Fifty-fourth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 27 - 30, 2016

Minimization of a Particular Singular Value

Alborz Alavian and Michael Rotkowitz

Abstract— **We consider the problem of minimizing a particular singular value of a matrix variable, which is then subject to some convex constraints. Convex heuristics for this problem are discussed, including some counter-intuitive results regarding which is best, which then provide upper bounds on the value of the problem. The use of polynomial optimization formulations is considered, particularly for obtaining lower bounds on the value of the problem.**

We show that the main problem can also be formulated as an optimization problem with a bilinear matrix inequality (BMI), and discuss the use of this formulation.

I. Introduction

In this paper we will discuss minimization of a particular singular value of a matrix. This problems becomes convex only when minimizing the largest singular value, or concave when maximizing the smallest one. However, in all the other cases the problem is neither convex nor concave.

When one wishes to obtain low-rank solutions, the convex heuristic of nuclear norm has been shown to be effective and even guaranteed to recover a low rank solution in some cases [1]–[3]. However, when we are trying to minimize a specific singular value the most common approach is to apply a non-smooth non-convex technique by using the subgradient of that singular value. In the case that all the singular values would be distinct, deriving a local subgradient would be a straightforward approach and resembles the subgradient of l_1-norm, however the general case requires a more detailed technical definition and derivation, for which we refer to [4], [5]. It has also been suggested that due to structural relation of the singular values, one can minimize the partial tail sum of singular values [6], which would also be a non-convex non-smooth problem.

In this work we consider minimization of a particular singular value (k-th singular value) subject to some constraints, for which we will provide algorithms for obtaining both upper bounds and lower bounds. When finding upper bounds we assume that the constraints are all convex, and will analyze a class of convex heuristics by taking a non-integer partial sum of the singular values from the greatest one up to a non-integer portion of $(k+1)$-th singular value. We will inspect this heuristic numerically and compare it against the conventional ones, both in presence and absence of low-rank solutions. It was widely observed that our counter intuitive convex heuristic for minimizing the k-th singular value would perform better in the absence of low-rank solutions. However, we will prove that if our convex heuristic recovers a low-rank solution, then the nuclear norm would also recover that solution, suggesting that if one is only concerned with the rank minimization, and not minimizing the k-th singular value even if it would ultimately be greater than zero, nuclear norm would be at least as good as our convex heuristic. We will also discuss using subgradient based methods to further improve the solution obtained from our convex heuristic.

The problem of minimizing the k-th singular value of a matrix is closely related to the problem of minimizing the k-th largest element of a vector, and we note that what we derive in this paper could be very similarly formulated for that case as well. Also, the same behavior regarding our convex heuristic has been seen when we apply it on the k-th largest element of a vector.

In order to obtain lower bounds on the k-th singular value, we will first derive a polynomial optimization problems so that the optimal value of one would be exactly equal to the optimal value of the considered singular, and the optimal value of another one would give a lower bound. We will then utilize Sum-of-Square (SOS) techniques to derive lower bounds on these polynomial programs, which would in turn result in lower bounds for the minimization problem that we are interested in. These different formulations would be comparable against each other in terms of tightness of the lower bound and the size of the corresponding optimization problem.

We will discuss how we can alternatively formulate the problem of minimizing a particular singular value subject to convex constraints by a Bilinear Matrix Inequality (BMI) that will be further subjected to an orthonormality condition on one of its variables.

One application of minimizing such a singular value is in control theory, where one would like to know how far a linear time invariant plant is from losing controllability or observability. When every actuator has access to the information from all the other sensors (centralized settings), this reduces to minimizing the smallest singular value of a matrix over a single variable in the complex plane. This special case has attracted much attention and is referred to as controllability radius. There are methods that try to find the global optimum of the controllability radius through suggesting polynomial time algorithms to check if the objective can become less than a selected value, which then could be used in conjunction with a bisection method to get as close as desired to the global optimal value [7], [8]. There has also been results

A. Alavian is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA, alavian@umd.edu.

M. C. Rotkowitz is with the Institute for Systems Research and the Department of Electrical and Computer Engineering, The University of Maryland, College Park, MD 20742 USA, mcrotk@umd.edu.

This material is based upon work supported by the National Science Foundation under Grant No. 1351674.

on finding lower bounds for this special case by formulating the problem as a polynomial optimization problem [9] and use SOS techniques to derive lower bounds on the global minimum. When each actuator has access to a specific set of sensor measurements (decentralized settings), the problem of how far we are from losing decentralized controllability becomes more complex [10] and can be transformed into minimization of a particular singular value (between the largest and smallest one) over a variable in the complex plane and a binary vector, for which some convex heuristics and an ADMM based approach has been suggested [11]. Lower bounds resembling the ones that will be given in this paper are also considered for this special case in [12].

The organization of this paper is as follows. We will review the sum-of-squares techniques for deriving lower bounds on polynomial optimization problems in Section II, and formulate the problem of minimizing a particular singular value in Section III. We will analyze and numerically inspect our convex heuristic along with an enhancement step using subgradient methods in Section IV. We formulate the polynomial optimization problem for obtaining lower bounds in Section V. In Section V-A we will provide a polynomial optimization problem that will exactly be equal to the singular value of consideration using factorization of the semidefinite matrices, however since it involves a large number of variables and constraints, we give an alternative approximate form via characterization of positive definite matrices using leading principle minors in Section V-B. This alternative form would involve fewer variables and constraints, but of higher degrees. We will provide another form by sampling from the non-convex constraint in the Courant-Fischer variational formulation of the singular values in Section V-C. Finally, we will provide an equivalent BMI in Section VI.

II. REVIEW

We will first give a brief overview of polynomial optimization problems, and then review some related results that we will use later in this paper. The materials in this section are mostly adopted from [13].

Definition 1 (Polynomial Optimization Problem): Given real-valued polynomials $p(x)$, $g_1(x), \cdots, g_r(x)$ all from $\mathbb{R}^n$ to $\mathbb{R}$, the following optimization problem is called a Polynomial Optimization (P.O.) problem:

$$p_K^* \triangleq \min_{x \in K} p(x), \tag{1}$$

with variable $x \in \mathbb{R}^n$, and where the set $K \subseteq \mathbb{R}^n$ is defined by polynomial inequalities as $K \triangleq \{x \in \mathbb{R}^n | g_i(x) \geq 0,$ for $i = 1, \cdots, r\}$.

