On Sampling Complexity of the Semidefinite Affine Rank Feasibility Problem

Igor Molybog and **Javad Lavaei**
University of California at Berkeley
Berkeley, CA 94720, USA

Abstract

In this paper, we study the semidefinite affine rank feasibility problem, which consists in finding a positive semidefinite matrix of a given rank from its linear measurements. We consider the semidefinite programming relaxations of the problem with different objective functions and study their properties. In particular, we propose an analytical bound on the number of relaxations that are sufficient to solve in order to obtain a solution of a generic instance of the semidefinite affine rank feasibility problem or prove that there is no solution. This is followed by a heuristic algorithm based on semidefinite relaxation and an experimental proof of its performance on a large sample of synthetic data.

Introduction

This paper is concerned with the problem of obtaining a real matrix $\mathbf{X}$ from the set $\mathbb{S}^+_{n;k}$ of $n \times n$ symmetric ($\mathbb{S}_n$) and positive semidefinite matrices of rank k that satisfies m linear specifications. In case there is no such a matrix up to a given accuracy, we expect to receive the corresponding information within finite time. We refer to this problem as the *Semidefinite Affine Rank Feasibility (SARF)*, which can be formally written as:

$$
\begin{aligned}
&\text{find} && \mathbf{X} \in \mathbb{S}_n \\
&\text{subject to} && \langle \mathbf{M}_r, \mathbf{X} \rangle = y_r, \quad r = 1, \ldots, m, && \text{(1a)} \\
&&& \mathbf{X} \succeq 0, && \text{(1b)} \\
&&& \operatorname{rank}\{\mathbf{X}\} = k. && \text{(1c)}
\end{aligned}
$$

where $\mathbf{M}_1, \ldots \mathbf{M}_m \in \mathbb{S}_n$ are some known symmetric matrices, $y_1, \ldots, y_m \in \mathbb{R}$ are some scalars that parametrize the problem, and $\langle \cdot, \cdot \rangle$ denotes the Frobenius inner product.

The above problem is closely related to the *Affine Rank Feasibility (ARF)* problem, which can be obtained by dropping the constraint (1b). It is clear that a solution of the Semidefinite Affine Rank Feasibility problem must also be a solution of its counterpart without any positive semidefinite constraint. More precisely, if $\mathbf{X}^* \succeq 0$ is the unique solution of the Affine Rank Feasibility problem, then it is also a solution of the corresponding SARF problem as well. In addition, every ARF problem can be equivalently converted to a SARF problem (Fazel, 2002; Madani et al., 2017).

Both of the above feasibility problems encompass several practically important cases. For example, the Phase Retrieval (Candes et al., 2015) and Quantum State Tomography (Kueng, Rauhut, and Terstiege, 2017) problems can be treated as special cases of the Semidefinite Affine Rank Feasibility problem with rank-one matrices $\mathbf{M}_r$ and $k = 1$. In Power Systems Engineering, the State Estimation problem consists in recovering complex voltages based on measured power flows and select voltage magnitudes for a power grid, which can be formulated as the Semidefinite Affine Rank-one Feasibility problem (Zhang, Madani, and Lavaei, 2018). In this case, the matrices $\mathbf{M}_r$ depend on the topology of the power system and change a few times a day, while the measurements y_r change every 5-15 minutes based on the data updated by system operators. This defines a class of problems with a fixed set of matrices $\mathbf{M}_r$'s. In wireless communication systems, the problem of Feasible Downlink Beamforming (Morency and Vorobyov, 2016) can be formulated as a block-diagonally shaped Affine Rank-one Feasibility problem. The most commonly studied problem with the target rank being not necessarily equal to 1 is the Low-rank Matrix Completion problem for which each matrix $\mathbf{M}_r$ has exactly one nonzero element in its upper/lower triangular part (Candès and Recht, 2009).

While the *Affine Rank Minimization (ARM)* (Recht, Fazel, and Parrilo, 2010) is probably the most studied rank-constrained problem, the solution of any Affine Rank Minimization problem can also be obtained from the solutions of $\mathcal{O}(n)$ Affine Rank Feasibility problems of the same size. The same is also true for those problems with a positive semidefinite constraint, so the results of this paper can be extended to rank minimization problems as well. From the point of view of computational complexity, Affine Rank Minimization is an $\mathcal{NP}$-hard, since it contains Cardinality Minimization as a special case (Natarajan, 1995). Moreover, the work by Marianna et al. (2013) establishes the result that the rank-constrained completion of a semidefinite matrix with all diagonal entries equal to 1 is $\mathcal{NP}$-hard as soon as $k \geq 2$. Hence, the Semidefinite Affine Rank Feasibility is also $\mathcal{NP}$-hard.

Nevertheless, the above problems have been extensively studied over the past decade, and practical approaches have been developed in various special cases. A major line of research assumes that the matrices $\mathbf{M}_r$ form a linear operator that satisfies a Restricted Isometry Property. For example, those matrices whose elements are independently sampled from the standard Gaussian distribution satisfy this property with high probability. Classical papers with this assumption adopt Nuclear Norm Minimization (Recht, Xu, and Hassibi, 2008; Candès and Recht, 2009; Recht, Fazel, and Parrilo, 2010; Cai and Zhang, 2013; Kueng, Rauhut, and Terstiege, 2017), while more recent developments deploy the minimization of other surrogate functions (Cui, Peng, and Li, 2018). In particular, the minimization of a rank function has received a lot of attention for applications related to artificial intelligence. For example, Xu, Lin, and Zha (2017) propose a surrogate of the Schatten-p norm as a spectral regularization, while Zhang and Zhang (2017) attack the problem of rank-constrained distance matrices. For the Low-rank Matrix Completion problem, a number of nonconvex techniques have been developed recently. More importantly, Bhojanapalli, Neyshabur, and Srebro (2016); Ge, Lee, and Ma (2016); Ge, Jin, and Zheng (2017); Zhang et al. (2018) show that the classical l_2-norm regression has no spurious local minimum under a Restricted Isometry Property. In addition, Josz et al. (2018) prove a similar result for nonsmooth problems, including l_1-norm regression. One possible option to avoid the Restricted Isometry Property assumption is to require the distribution of the true solution $\mathbf{X}^*$ to belong to a certain ensemble (Xin and Wipf, 2015).

An important contribution into the Low-rank Matrix Recovery has been made by Madani et al. (2017), where conic relaxations of linear matrix inequalities (LMIs) are proposed and it is shown that there is a low-rank solution for any sparse LMI with an upper bound on the rank being a function of certain graph-theoretic parameters. Note that any LMI Rank Feasibility problem is indeed a Semidefinite Affine Rank Feasibility problem. From this perspective, there is another application of Semidefinite Affine Rank Feasibility problem, which can be found in reducing the complexity of large-scale semidefinite programs (Fukuda et al., 2001; Andersen, Hansson, and Vandenberghe, 2014).

Approaches to the problem also vary depending on the noise model. Besides Gaussian additive noise (Jain, Meka, and Dhillon, 2010; Mohan and Fazel, 2012), a particularly interesting case is the presence of sparse noise, which has been considered by Candès et al. (2011); Klopp, Lounici, and Tsybakov (2017); Akhriev, Marecek, and Simonetto (2018).

