Sharp Restricted Isometry Bounds for the Inexistence of Spurious Local Minima in Nonconvex Matrix Recovery

Richard Y. Zhang
University of California, Berkeley
ryz@alum.mit.edu

Somayeh Sojoudi
University of California, Berkeley
sojoudi@berkeley.edu

Javad Lavaei
University of California, Berkeley
lavaei@berkeley.edu

January 7, 2019

Abstract

Nonconvex matrix recovery is known to contain *no spurious local minima* under a restricted isometry property (RIP) with a sufficiently small RIP constant δ. If δ is too large, however, then counterexamples containing spurious local minima are known to exist. In this paper, we introduce a proof technique that is capable of establishing *sharp* thresholds on δ to guarantee the inexistence of spurious local minima. Using the technique, we prove that in the case of a rank-1 ground truth, an RIP constant of $\delta < 1/2$ is both necessary and sufficient for exact recovery from any arbitrary initial point (such as a random point). We also prove a local recovery result: given an initial point x_0 satisfying $f(x_0) \le (1-\delta)^2 f(0)$, any descent algorithm that converges to second-order optimality guarantees exact recovery.

1 Introduction

The *low-rank matrix recovery* problem seeks to recover an unknown $n\times n$ ground truth matrix $M^\star$ of low-rank $r \ll n$ from m linear measurements of $M^\star$. The problem naturally arises in recommendation systems [Rennie and Srebro, 2005] and clustering algorithms [Amit et al., 2007]—often under the names of matrix completion and matrix sensing—and also finds engineering applications in phase retrieval [Candes et al., 2013] and power system state estimation [Zhang et al., 2018b].

In the symmetric, noiseless variant of low-rank matrix recovery, the ground truth $M^\star$ is taken to be positive semidefinite (denoted as $M^\star \succeq 0$), and the m linear measurements are made without error, as in

$$b \equiv \mathcal{A}(M) \quad \text{where} \quad \mathcal{A}(M) = \begin{bmatrix} \langle A_1, M\rangle & \cdots & \langle A_m, M\rangle \end{bmatrix}^T. \tag{1}$$

To recover $M^\star$ from b, the standard approach in the machine learning community is to factor a candidate M into its low-rank factors XX^T, and to solve a nonlinear least-squares problem on X using a local search algorithm (usually stochastic gradient descent):

$$\underset{X \in \mathbb{R}^{n\times r}}{\text{minimize}} \quad f(X) \equiv \|\mathcal{A}(XX^T) - b\|^2. \tag{2}$$

The function f is nonconvex, so a "greedy" local search algorithm can become stuck at a spurious local minimum, especially if a random initial point is used. Despite this apparent risk of failure, the nonconvex approach remains both widely popular as well as highly effective in practice.

Recently, Bhojanapalli et al. [2016b] provided a rigorous theoretical justification for the empirical success of local search on problem (2). Specifically, they showed that the problem contains *no spurious local minima* under the assumption that $\mathcal{A}$ satisfies the *restricted isometry property* (RIP) of Recht et al. [2010] with a sufficiently small constant. The nonconvex problem is easily solved using local search algorithms because every local minimum is also a global minimum.

Definition 1 (Restricted Isometry Property)**.** The linear map $\mathcal{A} : \mathbb{R}^{n\times n} \to \mathbb{R}^m$ is said to satisfy δ-RIP if there is constant $\delta \in [0, 1)$ such that

$$(1-\delta)\|M\|_F^2 \le \|\mathcal{A}(M)\|^2 \le (1+\delta)\|M\|_F^2 \tag{3}$$

holds for all $M \in \mathbb{R}^{n\times n}$ satisfying $\operatorname{rank}(M) \le 2r$.

Theorem 2 (Bhojanapalli et al., 2016b, Ge et al., 2017)**.** *Let $\mathcal{A}$ satisfy δ-RIP with $\delta < 1/5$. Then, (2) has no spurious local minima:*

$$\nabla f(X) = 0, \quad \nabla^2 f(X) \succeq 0 \quad \iff \quad XX^T = M^\star.$$

Hence, any algorithm that converges to a second-order critical point is guaranteed to recover $M^\star$ exactly.