Remark 2: The equality constraints $h(x) = 0$ can be expressed by two inequality constraints $h(x) \geq 0$, and $(-h(x)) \geq 0$.

Furthermore, define the sum of squares polynomials as:

Definition 3 (Sum-of-Squares): A real-valued polynomial $p(x) : \mathbb{R}^n \to \mathbb{R}$ is called ***Sum-of-Squares*** (SOS), if it can be written as:

$$p(x) = \sum_{i=1}^{\tilde{i}} (p_i(x))^2,$$

for some $\tilde{i} \in \mathbb{N}$, and where $p_i(x) : \mathbb{R}^n \to \mathbb{R}$ are all polynomials in x for $i = 1, \cdots, \tilde{i}$.

We would like to derive lower bounds on one particular instance of such P.O. problems. This could be achieved if the set K satisfies the following assumption:

Assumption 4 ([13, Assumption 4.1]): The set K is compact and there exists a real-valued polynomial $u(x) : \mathbb{R}^n \to \mathbb{R}$ such that the set $\{x \in \mathbb{R}^n | u(x) \geq 0\}$ is compact, and:

$$u(x) = u_0(x) + \sum_{k=1}^{r} g_i(x) u_i(x), \qquad \text{for all} \quad x \in \mathbb{R}^n,$$

where the polynomials $u_i(x)$ are all SOS for $i = 0, \cdots, r$.

Remark 5: One way to ensure that this assumption holds is that the variable x would be bounded, i.e., it would be known that the solution of (1) would lie in some bounded region $\|x\|_2^2 \leq a$. In this case, one can add an inequality constraint $g_{r+1}(x) \geq 0$, with $g_{r+1}(x) = a - \|x\|_2^2$, to the set K, and take $u_i(x) = 1$, if $i = r + 1$, and 0 otherwise.

It can then be proved that with this assumption one could obtain a sequence of finite dimensional Semidefinite Programs (SDP) that would converge to the optimum value p_K^* from below. This is stated in the following theorem:

Theorem 6 ([13, Theorem 4.2]): Let $p(x)$, p_K^*, and the set K be given as in Definition 1. Assume that K is a compact set that satisfies Assumption 4, then there exists a sequence of finite dimensional SDP of order N, denoted by $\mathbb{Q}_K^N$, that converges to the optimal value p_K^* from below, i.e.:

$$\inf \mathbb{Q}_K^N \uparrow p_K^*, \qquad \text{as} \quad N \to \infty.$$

III. PROBLEM FORMULATION

We will state the problem that we are going to consider in this paper in this section, and will discuss upper and lower bounds for this problem in later sections.

Given a matrix $X \in \mathbb{R}^{m \times n}$, convex functions $f_1(X), \cdots, f_{\bar{i}}(X)$, and affine functions $h_1(X), \cdots, h_{\bar{j}}(X)$ all from $\mathbb{R}^{m \times n}$ to $\mathbb{R}$, we are interested in the following optimization problem:

$$\begin{array}{lll} \text{minimize} & \sigma_k(X) & \\ \text{subject to} & f_i(X) \leq 0 & i = 1, \cdots, \bar{i} \\ & h_j(X) = 0 & j = 1, \cdots, \bar{j}, \end{array} \tag{2}$$

with variable $X \in \mathbb{R}^{m \times n}$, and where $\sigma_k(X)$ denotes the k-th largest singular value of the matrix X. Without loss of generality we assume that $m \geq n$ and thus $\sigma_{n+1}(X) = \cdots = \sigma_m(X) = 0$. We will hence focus on the non-trivial cases $1 \leq k \leq n$, for which $\sigma_1(X) \geq \cdots \geq \sigma_n(X)$. This problem is convex only if $k = 1$ and we are interested in cases where $k > 1$.

IV. A Convex Heuristic

We will analyze a class of convex heuristics for problem (2), and inspect their performance via numerical simulations later in this section. To this end, define the non-integer partial sum of singular values of X, denoted by $s_l(X)$ as:

$$s_l(X) \triangleq \sum_{i=1}^{\lfloor l\rfloor}\sigma_i(X) + (l-\lfloor l\rfloor)\sigma_{\lfloor l\rfloor+1}(X), \qquad (3)$$

where $l \in [1, n]$ is a real variable, and where $\lfloor l\rfloor$ denotes floor of l.

Example 7: If we take $l = 2.4$ then $s_{2.4}(X) = \sigma_1(X) + \sigma_2(X) + 0.4\,\sigma_3(X)$.

Remark 8: $s_l(X)$ is convex in X for all $l \in [1, n]$.

Remark 9: The nuclear norm of X is by definition equal to $s_n(X)$.

Problem (2) is non-convex in its variable for all $k > 1$ and we would like to replace the objective with the convex heuristic $s_l(X)$ and inspect the best l, i.e., we are interested in solving the following problem:

$$\begin{array}{lll} X_l^* \in & \arg\min\limits_{X\in\mathbb{R}^{m\times n}} & s_l(X) \\ & \text{subject to} & f_i(X) \le 0 \quad i = 1, \cdots, \bar{i} \\ & & h_j(X) = 0 \quad j = 1, \cdots, \bar{j}, \end{array} \qquad (4)$$

and then reporting $\sigma_k(X_l^*)$ as the output of our convex heuristic. Also denote the best l value for a specific singular value k by l_k^*, i.e.:

$$l_k^* \triangleq \arg\min_{l\in[1,n]} \sigma_k(X_l^*).$$

Remark 10: Perhaps the first consideration would be that taking $l = k$ would generally be the best convex heuristic for minimizing σ_k, however our understanding from a wide variety of numerical examples shows that generally we have $l_k^* > k$.