One of the key questions studied in the above papers is the sampling complexity, which can be interpreted as the amount of information contained in the linear measurements (1a). This can be measured in terms of the number of measurements m and/or the sampling strategy of selecting $\mathbf{M}_r$'s such that the problem becomes polynomial-time solvable. The known results on this topic are related to the Matrix Completion problem (Candès and Recht, 2009; Wei et al., 2016). The present paper develops the first result in the literature that studies the sampling complexity of a general Semidefinite Affine Rank Feasibility problem.

Notation

The symbol $\mathbb{R}$ denotes the set of real numbers. $\mathbb{S}^+_{n;k}$ is the set of $n \times n$ positive semidefinite matrices of rank k, whereas $\mathcal{T}^+_{n;k} \subset \mathbb{S}^+_{n;k}$ is the set of matrices in $\mathbb{S}^+_{n;k}$ whose nonzero eigenvalues are all equal to 1 (Stiefel manifold). In a normed vector space $(V, \|\cdot\|)$, the symbol $B_{V,\|\cdot\|}(r)$ stands for the r-ball centred at zero. Matrices are shown in capital bold, and vectors are shown with small bold letters. The notation $\mathbf{W} \succeq 0$ means that $\mathbf{W}$ is a symmetric positive semidefinite matrix. $\|\cdot\|_2$ and $\|\cdot\|_F$ stand for the Euclidean and Frobenius matrix norms. $\mathbf{0}$ denotes a zero matrix of appropriate dimension. $\sigma_{min}(\cdot)$ and $\sigma_k(\cdot)$ are the minimum singular value operator and k-largest singular value operator, respectively. $[\mathbf{A}]_{ij}$ is the (i, j)-component of the matrix $\mathbf{A}$. The symbol $\mathcal{L}_{n;k} \triangleq \{\mathbf{V} \in \mathbb{R}^{n\times k} \mid V_{ij} = 0 \text{ if } i < j\}$ denotes the set of lower triangular matrices and δ_{ij} stands for the indicator function of $\{i = j\}$. For a set $A = \{a_1, \dots a_{|A|}\} \subseteq \{1, \dots, n\}$, with the complement $\{1, \dots, n\}\backslash A = \{b_1, \dots, b_{n-|A|}\}$ such that $a_i < a_j$, $b_i < b_j$ for every $i < j$, let

$$\mathbf{\Pi}_{n;A} = \begin{bmatrix} [\delta_{a_i,j}]_{i=1,\dots,k;j=1,\dots,n} \\ [\delta_{b_i,j}]_{i=1,\dots,n-k;j=1,\dots,n} \end{bmatrix}$$

denote the permutation matrix that puts the entries of a vector with their indexes in the set A instead of those which have the indexes $\{1, \dots, |A|\}$. Given two index sets A and B, the matrix $\mathbf{W}[A, B]$ is obtained by keeping only those rows of $\mathbf{W}$ that correspond to the elements of the set A and those columns of $\mathbf{W}$ that correspond to the elements of the set B. Let $\mathcal{V}_{n;k}(\cdot)$ be a lower triangular vectorization operator such that

$$\mathcal{V}_{n;k}(\mathbf{V}) = [V_{11}, V_{21}, \dots, V_{n1}, V_{22}, V_{32}, \dots, V_{n2}, \dots, V_{kk}, \dots, V_{nk}]^T \in \mathbb{R}^{nk-\frac{k(k-1)}{2}}$$

for every matrix $\mathbf{V} \in \mathbb{R}^{n\times k}$.

Problem Formulation

To make the computational complexity analysis of the Semidefinite Affine Rank Feasibility problem (1) meaningful, we consider **a class of feasibility problems specified by the matrices** $\{\mathbf{M}_1, \dots, \mathbf{M}_m\}$. This defines infinitely many feasibility problems, where every rank-k positive semidefinite matrix X is a solution to some feasibility problem in this class (by appropriately designing $y_1, \dots, y_m$). The analysis of this class of problems is partially motivated by data analytics for electric power systems, where the matrices $\mathbf{M}_r$ are designed based on the parameters of the infrastructure that are considered to be fixed (as long as there is no network reconfiguration) while the measurements $y_1, \dots, y_m$ used by power operators change every 5-15 minutes. In this regard, we have infinitely many feasibility problems with the same matrices $\mathbf{M}_1, \dots, \mathbf{M}_m$. The results of this paper are based on the idea of constructing a semidefinite programming (SDP) of the form

$$\underset{\mathbf{X}\in\mathbb{S}_n}{\text{minimize}} \quad \langle \mathbf{N}, \mathbf{X}\rangle \tag{2a}$$

$$\text{subject to} \quad \langle \mathbf{M}_r, \mathbf{X}\rangle = y_r, \quad r = 1, \dots, m, \tag{2b}$$

$$\mathbf{X} \succeq 0, \tag{2c}$$

corresponding to the Semidefinite Affine Rank Feasibility problem (1). Here, $\mathbf{N}$ is the matrix of an orthogonal projection onto an $(n-k)$–dimensional linear subspace. In this case, the convex problem (2) is called an **SDP relaxation with the parameters** $(\mathbf{N}, \mathbf{y})$. A solution $\mathbf{X}^*$ of the Semidefinite Affine Rank Feasibility problem (1) is said to be **recoverable** through the SDP relaxation (2) with $\mathbf{N}$ if it is the unique optimal solution of (2) with the parameters $(\mathbf{N}, [\langle \mathbf{M}_r, \mathbf{X}^* \rangle]_{r=1}^m)$.

The **recovery region** of the matrix $\mathbf{N}$ associated with the class $\{\mathbf{M}_1, \ldots, \mathbf{M}_m\}$ is the set of all rank-k positive semidefinite matrices that are recoverable for their corresponding feasibility problems via the SDP relaxation with the objective function $\langle \mathbf{N}, \mathbf{X} \rangle$. Figure 1 depicts a two-dimensional slice of the recovery region of a single matrix $\mathbf{N}$ for a randomly generated problem (see the section "Numerical Results" for more details). It can be observed that the recovery region is not convex in general, but there is a ball with a positive radius in the space of matrices such that every matrix in the ball is recoverable via an SDP relaxation with the single objective matrix $\mathbf{N}$ for the corresponding Affine Rank Feasibility problem. This observation was first noticed and formalized in Theorem 1 of Ashraphijuo, Madani, and Lavaei (2016).

