENMAN, *Reduced-rank regression for the multivariate linear model*, Journal of Multivariate Analysis, 5 (1975), pp. 248 – 264.

[31] V. LARSSON AND C. OLSSON, *Convex low rank approximation*, International Journal of Computer Vision, 120 (2016), pp. 194–214.

[32] A. S. LEWIS, *The convex analysis of unitarily invariant matrix functions*, Journal of Convex Analysis, 2 (1995), pp. 173–183.

[33] F. LIN, M. R. JOVANOVIĆ, AND T. T. GEORGIOU, *An admm algorithm for matrix completion of partially known state covariances*, in 52nd IEEE Conference on Decision and Control, 2013, pp. 1684–1689.

[34] X. LIU, Z. WEN, AND Y. ZHANG, *Limited memory block Krylov subspace optimization for computing dominant singular value decompositions*, SIAM Journal on Scientific Computing, 35 (2013), pp. A1641–A1668.

[35] D. G. LUENBERGER, *Optimization by Vector Space Methods*, John Wiley & Sons, 1968.

[36] A. M. MCDONALD, M. PONTIL, AND D. STAMOS, *New Perspectives on k-Support and Cluster Norms*, (2015), arXiv:1512.08204.

[37] N. PARIKH AND S. BOYD, *Proximal algorithms*, Foundations and Trends in Optimization, 1 (2014), pp. 127–239.

[38] D. PEAUCELLE, D. HENRION, Y. LABIT, AND K. TAITZ, *Users guide for SEDUMI INTERFACE 1.04*, (2002). LAAS-CNRS, Toulouse.

[39] B. RECHT, M. FAZEL, AND P. A. PARRILO, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Review, 52 (2010), pp. 471–501.

[40] G. C. REINSEL AND R. VELU, *Multivariate Reduced-Rank Regression: Theory and Applications*, vol. 136 of Lecture Notes in Statistics, Springer New York, 1998.

[41] R. T. ROCKAFELLAR, *Convex Analysis*, no. 28, Princeton University Press, 1970.

[42] R. TIBSHIRANI, *Regression shrinkage and selection via the lasso*, Journal of the Royal Statistical Society, 58 (1996), pp. 267–288.

[43] K.-C. TOH, M. J. TODD, AND R. H. TÜTÜNCÜ, *SDPT3 – a MATLAB software package for semidefinite programming, version 1.3*, Optimization Methods and Software, 11 (1999), pp. 545–581.

[44] K. C. TOH, R. H. TUTUNCU, AND M. J. TODD, *On the implementation of SDPT3 (version 3.1) – a MATLAB software package for semidefinite-quadratic-linear programming*, in 2004 IEEE International Conference on Robotics and Automation, 2004, pp. 290–296.

[45] R. VIDAL, Y. MA, AND S. S. SASTRY, *Generalized Principal Component Analysis*, vol. 40 of Interdisciplinary Applied Mathematics, Springer-Verlag New York, 2016.

[46] G. WATSON, *Characterization of the subdifferential of some matrix norms*, Linear Algebra and its Applications, 170 (1992), pp. 33 – 45.

[47] B. WU, C. DING, D. SUN, AND K.-C. TOH, *On the Moreau–Yosida regularization of the vector k-norm related functions*, SIAM Journal on Optimization, 24 (2014), pp. 766–794.

[48] A. ZARE, Y. CHEN, M. JOVANOVIĆ, AND T. T. GEORGIOU, *Low-complexity modeling of partially available second-order statistics: Theory and an efficient matrix completion algorithm*, IEEE Transactions on Automatic Control, (2016), doi:10.1109/TAC.2016.2595761. arXiv:1412.3399.

[49] A. ZARE, M. R. JOVANOVIĆ, AND T. T. GEORGIOU, *Alternating direction optimization algorithms for covariance completion problems*, in 2015 American Control Conference (ACC), 2015, pp. 515–520.

[50] A. ZARE, M. R. JOVANOVIĆ, AND T. T. GEORGIOU, *Color of turbulence*, Journal of Fluid Mechanics, (2016), doi:10.1017/jfm.2016.682. arXiv:1602.05105.