While Theorem 2 says that an RIP constant of $\delta < 1/5$ is *sufficient* for exact recovery, Zhang et al. [2018a] proved that $\delta < 1/2$ is *necessary.* Specifically, they gave a counterexample satisfying 1/2-RIP that causes randomized stochastic gradient descent to fail 12% of the time. A number of previous authors have attempted to close the gap between sufficiency and necessity, including Bhojanapalli et al. [2016b], Ge et al. [2017], Park et al. [2017], Zhang et al. [2018a], Zhu et al. [2018]. In this paper, we prove that in the rank-1 case, an RIP constant of $\delta < 1/2$ is both *necessary* and *sufficient* for exact recovery.

Once the RIP constant exceeds $\delta \ge 1/2$, global guarantees are no longer possible. Zhang et al. [2018a] proved that counterexamples exist *generically*: almost every choice of $x, z \in \mathbb{R}^n$ generates an instance of nonconvex recovery satisfying RIP with x as a spurious local minimum and $M^\star = zz^T$ as ground truth. In practice, local search may continue to work well, often with a 100% success rate as if spurious local minima do not exist. However, the inexistence of spurious local minima can no longer be assured.

Instead, we turn our attention to local guarantees, based on good initial guesses that often arise from domain expertise, or even chosen randomly. Given an initial point x_0 satisfies $f(x_0) \le (1-\delta)^2\|M^\star\|_F^2$ where δ is the RIP constant and $M^\star = zz^T$ is a rank-1 ground truth, we prove that a *descent* algorithm that converges to second-order optimality is guaranteed to recover the ground truth. Examples of such algorithms include randomized full-batch gradient descent [Jin et al., 2017] and trust-region methods [Conn et al., 2000, Nesterov and Polyak, 2006].

2 Main Results

Our main contribution in this paper is a proof technique capable of establishing RIP thresholds that are both *necessary* and *sufficient* for exact recovery. The key idea is to disprove the counterfactual. To prove for some $\lambda \in [0,1)$ that "λ-RIP implies no spurious local minima", we instead establish the inexistence of a counterexample that admits a spurious local minimum despite satisfying λ-RIP. In particular, if $\delta^\star$ is the *smallest* RIP constant associated with a counterexample, then any $\lambda < \delta^\star$ cannot admit a counterexample (or it would contradict the definition of $\delta^\star$ as the smallest RIP constant). Accordingly, $\delta^\star$ is precisely the sharp threshold needed to yield a necessary and sufficient recovery guarantee.

The main difficulty with the above line of reasoning is the need to optimize over the set of counterexamples. Indeed, verifying RIP for a fixed operator $\mathcal{A}$ is already NP-hard in general [Tillmann and Pfetsch, 2014], so it is reasonable to expect that optimizing over the set of RIP operators is at least NP-hard. Surprisingly, this is not the case. Consider finding the smallest RIP constant associated with a counterexample with *fixed* ground truth $M^\star = ZZ^T$ and *fixed* spurious point X:

$$\begin{aligned} \delta(X,Z) \quad \equiv \quad & \underset{\mathcal{A}}{\text{minimum}} \quad \delta \\ & \text{subject to} \quad f(X) = \frac{1}{2}\|\mathcal{A}(XX^T - ZZ^T)\|^2 \\ & \qquad\qquad\quad \nabla f(X) = 0, \quad \nabla^2 f(X) \succeq 0 \\ & \qquad\qquad\quad \mathcal{A} \text{ satisfies } \delta\text{-RIP}. \end{aligned} \tag{4}$$

In Section 5, we reformulate problem (4) into a *convex* linear matrix inequality (LMI) optimization, and prove that the reformulation is *exact* (Theorem 8). Accordingly, we can evaluate $\delta(X,Z)$ to arbitrary precision in polynomial time by solving an LMI using an interior-point method.

In the rank $r = 1$ case, the LMI reformulation is sufficiently simple that it can be relaxed and then solved in closed-form (Theorem 13). This yields a lower-bound $\delta_{\text{lb}}(x,z) \le \delta(x,z)$ that we optimize over all spurious choices of $x \in \mathbb{R}^n$ to prove that $\delta^\star \ge 1/2$. Given that $\delta^\star \le 1/2$ due to the counterexample of Zhang et al. [2018a], we must actually have $\delta^\star = 1/2$.