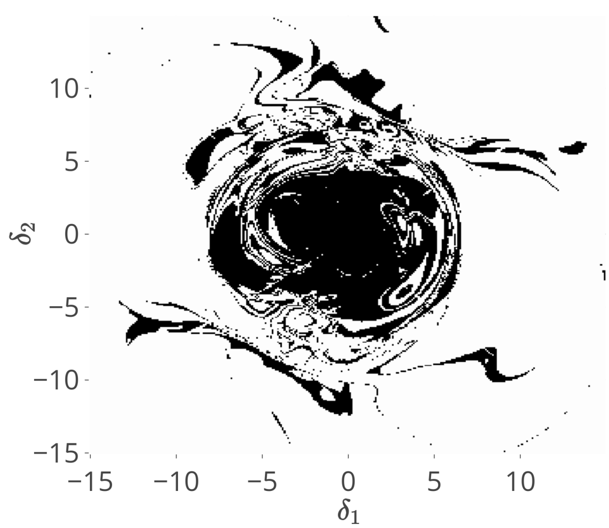


Figure 1: A slice of the recovery region of a single matrix. The x and y axes represent δ_1 and δ_2, respectively (for definition, see "Implementation" in the section "Numerical Results").

One of the main results of Ashraphijuo, Madani, and Lavaei (2016) is that under $m > k(2n-k)$, there are a finite number of SDPs with specially designed objectives such that every instance of the Semidefinite Affine Rank Feasibility problem (1) in the infinite class defined by $\{\mathbf{M}_1, \ldots, \mathbf{M}_m\}$ is recoverable via one of those SDPs. In other words, there is a finite list of matrices $\mathbf{N}$ defining a set of SDPs that can be used to solve any of the infinity many feasibility problems sharing the same model. The usefulness of this technique has been demonstrated by Ashraphijuo, Madani, and Lavaei (2015) for solving a set of polynomial equations and by Molybog, Madani, and Lavaei (2018) for nonlinear regression under sparse additive noise. The paper by Madani, Lavaei, and Baldick (2018) applies the above technique to the power flow / state estimation problem for power systems, where a small number of SDPs have solved many real-world feasibility problems and outperformed the existing methods.

The main objective of this paper is to find an upper bound on the number of SDP problems (2) to be solved in order to guarantee obtaining a solution of (1). Since we do not make any assumption on the structure of the matrices $\{\mathbf{M}_1, \ldots, \mathbf{M}_m\}$, the upper bound cannot be a polynomial function but the methodology pursued in this work could be used to study specialized problems (such as those appearing in power systems) to obtain a tighter upper bound for structured problems. Note also that although we do not make any assumption on the uniqueness of the solution of (1), it is well known that the solution is unique when m is large enough.

Main Results

We form a matrix $\mathbf{H}$ with $\mathcal{V}_{n;n}(\mathbf{M}_r)$ for $r = 1, ..., m$ as its columns. To set up a uniform bound on the size of a region that can be recovered through a single objective $\mathbf{N}$, define the following function of the measurements matrices:

$$r(\mathbf{H}) = \min_{\substack{\mathbf{Y} \in \mathcal{L}_{n;k} : \mathbf{Y}\mathbf{Y}^T \in \mathcal{T}_{n;k}^+; \\ A \subseteq \{1..n\} : |A| = k}}$$

$$\sigma_{min}([\mathcal{V}_{n;k}(\mathbf{\Pi}_{n;A}\mathbf{M}_1\mathbf{\Pi}_{n;A}^T\mathbf{Y}) \ldots \mathcal{V}_{n;k}(\mathbf{\Pi}_{n;A}\mathbf{M}_m\mathbf{\Pi}_{n;A}^T\mathbf{Y})])$$

Note that , according to Ashraphijuo, Madani, and Lavaei (2016), it holds that $r(\mathbf{H}) > 0$ if $m > k(2n-k)$, for a generic choice of $\{\mathbf{M}_r\}_{r=1}^m$. The genericity assumption in this statement implies that the inequality holds true for almost every choice of $\{\mathbf{M}_r \in \mathbb{S}^n\}_{r=1}^m$. A formal definition comes next.

Definition 1. *A property (Q) is said to hold for every generically chosen member of a topological space if there exists an open dense subset of it whose members all satisfy (Q).*

The main analytic result of this paper will be stated below, which sets up a bound on the sufficient number of SDP relaxations.

Theorem 1. *Given an arbitrary positive number κ, there are constants $C_1 = C_1(k, \{\mathbf{M}_r\}), C_2 = C_2(k, \{\mathbf{M}_r\}, \kappa, \|\mathbf{y}\|_2) \in \mathbb{R}$ and at most*

$$\min\left\{C_1^{k(n-k)}, C_2^{\frac{n(n+1)}{2}-m+1}\right\}$$

SDP relaxations of the form (2) such that any solution (satisfying $\frac{\sigma_1(\cdot)}{\sigma_k(\cdot)} \leq \kappa$) of a generic instance of the Semidefinite Affine Rank Feasibility problem (1) in the class of infinitely many feasibility problems defined by $\{\mathbf{M}_1, ..., \mathbf{M}_m\}$ can be obtained via one of those SDPs.

Proof. This is a direct corollary of Theorems 2 and 3 in the section "Proofs". $\square$

It can be observed that for a fixed k the bound in Theorem 1 is at most exponential in terms of the dimension of the problem for a generic choice of the matrices $\mathbf{M}_r$, while all known finite-time algorithms for the general ARF problem have at least doubly exponential running times (Recht, Fazel, and Parrilo, 2010). Moreover, the bound has a dependence on m, which implies the following corollary:

Corollary 1. *For every polynomial $p(\cdot)$, if the number of measurements obeys the inequality*

$$m \geq \frac{n(n+1)}{2} - \log p(n) \sim \mathcal{O}(n^2),$$

then a generic problem becomes polynomial-time solvable with any fixed sensing constant κ and up to an arbitrary nonzero precision error.

This captures the fact that the problem is expected to be easy when m is large because the feasible region of (2) for $m = \frac{n(n+1)}{2}$ collapses into at most a single point that should be the solution to the SARF problem (1).

The above result is the first one that studies the notion of "phase transition of complexity" for the generic rank-constrained feasibility problem. This notion has already been studied in the literature for other NP-hard problems (Cohen and Beck, 2017). We will continue this topic during the discussion of Numerical Results.

Proofs

The dual problem of (2) can be written in the form:

$$\underset{\mathbf{u}\in\mathbb{R}^m}{\text{minimize}} \qquad \mathbf{y}^{\mathrm{T}}\mathbf{u} \tag{3a}$$

$$\text{subject to} \qquad \mathbf{N} + \sum_{i=1}^{m} u_i \mathbf{M}_i \succeq 0 \tag{3b}$$

Since the primal and dual feasible matrices must be positive semidefinite, the Karush-Kuhn-Tucker (KKT) conditions impose the relationship:

$$\left(\mathbf{N} + \sum_{r=1}^{m} u_r \mathbf{M}_r\right)\mathbf{X} = \mathbf{0}. \tag{4}$$

Without loss of generality, we can assume that the vectors $\mathcal{V}_{n;n}(\mathbf{M}_r)$ form an orthonormal basis. To support this, suppose that the original problem (1) is associated with the matrices $\hat{\mathbf{M}}_r$ ($\langle \hat{\mathbf{M}}_r, \mathbf{X}\rangle = \hat{y}_r$). One can apply the Gram-Schmidt orthogonalization process to the system of vectors $\mathcal{V}_{n;n}(\hat{\mathbf{M}}_r)$ to obtain an orthonormal system of vectors that would generate symmetric matrices $\mathbf{M}_r$ in the obvious way. Note that $\mathbf{M}_r$ has a linear dependence on $\hat{\mathbf{M}}_r$ and, therefore, $y_r = \langle \mathbf{M}_r, \mathbf{X}\rangle$ can be computed by applying the same linear transformations to $\hat{y}_r$. Furthermore, now we can refer to m as to the maximum number of linearly independent matrices $\mathbf{M}_r$.