Example 11: As an example of (4) assume that we want to minimize the $\sigma_2(X)$, and that X_1 and X_2 both satisfy the feasibility constraints. Furthermore assume that the singular values of these two matrices are given as in Table I. If we

	σ_1	σ_2	σ_3	$\sigma_1+\sigma_2$	$\sigma_1+\sigma_2+\sigma_3$
X_1	10	9.9	9	19.9	28.9
X_2	12	8	7.9	20	27.9

TABLE I: Singular values for Example 11

only focus on $\sigma_1(X)+\sigma_2(X)$, i.e. $l = 2$ we would be worse off than taking the sum of the first three singular values. This could happen as the singular values are implicitly tied together via structural constraints $\sigma_1(X) \ge \cdots \ge \sigma_n(X)$. This would be further inspected via a variety of numerical examples in the rest of this paper.

For numerical examples through the rest of this section we will focus on the following optimization problem.

$$\begin{array}{lll} \text{minimize} & \sigma_k(X) & \\ \text{subject to} & X_{ij} = B_{ij} & \text{for} \quad (i,j) \in \mathcal{I} \\ & \underline{X}_{ij} \le X_{ij} \le \bar{X}_{ij} & \text{for} \quad (i,j) \in \bar{\mathcal{I}}, \end{array} \qquad (5)$$

with variable $X \in \mathbb{R}^{m\times n}$, and fixed $\mathcal{I} \subset \{1, \cdots, m\} \times \{1, \cdots, n\}$, $B \in \mathbb{R}^{m\times n}$, $\underline{X} \in \mathbb{R}^{m\times n}$, $\bar{X} \in \mathbb{R}^{m\times n}$, and where $\bar{\mathcal{I}}$ denotes the complement of the set $\mathcal{I}$. We want to inspect our convex heuristic for this problem by replacing the objective with $s_l(X)$, i.e., we will solve:

$$\begin{array}{llll} X_l^* \in & \arg\min & s_l(X) & \\ & \text{subject to} & X_{ij} = B_{ij} & \text{for} \quad (i,j) \in \mathcal{I} \\ & & \underline{X}_{ij} \le X_{ij} \le \bar{X}_{ij} & \text{for} \quad (i,j) \in \bar{\mathcal{I}}, \end{array} \qquad (6)$$

Example 12: In this example we will fix $m = n = 20$, and consider 200 instances of (6) by random sampling of $\mathcal{I}$, B, $\underline{X}$, and $\bar{X}$. We will vary l and look at the k-th singular value of the solution of (6). More precisely we will plot the difference of $\sigma_k(X_l^*)$ to its minimum when we vary l, i.e., plotting $\sigma_k(X_l^*) - \min_l \sigma_k(X_l^*)$ versus l. Results for $k = 4,\ 7$, and 16 are provided in Figures 1a, 1b and 1c respectively.

The vertical black dashed line indicates where $l = k$, the blue is the average of $\sigma_k(X_l^*) - \min_l \sigma_k(X_l^*)$ across 200 samples, the dashed cyan lines show the 5% and 95% quantiles for this metric, and the dashed blue shows where the average hits its minimum. It was observed that in for almost all k (even for the k that are not presented in this figure) the average hits its minimum is some point after $l > k$, suggesting that perhaps the best convex heuristic for minimizing σ_k in this class (by s_l) is achieved by a $l > k$.

We have tested our heuristic on random matrices where almost always the solutions would not be rank deficient. However it is interesting to see how this heuristic compares to nuclear norm for rank minimization. We will first show that how our heuristic and the nuclear norm are related to each other in the following theorem, and then discuss some experimental lessons when we apply them on a similar class of problems as in Example 12.

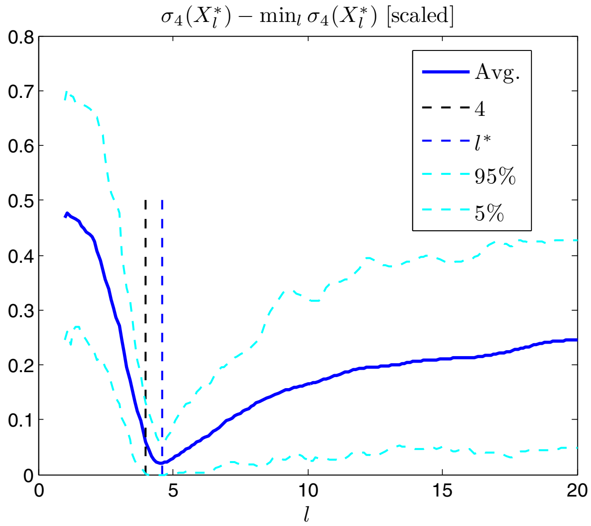
Theorem 13: If $X_{l_0}^*$ is a rank $\lceil l_0\rceil - 1$ minimizer of (4) for some $l_0 \in [1, n]$, then it also minimizes (4) for any $l \ge l_0$, i.e.: if $\sigma_{\lceil l_0\rceil}(X_{l_0}^*) = 0$ then $s_l(X_l^*) = s_l(X_{l_0}^*)$.

Proof: Proof is done by contradiction. Since X_l^* is a minimizer for l, the conclusion can be false only if $s_l(X_l^*) < s_l(X_{l_0}^*)$, for which we show that some singular value must be negative and achieves the contradiction. This condition can be equivalently written as:

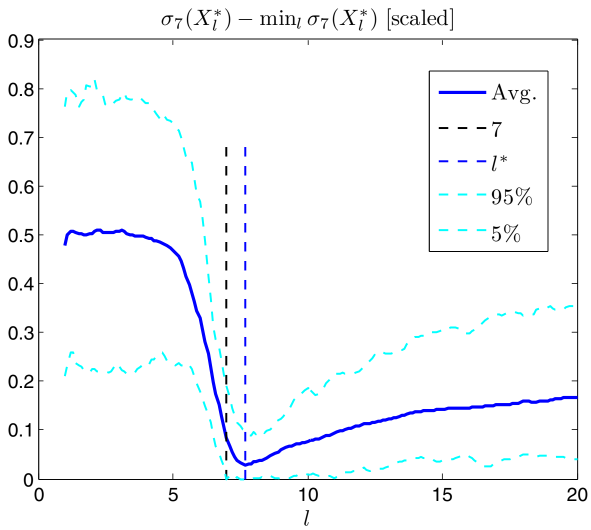
$$s_{l_0}(X_l^*) + s_l(X_l^*) - s_{l_0}(X_l^*) < s_l(X_{l_0}^*) = s_{l_0}(X_{l_0}^*) \qquad (7)$$

where the equality follows as $\sigma_{\lceil l_0\rceil}(X_{l_0}^*) = 0$ and thus $\sigma_k(X_{l_0}^*) = 0$ for all $k \ge \lceil l_0\rceil$, which by definition results in $s_l(X_{l_0}^*) = s_{l_0}(X_{l_0}^*)$ for all $l \ge l_0$. Also, since $X_{l_0}^*$ is a minimizer for l_0, we have that $s_{l_0}(X_l^*) \ge s_{l_0}(X_{l_0}^*)$ for all l, and thus (7) becomes true only if $s_l(X_l^*) - s_{l_0}(X_l^*) < 0$, meaning that some σ_k must be negative in the accumulative sum. This achieves the contradiction. ■