Theorem 3 (Global guarantee)**.** *Let $r = \operatorname{rank}(M^\star) = 1$, let $\mathcal{A}$ satisfy δ-RIP, and define $f(x) = \|\mathcal{A}(xx^T - M^\star)\|^2$.*

- *If $\delta < 1/2$, then f has no spurious local minima:*

$$\nabla f(x) = 0, \quad \nabla^2 f(x) \succeq 0 \qquad \iff \qquad xx^T = M^\star.$$

- *If $\delta \ge 1/2$, then there exists a counterexample $\mathcal{A}^\star$ satisfying δ-RIP, but whose $f^\star(x) = \|\mathcal{A}^\star(xx^T - M^\star)\|^2$ admits a spurious point $x \in \mathbb{R}^n$ satisfying:*

$$f(x) = \frac{3}{4}\|M^\star\|_F^2, \qquad \nabla f(x) = 0, \qquad \nabla^2 f(x) = 8\|M^\star\|_F\, xx^T.$$

We can also optimize $\delta_{\text{lb}}(x,z)$ over spurious choices $x \in \mathbb{R}^n$ within an ϵ-neighborhood of the ground truth. The resulting guarantee is applicable to much larger RIP constants δ, including those arbitrarily close to one.

Theorem 4 (Local guarantee)**.** *Let $r = \operatorname{rank}(M^\star) = 1$, and let $\mathcal{A}$ satisfy δ-RIP. If*

$$\delta < \left(1 - \frac{\epsilon^2}{2(1-\epsilon)}\right)^{1/2} \qquad \text{where } 0 \le \epsilon \le \frac{\sqrt{5}-1}{2}$$

then $f(x) = \|\mathcal{A}(xx^T - M^\star)\|^2$ has no spurious local minima within an ϵ-neighborhood of the solution:

$$\nabla f(x) = 0, \quad \nabla^2 f(x) \succeq 0, \quad \|xx^T - M^\star\|_F \le \epsilon \|M^\star\|_F \quad \iff \quad xx^T = M^\star.$$

Theorem 4 gives an RIP-based exact recovery guarantee for *descent* algorithms, such as randomized full-batch gradient descent [Jin et al., 2017] and trust-region methods [Conn et al., 2000, Nesterov and Polyak, 2006], that generate a sequence of iterates $x_1, x_2, \ldots, x_k$ from an initial guess x_0 with each iterate no worse than the one before:

$$f(x_k) \le \cdots \le f(x_2) \le f(x_1) \le f(x_0). \tag{5}$$

Heuristically, it also applies to nondescent algorithms, like stochastic gradient descent and Nesterov's accelerated gradient descent, under the mild assumption that the final iterate x_k is no worse than the initial guess x_0, as in $f(x_k) \le f(x_0)$.

Corollary 5. *Let $r = \operatorname{rank}(M^\star) = 1$, and let $\mathcal{A}$ satisfy δ-RIP. If $x_0 \in \mathbb{R}^n$ satisfies*

$$f(x_0) < (1-\delta)\epsilon^2 f(0) \qquad \text{where } \epsilon = \min\left\{\sqrt{1-\delta^2}, (\sqrt{5}-1)/2\right\},$$

where $f(x) = \|\mathcal{A}(xx^T - M^\star)\|^2$, then the sublevel set defined by x_0 contains no spurious local minima:

$$\nabla f(x) = 0, \quad \nabla^2 f(x) \succeq 0, \quad f(x) \le f(x_0) \quad \iff \quad xx^T = M^\star.$$

When the RIP constant satisfies $\delta \ge 0.787$, Corollary 5 guarantees exact recovery from an initial point x_0 satisfying $f(x_0) < (1-\delta)^2 f(0)$. In practice, such an x_0 can often be found using a spectral initializer [Keshavan et al., 2010a, Jain et al., 2013, Netrapalli et al., 2013, Candes et al., 2015, Chen and Candes, 2015]. If δ is not too close to one, then even a random point may suffice with a reasonable probability (see the related discussion by Goldstein and Studer [2018]).

In the rank-r case with $r > 1$, our proof technique continues to work, but the LMI reformulation becomes very difficult to solve in closed-form. Nevertheless, we can evaluate $\delta(X, Z)$ numerically using an interior-point method, and then heuristically optimize over the spurious point X and ground truth $M^\star = ZZ^T$. Doing this in Section 8, we obtain empirical evidence that higher-rank have larger RIP thresholds, and so are in a sense "easier" to solve.