The next lemma states that each matrix of a potential solution can be represented through a projection matrix onto the span of its eigenvectors.

Lemma 1. *For any orthogonal matrix $\mathbf{U}$, if $\mathbf{X} = \mathbf{U}\mathbf{U}^T$ is not a singular point of $[\langle \mathbf{M}_r, \cdot\rangle]_{r=1}^m$ and recoverable with $\mathbf{N}$, then $\mathbf{X}' = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$ is also recoverable with $\mathbf{N}$ for any positive definite diagonal matrix $\boldsymbol{\Lambda}$.*

Proof. Strong duality holds for the pair of problems (2)-(3) due to Lemma 4 by Ashraphijuo, Madani, and Lavaei (2016). Let $(\mathbf{X}, \mathbf{u})$ be a primal-dual optimal pair of the SDP relaxation (2) with the parameters $(\mathbf{N}, [\langle \mathbf{M}_r, \mathbf{X}\rangle]_{r=1}^m)$. Then, the equation (4) is still satisfied after being multiplied by $\mathbf{X}'$ on the right. It takes the form $(\mathbf{N} + \sum_{r=1}^m u_r \mathbf{M}_r)\mathbf{X}' = \mathbf{0}$, which implies that Complementary Slackness holds for the SDP relaxation (2) with the parameters $(\mathbf{N}, [\langle \mathbf{M}_r, \mathbf{X}'\rangle]_{r=1}^m)$ together with primal and dual feasibility. $\square$

Due to Lemma 1, it is enough to only consider the case when a solution $\mathbf{X}^*$ of (1) to be found belongs to $\mathcal{T}_{n;k}^+$. Therefore, we make this assumption henceforth, except for the proof of Theorem 2.

Consider a matrix $\mathbf{X} \in \mathcal{T}_{n;k}^+$, the set of indexes $A \subseteq \{1, \dots n\}$ ($|A| = k$) of linearly independent columns of $\mathbf{X}$, and its Cholesky embedding $\mathcal{C}_{n;A} : \mathcal{S}_{n;A}^+ \to \mathcal{L}_{n;k}$ defined as

$$\mathcal{C}_{n;A}(\mathbf{X}) \triangleq \left[\ \mathbf{L}^{\mathrm{T}}, \quad \mathbf{L}^{-1}\mathbf{X}\left[A, \{1, \dots, n\} \setminus A\right]\ \right]^{\mathrm{T}}$$

where $\mathbf{L}\mathbf{L}^{\mathrm{T}} = \mathbf{X}[A, A]$ is the Cholesky decomposition of $\mathbf{X}[A, A]$. Using the Guttman rank additivity formula, it is possible to show that $\mathbf{X}^* = \boldsymbol{\Pi}_{n;A}^T \mathcal{C}_{n;A}(\mathbf{X}^*)\mathcal{C}_{n;A}(\mathbf{X}^*)^T \boldsymbol{\Pi}_{n;A}$. We will rewrite this factorization in the form

$$\mathbf{X}^* = \mathbf{C}\mathbf{C}^T \tag{5}$$

Let us introduce

$$\mathbf{J} = [\mathcal{V}_{n;k}(\boldsymbol{\Pi}_{n;A}\mathbf{M}_1\mathbf{C}) \ \ \dots \ \ \mathcal{V}_{n;k}(\boldsymbol{\Pi}_{n;A}\mathbf{M}_m\mathbf{C})]^T,$$

It follows directly from the form of the pushforward function of the mapping that $\mathbf{J}$ has full column rank if and only if $\mathbf{X}^*$ is not a singular point of the mapping $[\langle \mathbf{M}_r, \cdot\rangle]_{r=1}^m$ (Ashraphijuo, Madani, and Lavaei, 2016). Consider the dual vector

$$\mathbf{u}^* = -(\mathbf{J}^+)^{\mathrm{T}}\mathcal{V}_{n;k}(\mathbf{N}\mathbf{C})$$

By multiplying both sides of this equation by $\mathbf{J}^T$, it is easy to observe that if $\mathbf{J}$ has full column rank, then $\mathbf{u}^*$ solves the equation (4). Let $\mathbf{R}$ be a matrix whose columns form an orthonormal basis in the space that is orthogonal to the span of the columns of $\mathbf{C}$.

Lemma 2. *Assume that $\mathbf{X}^*$ is not a singular point of $[\langle \mathbf{M}_r, \cdot\rangle]_{r=1}^m$. The pair $(\mathbf{X}^*, \mathbf{u}^*)$ is the primal-dual optimal pair for the SDP problem (2) with the parameters $(\mathbf{N}, [\langle \mathbf{M}_r, \mathbf{X}^*\rangle]_{r=1}^m)$ if and only if $\|\mathbf{R}^T\mathbf{F}\mathbf{R}\|_2 \leq 1$ for the matrix $\mathbf{F} \in \mathbb{S}^n$ defined through the equation*

$$\mathcal{V}_{n;n}(\mathbf{F}) = \mathbf{H}(\mathbf{J}^+)^T\mathcal{V}_{n;k}(\mathbf{N}\mathbf{C})$$

Proof. Primal feasibility of $\mathbf{X}^*$ is obvious. Complementary slackness (4) is satisfied by the construction of $\mathbf{u}^*$. Consider the dual feasible point:

$$\mathbf{N} + \sum_{r=1}^{m} u_r^* \mathbf{M}_r = \mathbf{N} + \sum_{i\geq j} \mathbf{E}_{ij} f_{ij}$$

where $\mathbf{E}_{ij}$ is a matrix with the only nonzero entries equal to 1 in the (i,j) and (j,i) locations (or just (i,i)), while $f_{ij} = \mathbf{m}_{ij}^T \mathbf{u}^*$ with $\mathbf{m}_{ij} = [M_{ij}^1 \quad \ldots \quad M_{ij}^m]^T$. Note that $\mathbf{m}_{ij}^T = \mathcal{V}_{n;n}(\mathbf{E}_{ij})^T \mathbf{H}$ and

$$f_{ij} = -\mathcal{V}_{n;n}(\mathbf{E}_{ij})^T \mathbf{H} (\mathbf{J}^+)^T \mathcal{V}_{n;k}(\mathbf{NC})$$