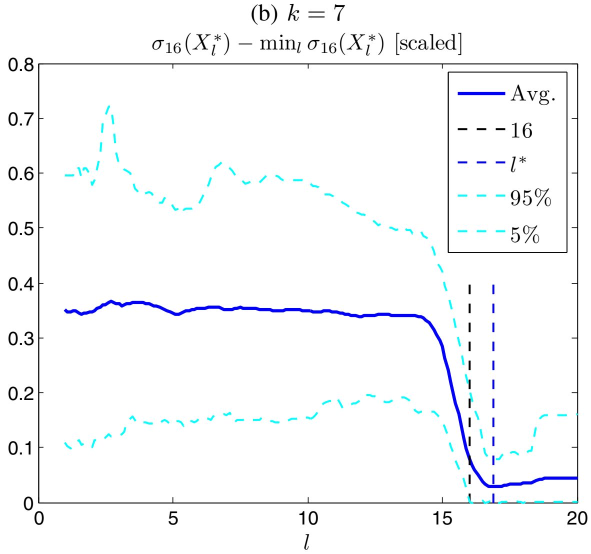
Remark 14: Theorem 13 suggests that there is no loss in considering the convex heuristic of nuclear norm ($l = n$) compared to the cases where $l < n$ when we have a sparse solution for some $l_0 < n$. This is contrary to the observations in Example 12 where the solutions were almost always full



(a) $k = 4$



(b) $k = 7$



(c) $k = 16$

Fig. 1: Non-integer partial sum as the convex heuristic for minimizing σ_k

rank.

In the next example we consider cases where the solution could be sparse and inspect our heuristic on this class also.

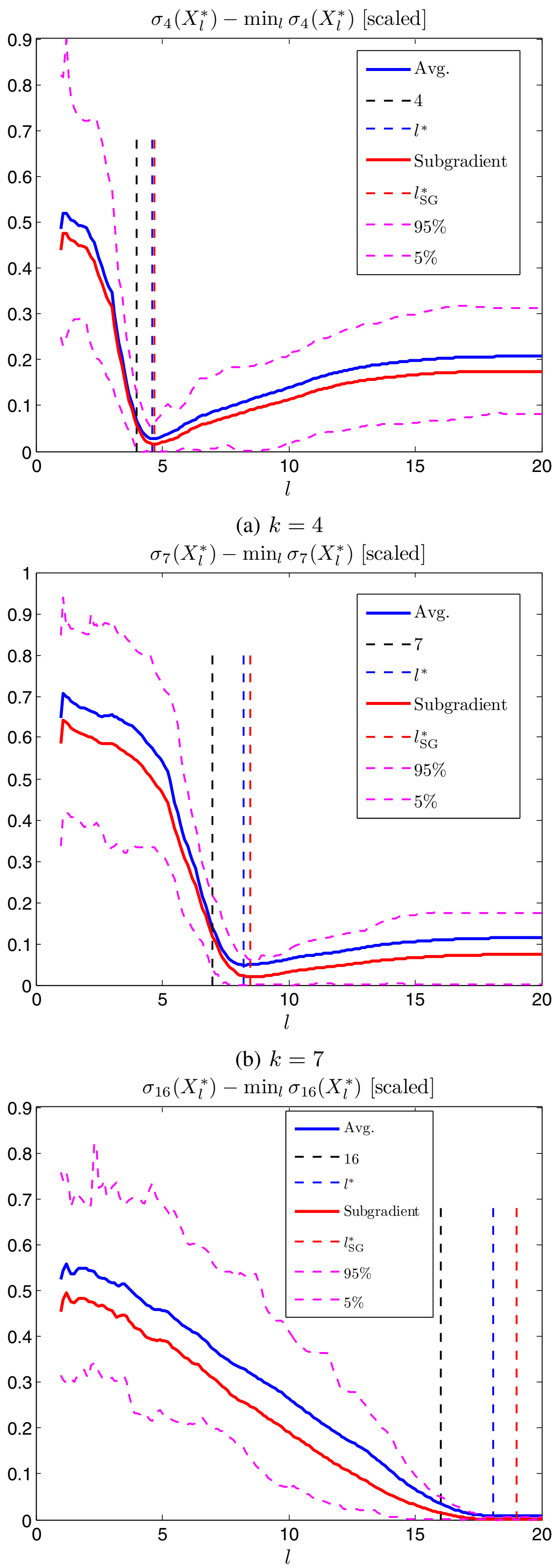
Example 15: We again fix $m = n = 20$, and consider 50 instances of (6). In each of these instances a random X_0 with rank 10 is generated first, then $\mathcal{I}$ is randomly selected and we set $B_{ij} = (X_0)_{ij}$ for all $(i,\ j) \in \mathcal{I}$, then $\underline{X}$ and $\bar{X}$ are randomly selected in a way that X_0 (the rank 10 solution) remains a feasible point. We again vary l and look at the k-th singular value of the solution of (6). We will further enhance $\sigma_k(X_l^*)$ with a subgradient based method. We will take X_l^* as our starting point, and do descent step along the subgradient of σ_k for each l. The subgradient is straightforward to derive when all the singular values are distinct, however more technical considerations would be needed when this would not be the case. See [4], [5] for a detailed derivation of the subgradient of the k-th singular value. Results for $k = 4,\ 7$, and 16 are provided in Figures 2a, 2b and 2c respectively.