3 Related work

3.1 No spurious local minima in matrix completion

Exact recovery guarantees like Theorem 2 have also been established for "harder" choices of $\mathcal{A}$ that do not satisfy RIP over its entire domain. In particular, the matrix completion problem has sparse measurement matrices $A_1, \dots, A_m$, with each containing just a single nonzero element. In this case, the RIP-like condition $\|\mathcal{A}(M)\|^2 \approx \|M\|_F^2$ holds only when M is both low-rank and sufficiently dense; see the discussion by Candès and Recht [2009]. Nevertheless, Ge et al. [2016] proved a similar result to Theorem 2 by adding a regularizing term to the objective.

Our recovery results are developed for the classical form of RIP—a much stronger notion than the RIP-like condition satisfied by matrix completion. Intuitively, if exact recovery cannot be guaranteed under standard RIP, then exact recovery under a weaker notion would seem unlikely. It remains future work to make this argument precise, and to extend our proof technique to these "harder" choices of $\mathcal{A}$.

3.2 Noisy measurements and nonsymmetric ground truth

Recovery guarantees for the noisy and/or nonsymmetric variants of nonconvex matrix recovery typically require a smaller RIP constant than the symmetric, noiseless case. For example, Bhojanapalli et al. [2016b] proved that the symmetric, zero-mean, σ^2-variance Gaussian noise case requires a rank-$4r$ RIP constant of $\delta < 1/10$ to recover an σ-accurate solution X satisfying $\|XX^T - M^\star\|_F \le 20\sigma\sqrt{(\log n)/m}$. Also, Ge et al. [2017] proved that the nonsymmetric, noiseless case requires a rank-$2r$ RIP constant $\delta < 1/10$ for exact recovery. By comparison, the symmetric, noiseless case requires only a rank-$2r$ RIP constant of $\delta < 1/5$ for exact recovery.

The main goal of this paper is to develop a proof technique capable of establishing *sharp* RIP thresholds for exact recovery. As such, we have focused our attention on the symmetric, noiseless case. While our technique can be easily modified to accommodate for the nonsymmetric, noisy case, the *sharpness* of the technique (via Theorem 8) may be lost. Whether an exact convex reformulation exists for the nonsymmetric, noisy case is an open question, and the subject of important future work.

3.3 Approximate second-order points and strict saddles

Existing "no spurious local minima" results [Bhojanapalli et al., 2016b, Ge et al., 2017] guarantee that satisfying second-order optimality to ϵ-accuracy will yield a point within an ϵ-neighborhood of the solution:

$$\|\nabla f(X)\| \le C_1\epsilon, \qquad \nabla^2 f(X) \succeq -C_2\sqrt{\epsilon} I \qquad \Longleftrightarrow \qquad \|XX^T - M^\star\|_F \le \epsilon.$$

Such a condition is often known as "strict saddle" [Ge et al., 2015]. The associated constants $C_1, C_2 > 0$ determine the rate at which gradient methods can converge to an ϵ-accurate solution [Du et al., 2017, Jin et al., 2017].

The proof technique presented in this paper can be extended in a straightforward way to the strict saddle condition. Specifically, we replace all instances of $\nabla f(X) = 0$, $\nabla^2 f(X) \succeq 0$, and $XX^T \ne M^\star$ with $\|\nabla f(X)\| \le C_1\epsilon$, $\nabla^2 f(X) \succeq C_2\sqrt{\epsilon} I$, and $\|XX^T - M^\star\| > \epsilon$ in Section 5, and derive a suitable version of Theorem 8. However, the resulting reformulation can no longer be solved in closed form, so it becomes difficult to extend the guarantees in Theorem 3 and Theorem 4. Nevertheless, quantifying its asymptotic behavior may yield valuable insights in understanding the optimization landscape.