It follows from (4) that the dual matrix has k zero eigenvalues in the subspace of the span of $\mathbf{X}^*$. We study the minimum eigenvalue of the matrix in the rest of the space:

$$\begin{aligned}\lambda_{\min}(\mathbf{N} + \sum_{r=1}^m u_r^* \mathbf{M}_r) = \\ 1 - \max_{\mathbf{v}:\|\mathbf{v}\|_2=1;\mathbf{C}^T\mathbf{v}=\mathbf{0}} 2\sum_{i\geq j} v_i v_j \mathcal{V}_{n;n}(\mathbf{E}_{ij})^T \mathcal{V}_{n;n}(\mathbf{F}) - \\ \sum_i v_i^2 \mathcal{V}_{n;n}(\mathbf{E}_{ii})^T \mathcal{V}_{n;n}(\mathbf{F}) = \\ 1 - \max_{\mathbf{v}:\|\mathbf{v}\|_2=1;\mathbf{C}^T\mathbf{v}=\mathbf{0}} \sum_{i,j=1}^n [\mathbf{v}\mathbf{v}^T]_{ij} F_{ij} = \\ 1 - \max_{\boldsymbol{\phi}\in\mathbb{R}^{n-k}:\|\boldsymbol{\phi}\|_2=1} \operatorname{trace}(\boldsymbol{\phi}^T \mathbf{R}^T \mathbf{F} \mathbf{R} \boldsymbol{\phi}),\end{aligned}$$

which is greater than or equal to zero if and only if $\|\mathbf{R}^T\mathbf{F}\mathbf{R}\|_2 \leq 1$. □

Note that if $m = \frac{n(n+1)}{2}$, then $\operatorname{rank}(\mathbf{H}) = \frac{n(n+1)}{2}$, and it is possible (e.g., using a Kronecker product representation) to show that $\mathbf{F} = \mathbf{\Pi}_{n,A}^T \mathbf{NCC}^+$, so $\|\mathbf{RFR}^T\|_2 \leq 1$, (this is expected since the only feasible point of the SDP relaxation should be the solution to the SARF problem).

Lemma 3. *If* $\mathbf{NX} = \mathbf{0}$, *then*

$$\|\mathbf{NC}\|_F \leq \sqrt{k}\|\mathbf{X} - \mathbf{X}^*\|_2$$

Proof.

$$\begin{aligned}\|\mathbf{NC}\|_F = \sqrt{\operatorname{trace}(\mathbf{NX}^*\mathbf{N})} = \sqrt{\operatorname{trace}(\mathbf{NX}^*\mathbf{X}^*\mathbf{N})} = \\ \|\mathbf{NX}^*\|_F \leq \sqrt{k}\|\mathbf{NX}^*\|_2 = \sqrt{k}\|\mathbf{N}(\mathbf{X}^* - \mathbf{X})\|_2 \leq \\ \sqrt{k}\|\mathbf{N}\|_2\|\mathbf{X}^* - \mathbf{X}\|_2 \leq \|\mathbf{X}^* - \mathbf{X}\|_2\sqrt{k}\end{aligned}$$

□

The previous lemma and the definition of $r(\mathbf{H})$ lead to the result stated below.

Corollary 2. *If* $\mathbf{NX} = \mathbf{0}$, *then then the matrices* $\mathbf{R}$ *and* $\mathbf{F}$ *in Lemma 2 satisfy the inequality*

$$\|\mathbf{RFR}^T\|_2 \leq \frac{\sqrt{k}}{r(\mathbf{H})}\|\mathbf{X}^* - \mathbf{X}\|_2$$

The above result will be used to prove Theorem 1. Define $\mathbf{\Phi} : (\mathbb{S}^n, \|\cdot\|_F) \to (\mathbb{S}^n, \|\cdot\|_F)$ such that $\mathbf{\Phi}(\mathbf{X}) = \mathbf{XX}^+$ and $\Xi_{n;k}(t) = \{\mathbf{X} \in B_{\mathbb{S}^n,\|\cdot\|_2}(1) \cap \mathbb{S}_{n;k}^+ : \sigma_k(\mathbf{X}) \geq t\}$, where $\sigma_k(\cdot)$ is the k-th largest singular value of a symmetric matrix.

Lemma 4. *For every* $t > 0$, *the operator* $\mathbf{\Phi}$ *is Lipschitz over* $\Xi_{n;k}(t)$ *with the constant* $L = \frac{1}{t}$

Proof. It is known that the projection of $\mathbf{X} \in (\mathbb{S}^n, \|\cdot\|_F)$ onto $B_{\mathbb{S}^n,\|\cdot\|_2}(1)$ is given by the matrix $\mathbf{X}'$ which can be obtained from $\mathbf{X}$ by replacing with 1 all eigenvalues that are greater than 1. Thus, $\mathbf{XX}^+$ can be viewed as the $\|\cdot\|_F$-projection of $\mathbf{X} \in \Xi_{n;k}(1)$ onto the convex set $B_{\mathbb{S}^n,\|\cdot\|_2}(1)$. Consequently, $\mathbf{\Phi}$ is Lipschitz with the constant 1 over $\Xi_{n;k}(1)$. Consider $\mathbf{X}, \mathbf{Y} \in \Xi_{n;k}(t)$:

$$\begin{aligned}\|\mathbf{\Phi}(\mathbf{X}) - \mathbf{\Phi}(\mathbf{Y})\|_F = \\ \|t^{-1}\mathbf{X}\left(t^{-1}\mathbf{X}\right)^+ - t^{-1}\mathbf{Y}\left(t^{-1}\mathbf{Y}\right)^+\|_F \leq \\ \|t^{-1}\mathbf{X} - t^{-1}\mathbf{Y}\|_F = t^{-1}\|\mathbf{X} - \mathbf{Y}\|_F\end{aligned}$$

□

Note that $\Xi_{n;1}(1) \supset \mathcal{T}_{n;1}^+$, so the function $\mathbf{\Phi}$ has the Lipschitz constant 1 over the entire $\mathcal{T}_{n;1}^+$.

Theorem 2. *The number*

$$\left(\frac{3\kappa\sqrt{k}}{r(\mathbf{H})}\frac{\max\{\|\mathbf{y}\|_2^2, 1\}}{\min\{\|\mathbf{y}\|_2^2, 1\}}\right)^{\frac{n(n+1)}{2} - m + 1}$$

is an upper bound on the number of SDP relaxations needed to find a solution with the property $\frac{\sigma_1(\cdot)}{\sigma_k(\cdot)} \leq \kappa$ *for every instance of the Semidefinite Affine Rank Feasibility problem in the class defined by* $\{\mathbf{M}_1, ..., \mathbf{M}_m\}$.