The vertical black dashed line indicates where $l = k$, the blue is the average of $\sigma_k(X_l^*) - \min_l \sigma_k(X_l^*)$ across 50 samples, the red line shows the average further enhancement resulted from applying the subgradient based method on each of the samples, the dashed magenta lines show the 5% and 95% quantiles for the red line, the dashed blue shows where the blue hits its minimum, and the dashed red shows where the red hits its minimum. The observation that both the minimum of the convex heuristic and its enhanced version (by subgradient method) would happen for some $l > k$ was widespread. It can also be seen that when $k = 16$ the solution, up to numerical errors, would remain the same after some point (see Remark 14), meaning that for rank-deficient solutions nuclear norm would still recover as good as the suggested heuristic in this paper. The solution obtained from the nuclear norm is displayed in the figures where $l = n = 20$.

Our heuristic for minimizing the k-th singular value results in an upper bound for the global minimizer, and we want to have optimality certificates on how good these upper bounds will be. In the following section we will discuss methods for obtaining such lower bounds.

V. Lower Bounds

We will derive lower bounds for the global minimizer of the k-th singular value in this section. In each of the subsequent subsections we will formulate the k-th singular value as a different polynomial optimization problem which can then be used in conjunction with SOS techniques in Section II to obtain lower bounds. We will first use a factorization of positive semidefinite matrices to derive a polynomial optimization problem in Section V-A, we will lay out an alternative form with fewer constraints of higher degrees in Section V-B. Finally we will utilize the Courant-Fischer variational formulation of singular values to obtain another form which would require less resources for implementation. Properties related to each of these specific formulations are discussed where appropriate.



(a) $k = 4$

(b) $k = 7$

(c) $k = 16$

Fig. 2: Non-integer partial sum as the convex heuristic for minimizing σ_k in presence of a low-rank solution

In the remainder of this section we directly address problem (2) with the assumption that $f_i(X)$ and $h_j(X)$ would all be polynomials in X, however we no longer require that they would be convex or affine.

A. Lower Bound via Factorization

In this section we will utilize the factorization of the semidefinite matrices to transfer problem (2) into a polynomial optimization problem.

We will use the following lemma:

Lemma 16: Given a matrix $X \in \mathbb{R}^{m\times n}$ with nonzero singular values $\sigma_1(X) \geq \cdots \geq \sigma_n(X)$, for any $k \in \{1, \cdots, n\}$ we have that:

$$\sigma_k(X) = \min_{\substack{R\in\mathbb{R}^{m\times n}\\ \text{rank}(R)=k-1}} \|X - R\|_2.$$

Proof: See, for example [14, Eq. (5.12.10), p. 417]. ■

We can further replace $\text{rank}(R) = k-1$ constraint with $\text{rank}(R) \leq k-1$ and at the same time extend the result to allow for zero singular values as in the following corollary:

Corollary 17: Given a matrix $X \in \mathbb{R}^{m\times n}$ we have that:

$$\sigma_k(X) = \min_{\substack{R\in\mathbb{R}^{m\times n}\\ \text{rank}(R)\leq k-1}} \|X - R\|_2.$$

Proof: Let $U_X \Sigma_X V_X^*$ denote a SVD of X, then an argmin in Lemma 16 would be achieved by taking $R = U_X \,\text{diag}(\sigma_1(X), \cdots, \sigma_{k-1}(X), 0, \cdots, 0) V_X^*$ (even if $\text{rank}(X) < n$). Any solution with the strict constraint $\text{rank}(R) < k-1$ would imply $\sigma_{k-1}(R^*) = 0$. This would render the minimum value to be equal to $\sigma_{k-1}(X) \geq \sigma_k(X)$, meaning that the aforementioned R would still be an optimal solution. ■

By combining the above corollary with the SDP representation of the matrix 2-norm, we would have:

Theorem 18: The optimization problem (2) is equivalent to the following:

$$\begin{array}{lll} \text{minimize} & \tau & \\ \text{subject to} & f_i(X) \leq 0 & i = 1, \cdots, \bar{i} \\ & h_j(X) = 0 & j = 1, \cdots, \bar{j} \\ & R = UV & \\ & \begin{bmatrix} \tau I & X - R \\ (X-R)^T & \tau I \end{bmatrix} \succeq 0, & \end{array}$$

with variables $\tau \in \mathbb{R}, X \in \mathbb{R}^{m\times n}, R \in \mathbb{R}^{m\times n}, U \in \mathbb{R}^{m\times(k-1)}, V \in \mathbb{R}^{(k-1)\times n}$.

Proof: Constraint $\text{rank}(R) \leq k-1$ is equivalent to $R = UV$ where U has $k-1$ columns and V has $k-1$ rows. Then the result follows from SDP representation of the 2-norm. ■

We can insert $R = UV$ into the optimization problem and factorize the semidefinite constraint as what follows to derive a polynomial optimization problem:

Corollary 19: The optimization problem (2) is equivalent to the following:

$$\begin{array}{ll} \text{minimize} & \tau \\ \text{subject to} & f_i(X) \le 0 \qquad\qquad i = 1, \cdots, \bar{i} \\ & h_j(X) = 0 \qquad\qquad j = 1, \cdots, \bar{j} \\ & \begin{bmatrix} \tau I & X - UV \\ (X-UV)^T & \tau I \end{bmatrix} = G^T G, \end{array}$$

with variables $\tau \in \mathbb{R}, X \in \mathbb{R}^{m\times n}, U \in \mathbb{R}^{m\times(k-1)}, V \in \mathbb{R}^{(k-1)\times n}, G \in \mathbb{R}^{(m+n)\times(m+n)}$.

Remark 20: The minimization problem in Corollary 19 is a polynomial optimization problem with $(m+n)^2 + (k-1)(m+n)+mn+1$ variables and $(m+n)^2+\bar{i}+\bar{j}$ constraints. Furthermore each of the constraints (aside from $f_i(X) \le 0$ and $h_j(X) = 0$) are of degree two.

The SOS technique in Section II can now be applied to derive a lower bound for this polynomial optimization problem, however due to the large number of variables and constraints, we will consider an alternative form which give us fewer variables and constraints but of higher degrees in the following section.