3.4 Special initialization schemes

Our local recovery result is reminiscent of classic exact recovery results based on placing an initial point sufficiently close to the global optimum. Most algorithms use the spectral initializer to chose the initial point [Keshavan et al., 2010a,b, Jain et al., 2013, Netrapalli et al., 2013, Candes et al., 2015, Chen and Candes, 2015, Zheng and Lafferty, 2015, Zhao et al., 2015, Bhojanapalli et al., 2016a, Sun and Luo, 2016, Sanghavi et al., 2017, Park et al., 2018], although other initializers have also been proposed [Wang et al., 2018, Chen et al., 2018, Mondelli and Montanari, 2018]. Our result differs from prior work in being completely agnostic to the specific application and the initializer. First, it requires only a suboptimality bound $f(x_0) \le (1-\delta)^2 f(0)$ to be satisfied by the initial point x_0. Second, its sole parameter is the RIP constant δ, so issues of sample complexity are implicitly resolved in a universal way for different measurement ensembles. On the other hand, the result is not directly applicable to problems that only approximately satisfy RIP, including matrix completion.

3.5 Comparison to convex recovery

Classical theory for the low-rank matrix recovery problem is based on a quadratic lift: replacing XX^T in (2) by a convex term $M \succeq 0$, and augmenting the objective with a trace penalty $\lambda \cdot \mathrm{tr}(M)$ to induce a low-rank solution [Candès and Recht, 2009, Recht et al., 2010, Candès and Tao, 2010, Candes and Plan, 2011, Candes et al., 2013]. The convex approach also enjoys RIP-based exact recovery guarantees: in the noiseless case, Cai and Zhang [2013] proved that $\delta \le 1/2$ is sufficient, while the counterexample of Wang and Li [2013] shows that $\delta \le 1/\sqrt{2}$ is necessary. While convex recovery may be able to solve problems with larger RIP constants than nonconvex recovery, it is also considerably more expensive. In practice, convex recovery is seldom used for large-scale datasets with n on the order of thousands to millions.

Recently, several authors have proposed *non-lifting* convex relaxations, motivated by the desire to avoid squaring the number of variables in the classic quadratic lift. In particular, we mention the PhaseMax method studied by Bahmani and Romberg [2017] and Goldstein and Studer [2018], which avoids the need to square the number of variables when both the measurement matrices $A_1, \ldots, A_m$ and the ground truth $M^\star$ are rank-1. These methods also require a good initial guess as an input, and so are in a sense very similar to nonconvex recovery.

4 Preliminaries

4.1 Notation

Lower-case letters are vectors and upper-case letters are matrices. The sets $\mathbb{R}^{n\times n} \supset \mathbb{S}^n$ are the space of $n \times n$ real matrices and real symmetric matrices, and $\langle X, Y\rangle \equiv \mathrm{tr}(X^T Y)$ and $\|X\|_F^2 \equiv \langle X, X\rangle$ are the Frobenius inner product and norm. We write $M \succeq 0$ (resp. $M \succ 0$) to mean that M is positive semidefinite (resp. positive definite), and $M \succeq S$ to denote $M - S \succeq 0$ (resp. $M \succ S$ to denote $M - S \succ 0$).

Throughout the paper, we use $X \in \mathbb{R}^{n\times r}$ (resp. $x \in \mathbb{R}^n$) to refer to any candidate point, and $M^\star = ZZ^T$ (resp. $M^\star = zz^T$) or to refer to a rank-r (resp. rank-1) factorization of the ground truth $M^\star$. The vector $\mathbf{e}$ and matrix $\mathbf{X}$ are defined in (11). We also denote the optimal value of the nonconvex problem (15) as $\delta(X, Z)$, and later show it to be equal to the optimal value of the convex problem (21) denoted as $\mathrm{LMI}(X, Z)$.

4.2 Basic definitions

The **vectorization** operator stacks the columns of an $m \times n$ matrix A into a single column vector:

$$\mathrm{vec}\,(A) = \begin{bmatrix} A_{1,1} & \cdots & A_{m,1} & A_{1,2} & \cdots & A_{m,2} & \cdots & A_{1,n} & \cdots & A_{m,n} \end{bmatrix}^T.$$

It defines an isometry between the $m \times n$ matrices A, B and their mn underlying degrees of freedom $\mathrm{vec}\,(A), \mathrm{vec}\,(B)$:

$$\langle A, B \rangle \equiv \mathrm{tr}(A^T B) = \mathrm{vec}\,(A)^T \mathrm{vec}\,(B) \equiv \langle \mathrm{vec}\,(A), \mathrm{vec}\,(B) \rangle.$$

The **matricization** operator is the inverse of vectorization, meaning that $A = \mathrm{mat}(a)$ if and only if $a = \mathrm{vec}\,(A)$.