Proof. By Lemma 1, Lemma 2 and Corollary 2, given $\mathbf{A} \in \mathcal{T}_{n;k}^+$, every $\mathbf{X}^* \in \mathbb{S}_{n;k}^+$ with the property $\|(\mathbf{X}^*)^+\mathbf{X}^* - \mathbf{A}\|_2 \leq \frac{r(\mathbf{H})}{\sqrt{k}}$ is recoverable through $\mathbf{N}$ such that $\mathbf{NA} = \mathbf{0}$. Now, we aim to compute the covering number for the set of all possible solutions to the problem. Define

$$\mathbf{X}(\mathbf{x}) = x_0 \sum_{r=1}^m y_r \mathbf{M}_r + \sum_{r=1}^{\frac{n(n+1)}{2} - m} x_r \mathbf{K}_r$$

and

$$\mathcal{S} = \{\mathbf{X}(\mathbf{x}) \big| \mathbf{x} \in \mathbb{R}^{\frac{n(n+1)}{2} - m + 1}\},$$

where $\{\mathbf{K}_r \in \mathbb{S}_n\}$ are normalized as vectors that are orthogonal to $\{\mathbf{M}_r\}$ and to each other. $\mathbf{X}(\{x_0 = 1\})$ includes the set of all possible solutions to the SARF. By considering $\mathbf{X}^* = \mathbf{X}(\mathbf{x}^*)$ as the solution to be found, $\frac{\mathbf{X}^*}{\|\mathbf{X}^*\|_2} = \mathbf{X}(\frac{\mathbf{x}^*}{\|\mathbf{X}^*\|_2})$ belongs to $\mathcal{S}$. In light of Lemma 1, if $\mathbf{X}(\frac{\mathbf{x}^*}{\|\mathbf{X}^*\|_2})$ belongs to the recovery region of $\mathbf{N}$, then $\mathbf{X}^*$ belongs to it as well. Therefore, it is enough to cover $\mathcal{S} \cap B_{\mathbb{S}^n,\|\cdot\|_F}(\sqrt{k}) \cap \Xi_{n;k}(\frac{\sigma_k(\mathbf{X}^*)}{\|\mathbf{X}^*\|_2})$ with recovery regions to guarantee that $\mathbf{X}^*$ lies in one of them.

Lemma 4 yields that $\|\mathbf{X}(\frac{\mathbf{x}^*}{\|\mathbf{X}^*\|_2})\mathbf{X}(\frac{\mathbf{x}^*}{\|\mathbf{X}^*\|_2})^+ - \mathbf{A}\|_2 \leq \frac{\|\mathbf{X}^*\|_2}{\sigma_k(\mathbf{X}^*)}\|\mathbf{X}(\frac{\mathbf{x}^*}{\|\mathbf{X}^*\|_2}) - \mathbf{A}\|_F$. Therefore, it is sufficient to cover $\mathcal{S} \cap B_{\mathbb{S}^n,\|\cdot\|_F}(\sqrt{k}) \cap \Xi_{n;k}(\frac{\sigma_k(\mathbf{X}^*)}{\|\mathbf{X}^*\|_2})$ with the $\|\cdot\|_F$-balls of radius $\frac{r(\mathbf{H})\sigma_k(\mathbf{X}^*)}{\sqrt{k}\|\mathbf{X}^*\|_2}$. After noticing

$$\|\mathbf{X}(\mathbf{x})\|_F^2 = x_0^2(\|\mathbf{y}\|_2^2 - 1) + \|\mathbf{x}\|_2^2 \geq \|\mathbf{x}\|_2^2 \min\{\|\mathbf{y}\|_2^2, 1\},$$

$$\|\mathbf{X}(\mathbf{x})\|_F^2 \leq \|\mathbf{x}\|_2^2 \max\{\|\mathbf{y}\|_2^2, 1\},$$

we conclude

$$\mathcal{S} \cap B_{\mathbb{S}^n,\|\cdot\|_F}(R) \subset \mathbf{X}\left(\{\|\mathbf{x}\|_2 \leq \frac{R}{\sqrt{\min\{\|\mathbf{y}\|_2^2,1\}}}\}\right)$$

$$\mathbf{X}\left(\{\|\mathbf{x}\|_2 \leq \frac{r}{\sqrt{\max\{\|\mathbf{y}\|_2^2,1\}}}\}\right) \subset \mathcal{S} \cap B_{\mathbb{S}^n,\|\cdot\|_F}(r).$$

It is known that the covering number of the ball $B_{\mathbb{R}^{\frac{n(n+1)}{2}-m+1},\|\cdot\|_2}(R)$ with balls of radius r obeys the bound $\left(\frac{3}{r}R\right)^{\frac{n(n+1)}{2}-m+1}$ (Pajor, 1998). Applying the function $\mathbf{X}(\cdot)$ to this cover, one can obtain that $\mathcal{S} \cap B_{\mathbb{S}^n,\|\cdot\|_F}(\sqrt{k}) \cap \Xi_{n;k}(\frac{\sigma_k(\mathbf{X}^*)}{\|\mathbf{X}^*\|_2})$ belongs to the union of $\left(\frac{3\sqrt{\max\{\|\mathbf{y}\|_2^2,1\}}}{\frac{r(\mathbf{H})\sigma_k(\mathbf{X}^*)}{\sqrt{k}\|\mathbf{X}^*\|_2}}\frac{\sqrt{k}}{\sqrt{\min\{\|\mathbf{y}\|_2^2,1\}}}\right)^{\frac{n(n+1)}{2}-m+1}$ balls having radius $\frac{r(\mathbf{H})\sigma_k(\mathbf{X}^*)}{\sqrt{k}\|\mathbf{X}^*\|_2}$ in Frobenius norm. This concludes the proof. □

Theorem 3. *There is an absolute constant C such that*

$$\left(\frac{C\sqrt{k}}{r(\mathbf{H})}\right)^{k(n-k)}$$

is an upper bound on the number of SDP relaxations needed to solve every instance of the Semidefinite Affine Rank Feasibility problem in the class defined by $\{\mathbf{M}_1, ..., \mathbf{M}_m\}$.

Proof. Due to Proposition 8 by Szarek (1982), there is an absolute constant C such that the covering number of the Grassmann manifold $\mathcal{G}_{n;k}$ obeys the inequality $N_{\mathcal{G}_{n;k}}(\varepsilon) \leq \left(\frac{C\ diam(G_{n;k})}{\varepsilon}\right)^{k(n-k)}$. Similarly to the proof of Theorem 2, the diameter of the Grassmann manifold for the l_2-induced norm is equal to 1. Therefore, the proof is completed by noting that $\varepsilon = \frac{r(\mathbf{H})}{\sqrt{k}}$. □

Numerical results

In this section, we present and study a randomized algorithm for solving the SARF problem via an SDP relaxation that is based on the theoretical results of this paper. Algorithm 1 iteratively solves SDP relaxations of the problem with randomly sampled objective matrices. Under the assumption that no prior information is available about the unknown solution, we sample $\mathbf{N}$ uniformly since it belongs to the compact set $\mathcal{T}^+_{n;k}$ that is isomorphic to the Grassmann manifold $\mathcal{G}_{n;k}$.

We present experimental results on the performance of Algorithm 1 on a large set of synthetic data. The main goal of these experiments is to study the dependence between the probability of success of the convex (SDP) procedure and the number of linearly independent measurements in the problem. For a number of values of m and k, we sample a random problem with the data $\{\mathbf{M}_r\}_{r=1}^m$ and a random solution $\mathbf{X}^*$ (from which we design $y_1, ..., y_m$), and aim to solve it with Algorithm 1. After a successful ending, we sample another dataset together with a solution and then start over. After constructing and solving 300 SDP relaxations for a particular value of m, we proceed to the next value. The details on the sampling strategy are given below in the Implementation paragraph, and the results are summarized in Figure 2. It can be observed that it is easy to design an SDP that recovers the true solution even for those values of m that are much smaller than $n(n+1)/2$. Notice that the frequency of recovery decreases linearly at first, and then turns to exponential at a certain point, which appears to be a constant loosely related to n but closely connected to the value of k. This shows the existence and characterizes the behavior of the point of "complexity phase transition" of the problem from easy to hard, at least with respect to the algorithm.