B. Lower Bound via Leading Principle Minors

We will form an alternative formulation with fewer variables and constraints in this section. This form will only replace the semidefinite factorization in Corollary 19 with a constraint on the leading principle minors, which we define below.

Denote the d-dimensional identity matrix by $I_d \in \mathbb{R}^{d\times d}$, then the leading principle minors can be defined as:

Definition 21 (Leading Principle Minors): Given a square matrix $A \in \mathbb{R}^{n\times n}$ the leading principle minors of A are the determinants of the $d \times d$ sub-matrices obtained from only considering the first d rows and columns of A, where $d \in \{1, \cdots, n\}$. We denote them by $g_d(A)$:

$$g_d(A) \triangleq \det\left(\begin{bmatrix} I_d & 0 \\ 0 & 0 \end{bmatrix} A \begin{bmatrix} I_d & 0 \\ 0 & 0 \end{bmatrix}\right),$$

for $d \in \{1, \cdots, n\}$.

Positive definiteness of a matrix can be equivalently stated in terms of its leading principle minors:

Lemma 22: Given a symmetric matrix $A \in \mathbb{R}^{n\times n}$, we have that $A \succ 0$ if and only if each of the n leading principle minors of A are strictly positive.

Proof: See, for example [15, Theorem 3, p. 306], or [16, Sylvester's Criterion] ■

Remark 23: It is noteworthy that although checking for positive definiteness of a matrix is equivalent to n leading principle minors being positive, checking if a matrix is positive semidefinite requires that all the principle minors would be nonnegative [15, Theorem 4, p. 307]. There are $2^n - 2$ principle minors of A.

Although it is possible to derive an exact equivalent to the optimization problem in Theorem 18 by principle minors, however due to excessive number of resulting constraints we will derive an approximate one based on leading principle minors by first tightening the positive semidefinite constraint in Theorem 18 to the positive definiteness and then use Lemma 22 to derive an approximate lower bound with fewer constraints and variables in the following corollary.

Corollary 24: The optimization problem in Theorem 18 when replacing the positive semidefinite constraint with positive definite constraint is equivalent to:

$$\begin{array}{ll} \text{minimize} & \tau \\ \text{subject to} & f_i(X) \le 0 \qquad i = 1, \cdots, \bar{i} \\ & h_j(X) = 0 \qquad j = 1, \cdots, \bar{j} \\ & g_d\left(\begin{bmatrix} \tau I & X - UV \\ (X-UV)^T & \tau I \end{bmatrix}\right) > 0 \\ & \qquad\qquad\qquad\qquad d = 1, \cdots, m+n, \end{array}$$

with variables $\tau \in \mathbb{R}, X \in \mathbb{R}^{m\times n}, U \in \mathbb{R}^{m\times(k-1)}, V \in \mathbb{R}^{(k-1)\times n}$.

Proof: The corollary is a direct consequence of Lemma 22 ■

Remark 25: The only approximation in Corollary 24 compared to the exact formulation in Corollary 19 is due to tightening the semidefinite constraint, which when implemented numerically would not be of concern. Furthermore, the minimization problem in Corollary 24 is a polynomial optimization problem with $(k-1)(m+n)+mn+1$ variables and $m + n + \bar{i} + \bar{j}$ constraints. These constraints (aside from $f_i(X) \le 0$ and $h_j(X) = 0$) are of degree $m + n$ at most.

C. Lower Bound by Sampling from Courant-Fischer

We will first review the Courant-Fischer variational formulation of the singular values, and will transform this formulation into a polynomial program by rewriting some of the constraints. To this end, given a symmetric matrix M, and a non-zero real vector v with compatible dimension, define the ***Rayleigh quotient***, denoted by $R(M, v)$, as:

$$R(M, v) \triangleq \frac{v^T M v}{v^T v}.$$

The next theorem reviews the Courant-Fischer formulation of the singular values, and is adopted to the notation used in this paper.

Theorem 26 (Courant-Fischer): Given a matrix $X \in \mathbb{R}^{m\times n}$ the following relations holds:

$$\sigma_k^2(X) = \min_{v_1, v_2, \cdots, v_{k-1} \in \mathbb{R}^n} \max_{\substack{v \neq 0, v \in \mathbb{R}^n, \\ v \perp v_1, \cdots, v_{k-1}}} R\left(X^T X, v\right) \quad (8)$$

$$\sigma_k^2(X) = \max_{v_1, v_2, \cdots, v_{n-k} \in \mathbb{R}^n} \min_{\substack{v \neq 0, v \in \mathbb{R}^n, \\ v \perp v_1, \cdots, v_{n-k}}} R\left(X^T X, v\right). \quad (9)$$

Proof: The proof is a direct consequence of [17, Theorem, 4.2.11, p. 179] when considering that $X^T X$ is an $n\times n$ symmetric matrix, for which we have that $\lambda_k(X^T X) = \sigma_k^2(X)$. ■

Remark 27 (Rayleigh-Ritz): This theorem extends Rayleigh-Ritz theorem on the largest and smallest eigenvalues of a Hermitian matrix [17, Theorem 4.2.2, p. 176], which states that for a symmetric matrix $M \in \mathbb{R}^{n\times n}$, with eigenvalues $\lambda_1(M) \ge \cdots \ge \lambda_n(M)$, we have that:

$$\lambda_{\max}(M) = \lambda_1(M) = \max_{v \in \mathbb{R}^n} R(M, v),$$

and

$$\lambda_{\min}(M) = \lambda_n(M) = \min_{v\in\mathbb{R}^n} R(M, v).$$

We can equivalently write (8) and (9) using rank constraints on the considered subspaces:

Corollary 28: Given a matrix $X \in \mathbb{R}^{m\times n}$ we have:

$$\sigma_k^2(X) = \min_{\substack{V\in\mathbb{R}^{(k-1)\times n},\\ \mathrm{rank}(V)=k-1}} \max_{\substack{v\in\mathbb{R}^n,\\ Vv=0\\ v^Tv=1}} v^T X^T X v \quad (10)$$

$$\sigma_k^2(X) = \max_{\substack{V\in\mathbb{R}^{(n-k)\times n},\\ \mathrm{rank}(V)=n-k}} \min_{\substack{v\in\mathbb{R}^n,\\ Vv=0\\ v^Tv=1}} v^T X^T X v. \quad (11)$$