The **Kronecker product** between the $m \times n$ matrix A and the $p \times q$ matrix B is the $mp \times pq$ matrix defined

$$A \otimes B = \begin{bmatrix} A_{1,1}B & \cdots & A_{1,n}B \\ \vdots & \ddots & \vdots \\ A_{m,1}B & \cdots & A_{m,n}B \end{bmatrix}$$

to satisfy the Kronecker identity

$$\mathrm{vec}\,(AXB^T) = (B \otimes A)\,\mathrm{vec}\,(X).$$

The **orthogonal basis** of a given $m \times n$ matrix A (with $m \geq n$) is a matrix $P = \mathrm{orth}(A)$ comprising $\mathrm{rank}\,(A)$ orthonormal columns of length-m that span $\mathrm{range}(A)$:

$$P = \mathrm{orth}(A) \qquad \Longleftrightarrow \qquad PP^T A = A, \qquad P^T P = I_{\mathrm{rank}\,(A)}.$$

We can compute P using either a rank-revealing QR factorization [Chan, 1987] or a (thin) singular value decomposition [Golub and Van Loan, 1996, p. 254] in $O(mn^2)$ time and $O(mn)$ memory.

4.3 Global optimality and local optimality

Given a choice of $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ and the rank-r ground truth $M^\star \succeq 0$, we define the nonconvex objective

$$f : \mathbb{R}^{n \times r} \to \mathbb{R} \qquad \text{such that} \qquad f(X) = \frac{1}{2}\|\mathcal{A}(XX^T - M^\star)\|^2. \tag{6}$$

If the point X attains $f(X) = 0$, then we call it a *globally minimum*; otherwise, we call it a *spurious* point. If $\mathcal{A}$ satisfies δ-RIP, then X is a global minimum if and only if $XX^T = M^\star$ [Recht et al., 2010, Theorem 3.2].

The point X is said to be a *local minimum* if $f(X) \le f(X')$ holds for all X' within a local neighborhood of X. If X is a local minimum, then it must satisfy the second-order *necessary* condition for local optimality:

$$\nabla f(X) = 0, \qquad \nabla^2 f(X) \succeq 0. \tag{7}$$

Conversely, a point X satisfying (7) is called a *second-order critical point*, and can be either a local minimum or a saddle point. It is worth emphasizing that local search algorithms can only guarantee convergence to a second-order critical point, and not necessarily a local minimum; see Ge et al. [2015], Lee et al. [2016], Jin et al. [2017], Du et al. [2017] for the literature on gradient methods, and Conn et al. [2000], Nesterov and Polyak [2006], Cartis et al. [2012], Boumal et al. [2018] for the literature on trust-region methods.

If a point X satisfies the second-order *sufficient* condition for local optimality (with $\mu > 0$):

$$\nabla f(X) = 0, \qquad \nabla^2 f(X) \succeq \mu I \tag{8}$$

then it is guaranteed to be a local minimum. However, it is also possible for X to be a local minimum without satisfying (8). Indeed, certifying X to be a local minimum is NP-hard in the worst case [Murty and Kabadi, 1987]. Hence, the finite gap between necessary and sufficient conditions for local optimality reflects the inherent hardness of the problem.

4.4 Explicit expressions for $\nabla f(X)$ and $\nabla^2 f(X)$

Define $f(X)$ as the nonlinear least-squares objective shown in (6). While not immediately obvious, both the gradient $\nabla f(X)$ and the Hessian $\nabla^2 f(X)$ are *linear* with respect to the the *kernel* operator $\mathcal{H} \equiv \mathcal{A}^T \mathcal{A}$. To show this, we define the matrix representation of the operator $\mathcal{A}$

$$\mathbf{A} = \begin{bmatrix} \mathrm{vec}\,(A_1) & \mathrm{vec}\,(A_2) & \cdots & \mathrm{vec}\,(A_m) \end{bmatrix}^T, \tag{9}$$

which satisfies

$$\mathcal{A}(M) = \begin{bmatrix} \langle A_1, M \rangle \\ \vdots \\ \langle A_m, M \rangle \end{bmatrix} = \begin{bmatrix} \mathrm{vec}\,(A_1)^T \mathrm{vec}\,(M) \\ \vdots \\ \mathrm{vec}\,(A_m)^T \mathrm{vec}\,(M) \end{bmatrix} = \begin{bmatrix} \mathrm{vec}\,(A_1)^T \\ \vdots \\ \mathrm{vec}\,(A_m)^T \end{bmatrix} \mathrm{vec}\,(M) = \mathbf{A}\,\mathrm{vec}\,(M).$$