Data: $\{\mathbf{M}_r, y_r\}_{r=1}^m$
Result: $\mathbf{X}$
initialization;
for $t \in \{1, \ldots, T\}$ **do**
    Sample a random $\mathbf{N}$;
    $\mathbf{X}$ = solution of (2) with $\mathbf{N}$;
    if $\text{rank}(\mathbf{X}) = k$ **then**
        return $\mathbf{X}$;
    end
end

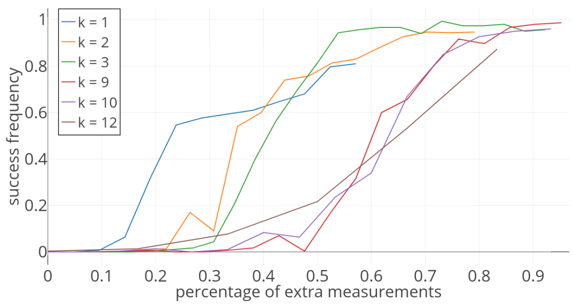
Algorithm 1: Heuristic algorithm for solving the Semidefinite Affine Rank Feasibility problem (1)

Implementation To build the region for the example in Figure 1, we randomly sample $\mathbf{M}_r \in \mathbb{S}^n$ and the matrix $\mathbf{X} \in \mathbb{S}^+_{n;k}$ following the procedure to be explained later. Afterwards, we select $\mathbf{N}$ in such a way that $\mathbf{N}\mathbf{X} = \mathbf{0}$. Let $\mathbf{Q}_0 \in \mathbb{R}^{n\times k}$ be the matrix with orthonormal columns such that $\mathbf{X} = \mathbf{Q}_0\mathbf{Q}_0^T$. In this notation, y_r is obtained as follows: $y_r = \langle \mathbf{M}_r, \mathbf{Q}(\delta_1,\delta_2)\mathbf{Q}(\delta_1,\delta_2)^T\rangle$, where $\mathbf{Q}(\delta_1,\delta_2) = \mathbf{Q}_0 + \delta_1\mathbf{E}_{00} + \delta_2\mathbf{E}_{01}$. We solve problem (2) with the parameters $(\mathbf{N}, \mathbf{y})$ for different values of δ_1 and δ_2, compare the result to the matrix $\mathbf{Q}(\delta_1,\delta_2)\mathbf{Q}(\delta_1,\delta_2)^T$, and mark the points where they almost coincide.

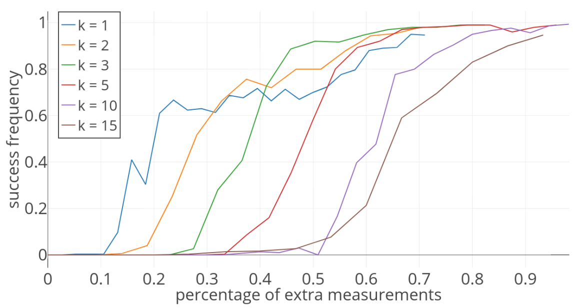
Now, let us turn to the generation procedures. $u(\{1,\ldots,n\})$ denotes the uniform distribution over the set $\{1,\ldots,n\}$; in particular, the uniform distribution over all $n \times n$ orthogonal matrices, named $O(n)$ Haar distribution, is denoted as Haar$(O(n))$.

- Generating $\{\mathbf{M}_r\}$: We denote the uniformly distributed subset of indexes as $\gamma \sim u(\{1,\ldots,\frac{n(n+1)}{2}\})$, where $|\gamma| = m$. For sampling a random matrix, we obtain $\mathbf{B}_H \sim \text{Haar}(O(\frac{n(n+1)}{2}))$ and subsample $\mathbf{H} = \mathbf{B}_H[\{1,\ldots,\frac{n(n+1)}{2}\},\gamma]$. Afterwards, $\mathbf{M}_r$ is the only symmetric matrix such that $\mathcal{V}_{n;n}(\mathbf{M}_r) = \mathbf{H}[\{1,\ldots,\frac{n(n+1)}{2}\},\{r\}]$.
- Generating $\mathbf{X}^*$: Similarly, we set up indexes $\alpha \sim u(\{1,\ldots,n\})$, where $|\alpha| = k$. Then, we obtain $\mathbf{B}_X \sim \text{Haar}(O(n))$ and set $\mathbf{Q} = \mathbf{B}_X[\{1,\ldots,n\},\alpha]$. In this notation, $\mathbf{X}^* = \mathbf{Q}\mathbf{Q}^T$

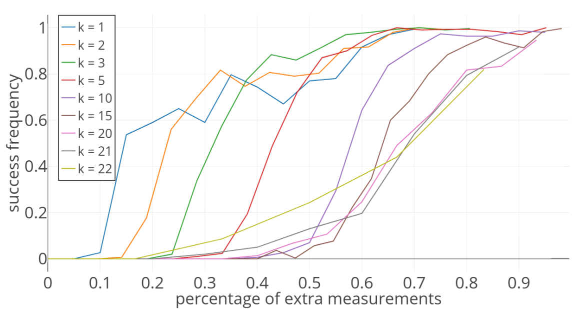
For more statistically significant results, we also use the same scheme to design a random objective matrix:



$n = 15$



$n = 20$



$n = 25$

Figure 2: These plots show the frequency of recovery for synthetic data. The x axis is the percentage (normalized) of total number of extra measurements available. This means that 0 corresponds to $m = m_{min} = nk - k(k-1)/2$, and 1 corresponds to $m = m_{max} = n(n+1)/2$. The y axis shows the probability of successful recoveries.

- Generating $\mathbf{N}$: We obtain $\beta \sim u(\{1, \ldots, n\})$, $|\beta| = n - k$ and $\mathbf{B}_N \sim \text{Haar}(O(n))$. Similarly to the previous cases, set $\mathbf{K} = \mathbf{B}_X[\{1, \ldots, n\}, \beta]$ and $\mathbf{N} = \mathbf{K}\mathbf{K}^T$.

The experiments have been scripted in Python with the use of CVXOPT as the mathematical optimization library.

Conclusion

In this work, we consider an arbitrary Semidefinite Affine Rank Feasibility problem and associate it with a class of infinity many feasibility problems. We study how many convex programs should be designed so that each member of this infinite class of feasibility problems can be solved via one of those convex programs. As a by-product, we derive the first nontrivial theoretical guarantee on the number of linearly independent measurements that is sufficient to make a generic Semidefinite Affine Rank Feasibility problem polynomial-time solvable. Besides theoretical results, we propose a randomized algorithm for solving this problem and study its performance on a large synthetic dataset. We obtain the approximate characteristic for the point of "phase transition" of a uniformly sampled problem in the space of parameters (n, m, k). These results lay the foundation for a further analysis of both theoretical and practical aspects of rank-constrained problems. Among the raised open questions, the following can be mentioned:

- How can one characterize the "phase transition" of the problem more precisely? What bounds can be set from above and from below? How would extra information and noise affect them?
- How can one exploit the underlying structure of $\mathbf{M}_r$'s more efficiently than the function $r(\mathbf{H})$?