Proof: It is straightforward to replace the Rayleigh quotient with only its numerator, while enforcing the denominator to be equal to one, as any non-zero v can be scaled to have unit norm. For the min-max formulation of (8), we can gather $v_1, \cdots, v_{k-1}$ in a $(k-1)\times n$ matrix as $V = \begin{bmatrix} v_1 & \cdots & v_{k-1}\end{bmatrix}^T$. Then, the minimum would be achieved when all the $v_1, \cdots, v_{k-1}$ are independent of each other so as to make the feasible set for the v in the max part as small as possible, which is equivalent to the constraint $\mathrm{rank}(V) = k-1$. Similar reasoning applies to the max-min formulation. ■

We would further relax (10) and (11) by considering finite samples of the rank-constrained subspaces in those equations, and show that this would lead to upper and lower bounds on the singular values. This is illustrated in the following corollary:

Corollary 29: Suppose that a matrix $X \in \mathbb{R}^{m\times n}$ is given, Let $q \in \mathbb{N}$ be the desired number of the samples from the rank constrained subspaces, and for $1 \le k \le n$, let $\bar{V}_1, \cdots, \bar{V}_q$ be matrices that are sampled from the subspace specified by the rank constraint in (10), i.e., they are all in $\mathbb{R}^{(k-1)\times n}$ and have rank $k-1$. Similarly let $\underline{V}_1, \cdots, \underline{V}_q$ be matrices that are sampled from the rank constraint in (11), that are all in $\mathbb{R}^{(n-k)\times n}$ and have rank $n-k$. Then we have:

$$\begin{array}{llll} \sigma_k^2(X) \le & \max & \gamma & \\ & \text{s.t.} & \gamma \le \bar{v}_i^T X^T X \bar{v}_i & i = 1, \cdots, q \\ & & \bar{V}_i \bar{v}_i = 0 & i = 1, \cdots, q \\ & & \bar{v}_i^T \bar{v}_i = 1 & i = 1, \cdots, q, \end{array} \quad (12)$$

with variables $\gamma \in \mathbb{R}$, and $\bar{v}_1, \cdots, \bar{v}_q \in \mathbb{R}^n$. Similarly, we have the following lower bound

$$\begin{array}{llll} \sigma_k^2(X) \ge & \min & \gamma & \\ & \text{s.t.} & \gamma \ge \underline{v}_i^T X^T X \underline{v}_i & i = 1, \cdots, q \\ & & \underline{V}_i \underline{v}_i = 0 & i = 1, \cdots, q \\ & & \underline{v}_i^T \underline{v}_i = 1 & i = 1, \cdots, q, \end{array} \quad (13)$$

with variables $\gamma \in \mathbb{R}$ and $\underline{v}_1, \cdots, \underline{v}_q \in \mathbb{R}^n$.

Proof: The upper-bound in (12) is achieved due to the fact that by finite sampling we are minimizing over a smaller set rather than the rank constrained subspace in (10), and hence the minimum value would increase. Similarly, in (13) we are considering only finite numbers of $\underline{V}_1, \cdots, \underline{V}_q$, rather than the original subspace specified by the rank constraint in (11), and hence the maximum value would decrease and we would have a lower bound for $\sigma_k(X)$. ■

We will next utilize this lower bound for the optimization problem (2).

Corollary 30: Given a finite number $q \in \mathbb{N}$, let $\underline{V}_1, \cdots, \underline{V}_q$ be q rank $n-k$ samples from $\mathbb{R}^{(n-k)\times n}$, then a lower bound for the optimization problem (2) can be obtained by taking $\sqrt{\gamma}$ from the following polynomial optimization problem:

$$\begin{array}{lll} \text{minimize} & \gamma & \\ \text{subject to} & f_i(X) \le 0 & i = 1, \cdots, \bar{i} \\ & h_j(X) = 0 & j = 1, \cdots, \bar{j} \\ & \gamma \ge \underline{v}_i^T X^T X \underline{v}_i & i = 1, \cdots, q \\ & \underline{V}_i \underline{v}_i = 0 & i = 1, \cdots, q \\ & \underline{v}_i^T \underline{v}_i = 1 & i = 1, \cdots, q, \end{array}$$

with variables $\gamma \in \mathbb{R}$, $X \in \mathbb{R}^{m\times n}$, and $\underline{v}_1, \cdots, \underline{v}_q \in \mathbb{R}^n$.

Remark 31: The tightness of this lower bound would be dependent on the number (q) and choices of the samples. However, the advantage of this approach compared to the ones in the previous sections is that the required resources (memory and computation time) can be implicitly controlled by choosing a moderate value for q.

VI. An Equivalent Bilinear Matrix Inequality

We will derive an equivalent BMI to (2) in this section, which would be based on the min-max form in the Courant-Fischer formulation. This BMI is further subjected to have an orthonormal constraint on one of its variable. To this end we will equivalently write (8) as:

$$\sigma_k^2(X) = \min_{\substack{\mathcal{V}_{n-k+1}\subseteq\mathbb{R}^n\\ \dim(\mathcal{V}_{n-k+1})=n-k+1}} \max_{\substack{v\in\mathcal{V}_{n-k+1}\\ v^Tv=1}} v^T X^T X v \quad (14)$$

Which is further equivalent to the form stated in the following corollary:

Corollary 32: Given a matrix $X \in \mathbb{R}^{m\times n}$ we have:

$$\sigma_k^2(X) = \min_{\substack{R\in\mathbb{R}^{n\times(n-k+1)}\\ R^TR=I}} \max_{\substack{v\in\mathbb{R}^n, x\in\mathbb{R}^{n-k+1}\\ v^Tv\le 1\\ v=Rx}} v^T X^T X v \quad (15)$$

Proof: We can represent the $n-k+1$ dimensional subspaces $\mathcal{V}_{n-k+1}$ in $\mathbb{R}^n$ via range-spaces of full column rank matrices $R \in \mathbb{R}^{n\times(n-k+1)}$. Since we are only interested in the directions specified by this full column rank R, without loss of generality it can also be assumed that it is orthonormal. It is also straightforward to check that given any direction $v \in \mathbb{R}^n$, the inner maximum would occur at the boundary of the unit circle, and thus we can replace the $v^Tv = 1$ constraint with $v^Tv \le 1$. ■