Acknowledgment

This work was supported by the DARPA YFA Award, ARO Award, AFOSR YIP Award, NSF 1807260, and ONR N000141712933.

References

Akhriev, A.; Marecek, J.; and Simonetto, A. 2018. Pursuit of low-rank models of time-varying matrices robust to sparse and measurement noise. *arXiv preprint arXiv:1809.03550*.

Andersen, M. S.; Hansson, A.; and Vandenberghe, L. 2014. Reduced-complexity semidefinite relaxations of optimal power flow problems. *IEEE Transactions on Power Systems* 29(4):1855–1863.

Ashraphijuo, M.; Madani, R.; and Lavaei, J. 2015. Inverse function theorem for polynomial equations using semidefinite programming. In *IEEE 54th Annual Conference on Decision and Control (CDC)*, 6589–6596. IEEE.

Ashraphijuo, M.; Madani, R.; and Lavaei, J. 2016. Characterization of rank-constrained feasibility problems via a finite number of convex programs. In *IEEE 55th Conference on Decision and Control (CDC)*, 6544–6550. IEEE.

Bhojanapalli, S.; Neyshabur, B.; and Srebro, N. 2016. Global optimality of local search for low rank matrix recovery. In *Advances in Neural Information Processing Systems*, 3873–3881.

Cai, T. T., and Zhang, A. 2013. Compressed sensing and affine rank minimization under restricted isometry. *IEEE Transactions on Signal Processing* 61(13):3279–3290.

Candès, E. J., and Recht, B. 2009. Exact matrix completion via convex optimization. *Foundations of Computational mathematics* 9(6):717.

Candès, E. J.; Li, X.; Ma, Y.; and Wright, J. 2011. Robust principal component analysis? *Journal of the ACM (JACM)* 58(3):11.

Candes, E. J.; Eldar, Y. C.; Strohmer, T.; and Voroninski, V. 2015. Phase retrieval via matrix completion. *SIAM review* 57(2):225–251.

Cohen, E., and Beck, J. C. 2017. Problem difficulty and the phase transition in heuristic search. In *AAAI*, 780–786.

Cui, A.; Peng, J.; and Li, H. 2018. Exact recovery low-rank matrix via transformed affine matrix rank minimization. *Neurocomputing*.

Fazel, M. 2002. *Matrix rank minimization with applications*. Ph.D. Dissertation, PhD thesis, Stanford University.

Fukuda, M.; Kojima, M.; Murota, K.; and Nakata, K. 2001. Exploiting sparsity in semidefinite programming via matrix completion i: General framework. *SIAM Journal on Optimization* 11(3):647–674.

Ge, R.; Jin, C.; and Zheng, Y. 2017. No spurious local minima in nonconvex low rank problems: A unified geometric analysis. *arXiv preprint arXiv:1704.00708*.

Ge, R.; Lee, J. D.; and Ma, T. 2016. Matrix completion has no spurious local minimum. In *Advances in Neural Information Processing Systems*, 2973–2981.

Jain, P.; Meka, R.; and Dhillon, I. S. 2010. Guaranteed rank minimization via singular value projection. In *Advances in Neural Information Processing Systems*, 937–945.

Josz, C.; Ouyang, Y.; Zhang, R.; Lavaei, J.; and Sojoudi, S. 2018. A theory on the absence of spurious solutions for nonconvex and nonsmooth optimization. *Advances in Neural Information Processing Systems*.

Klopp, O.; Lounici, K.; and Tsybakov, A. B. 2017. Robust matrix completion. *Probability Theory and Related Fields* 169(1-2):523–564.

Kueng, R.; Rauhut, H.; and Terstiege, U. 2017. Low rank matrix recovery from rank one measurements. *Applied and Computational Harmonic Analysis* 42(1):88–116.

Madani, R.; Sojoudi, S.; Fazelnia, G.; and Lavaei, J. 2017. Finding low-rank solutions of sparse linear matrix inequalities using convex optimization. *SIAM Journal on Optimization* 27(2):725–758.

Madani, R.; Lavaei, J.; and Baldick, R. 2018. Convexification of power flow equations in the presence of noisy measurements. *to appear in IEEE Transactions on Automatic Control*.

Marianna, E.; Laurent, M.; Varvitsiotis, A.; et al. 2013. Complexity of the positive semidefinite matrix completion problem with a rank constraint. In *Discrete Geometry and Optimization*. Springer. 105–120.

Mohan, K., and Fazel, M. 2012. Iterative reweighted algorithms for matrix rank minimization. *Journal of Machine Learning Research* 13(Nov):3441–3473.

Molybog, I.; Madani, R.; and Lavaei, J. 2018. Conic optimization for robust quadratic regression: Deterministic bounds and statistical analysis. *IEEE 57th Conference on Decision and Control (CDC)*.

Morency, M. W., and Vorobyov, S. A. 2016. An algebraic approach to a class of rank-constrained semidefinite programs with applications. *arXiv preprint arXiv:1610.02181*.

Natarajan, B. K. 1995. Sparse approximate solutions to linear systems. *SIAM journal on computing* 24(2):227–234.

Pajor, A. 1998. Metric entropy of the grassmann manifold. *Convex Geometric Analysis* 34:181–188.

Recht, B.; Fazel, M.; and Parrilo, P. A. 2010. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review* 52(3):471–501.

Recht, B.; Xu, W.; and Hassibi, B. 2008. Necessary and sufficient conditions for success of the nuclear norm heuristic for rank minimization. In *IEEE 47th Conference on Decision and Control (CDC)*, 3065–3070.

Szarek, S. J. 1982. Nets of grassmann manifold and orthogonal group. In *Proceedings of research workshop on Banach space theory (Iowa City, Iowa, 1981)*, volume 169, 185.

Wei, K.; Cai, J.-F.; Chan, T. F.; and Leung, S. 2016. Guarantees of riemannian optimization for low rank matrix completion. *arXiv preprint arXiv:1603.06610*.

Xin, B., and Wipf, D. 2015. Pushing the limits of affine rank minimization by adapting probabilistic pca. In *International Conference on Machine Learning*, 419–427.

Xu, C.; Lin, Z.; and Zha, H. 2017. A unified convex surrogate for the schatten-p norm. In *AAAI*, 926–932.

Zhang, J., and Zhang, L. 2017. Efficient stochastic optimization for low-rank distance metric learning. In *AAAI*, 933–940.

Zhang, R. Y.; Josz, C.; Sojoudi, S.; and Lavaei, J. 2018. How much restricted isometry is needed in nonconvex matrix recovery? *Advances in Neural Information Processing Systems*.

Zhang, Y.; Madani, R.; and Lavaei, J. 2018. Conic relaxations for power system state estimation with line measurements. *IEEE Transactions on Control of Network Systems* 5(3):1193 – 1205.