We will further insert $v = Rx$ for v in (15) to have:

$$\sigma_k^2(X) = \min_{\substack{R\in\mathbb{R}^{n\times(n-k+1)}\\ R^TR=I}} \max_{\substack{x\in\mathbb{R}^{n-k+1}\\ x^Tx\le 1}} x^T R^T X^T X R x. \quad (16)$$

The inner maximization has a quadratic objective subject to only a single quadratic constraint, and thus by strong duality we have:

$$\max_{\substack{x\in\mathbb{R}^{n-k+1}\\ x^Tx\le 1}} x^T R^T X^T X R x = \min_{\substack{\tau\in\mathbb{R},\ \tau\ge 0\\ R^TX^TXR\preceq\tau^2 I}} \tau^2 \quad (17)$$

By combining (17), (16) and (2) we would have the following theorem:

Theorem 33: Problem (2) can be equivalently written as:

$$\begin{array}{lll} \text{minimize} & \tau & \\ \text{subject to} & f_i(X) \leq 0 & i = 1, \cdots, \bar{i} \\ & h_j(X) = 0 & j = 1, \cdots, \bar{j} \\ & R^T R = I & \\ & \begin{bmatrix} \tau I & XR \\ (XR)^T & \tau I \end{bmatrix} \succeq 0, & \end{array} \quad (18)$$

with variables $\tau \in \mathbb{R}$, $X \in \mathbb{R}^{m \times n}$, and $R \in \mathbb{R}^{n \times (n-k+1)}$.

Proof: Insert (17) into (16). We are interested in $\sigma_k(X)$ and not $\sigma_k^2(X)$, hence noting that τ^2 is a monotonic function of τ, we can replace the objective with τ. We will further insert this for $\sigma_k(X)$ in (2) which would give us (18). ∎

VII. Future Directions

We will discuss some likely future directions in this section.

- Search for further applications in different domains, and the properties related to these specific applications is a future direction.
- We demonstrated that minimizing $s_l(X)$ generally works best as a convex heuristic for σ_k when l is chosen somewhere in between k and n; it would be desirable to express this as a precise mathematical statement, perhaps whereby a particular function of k and n determines the l such that an expectation of σ_k of the solution to the heuristic problem is minimized.
- Perhaps obtaining more resource-efficient lower bounds could be pursued by either different formulations, or a more domain specific choice of samples in the Courant-Fischer formulation by using the upper bound information.
- Benefiting from available methods on computational aspects of bilinear matrix inequalities and combining those with the ones regarding orthonormal constraints could be interesting for deriving better upper or lower bounds.

VIII. Conclusions

We considered the minimization of a particular singular value of a matrix in this paper. We analyzed a class of convex heuristics which resulted in an initially counter-intuitive numerical observations. It was observed that this convex heuristic would perform better than the usual ones when the other methods can not make the considered singular value zero. We further inspected how this general observation would transform when we enhance the minimization of the considered singular value using its subgradient. We also provided an equivalent bilinear matrix inequality subject to one of the variables being orthonormal for our problem.

We formulated minimization of a particular singular value subjected to polynomial constraints as a polynomial optimization problem for which we can use SOS techniques to derive lower bounds. We provided different polynomial optimization problems, which differ from one another in terms of required resources and tightness of the theoretical bound that will be obtained from each one.

References

[1] B. Recht, W. Xu, and B. Hassibi, "Necessary and sufficient conditions for success of the nuclear norm heuristic for rank minimization," in *Proc. IEEE Conference on Decision and Control*. IEEE, 2008, pp. 3065–3070.

[2] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM review*, vol. 52, no. 3, pp. 471–501, 2010.

[3] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[4] A. S. Lewis and H. S. Sendov, "Nonsmooth analysis of singular values. Part I: Theory," *Set-Valued Analysis*, vol. 13, no. 3, pp. 213–241, 2005.

[5] ——, "Nonsmooth analysis of singular values. Part II: Applications," *Set-Valued Analysis*, vol. 13, no. 3, pp. 243–264, 2005.

[6] T.-H. Oh, Y.-W. Tai, J.-C. Bazin, H. Kim, and I. S. Kweon, "Partial sum minimization of singular values in robust pca: Algorithm and applications," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 4, pp. 744–758, 2016.

[7] M. Gu, E. Mengi, M. L. Overton, J. Xia, and J. Zhu, "Fast methods for estimating the distance to uncontrollability," *SIAM journal on matrix analysis and applications*, vol. 28, no. 2, pp. 477–502, 2006.

[8] J. V. Burke, A. S. Lewis, and M. L. Overton, "Pseudospectral components and the distance to uncontrollability," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 2, pp. 350–361, 2004.

[9] B. Dumitrescu, B. Sicleru, and R. Stefan, "Computing the controllability radius: a semi-definite programming approach," *IET Control Theory & Applications*, vol. 3, no. 6, pp. 654–660, 2009.

[10] A. Vaz and E. Davison, "A measure for the decentralized assignability of eigenvalues," *Systems and Control Letters*, vol. 10, no. 3, pp. 191 – 199, 1988.

[11] A. Alavian and M. C. Rotkowitz, "An optimization-based approach to decentralized assignability," in *Proc. American Control Conference*, 2016, pp. 5199–5204.

[12] ——, "Polynomial optimization methods for determining lower bounds on decentralized assignability," to appear in Proc. of the Annual Allerton Conference on Communication Control and Computing, 2016.

[13] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[14] C. D. Meyer, *Matrix analysis and applied linear algebra*. Siam, 2000, vol. 2.

[15] F. R. Gantmakher, *The theory of matrices*. American Mathematical Soc., 1998, vol. 131.

[16] G. T. Gilbert, "Positive definite matrices and sylvester's criterion," *The American Mathematical Monthly*, vol. 98, no. 1, pp. 44–46, 1991.

[17] R. Horn and C. R. Johnson, Eds., *Matrix Analysis*. New York, NY, USA: Cambridge University Press, 1985.