Then, some linear algebra reveals

$$f(X) = \frac{1}{2} \mathbf{e}^T \mathbf{A}^T \mathbf{A} \mathbf{e}, \tag{10a}$$

$$\nabla f(X) = \mathbf{X}^T \mathbf{A}^T \mathbf{A} \mathbf{e}, \tag{10b}$$

$$\nabla^2 f(X) = 2 \cdot [I_r \otimes \mathrm{mat}(\mathbf{A}^T \mathbf{A} \mathbf{e})] + \mathbf{X}^T \mathbf{A}^T \mathbf{A} \mathbf{X}, \tag{10c}$$

where $\mathbf{e}$ and $\mathbf{X}$ are defined with respect to X and $M^\star$ to satisfy

$$\mathbf{e} = \mathrm{vec}\,(XX^T - M^\star), \tag{11a}$$

$$\mathbf{X}\,\mathrm{vec}\,(U) = \mathrm{vec}\,(XU^T + UX^T) \qquad \forall U \in \mathbb{R}^{n \times r}. \tag{11b}$$

(Note that $\mathbf{X}$ is simply the Jacobian of $\mathbf{e}$ with respect to X.) Clearly, $f(X)$, $\nabla f(X)$, and $\nabla^2 f(X)$ are all linear with respect to $\mathbf{H} = \mathbf{A}^T\mathbf{A}$. In turn, $\mathbf{H}$ is simply the matrix representation of the kernel operator $\mathcal{H}$.

As an immediate consequence noted by Zhang et al. [2018a], both the second-order necessary condition (7) and the second-order sufficient condition (8) for local optimality are *linear matrix inequalities* (LMIs) with respect to $\mathbf{H}$. In particular, this means that finding an instance of (2) with a fixed $M^\star$ as the ground truth and X as a spurious local minimum is a *convex* optimization problem:

$$\begin{aligned} &\text{find } \mathcal{A} \\ &\text{such that } f(X) = \frac{1}{2}\|\mathcal{A}(XX^T - M^\star)\|^2, \\ &\quad \nabla f(X) = 0, \\ &\quad \nabla^2 f(X) \succeq \mu I. \end{aligned} \quad \iff \quad \begin{aligned} &\text{find } \mathbf{H} \succeq 0 \\ &\text{such that } \mathbf{X}^T\mathbf{H}\mathbf{e} = 0, \\ &\quad 2 \cdot [I_r \otimes \mathrm{mat}(\mathbf{A}^T\mathbf{A}\mathbf{e})] \\ &\quad\quad + \mathbf{X}^T\mathbf{A}^T\mathbf{A}\mathbf{X} \succeq \mu I. \end{aligned} \tag{12}$$

Given a feasible point $\mathbf{H}$, we compute an $\mathbf{A}$ satisfying $\mathbf{H} = \mathbf{A}^T\mathbf{A}$ using Cholesky factorization or an eigendecomposition. Then, matricizing each row of $\mathbf{A}$ recovers the matrices $A_1, \dots, A_m$ implementing a feasible choice of $\mathcal{A}$.

5 Main idea: The inexistence of counterexamples

At the heart of this work is a simple argument by the inexistence of counterexamples. To illustrate the idea, consider making the following claim for a fixed choice of $\lambda \in [0, 1)$ and $X, Z \in \mathbb{R}^{n\times r}$:

If $\mathcal{A}$ satisfies λ-RIP, then X is *not* a spurious second-order critical point for the nonconvex recovery of $M^\star = ZZ^T$. (13)

The claim is refuted by a counterexample: an instance of (2) satisfying λ-RIP with ground truth $M^\star = ZZ^T$ and spurious local minimum X. The problem of finding a counterexample is a nonconvex feasibility problem:

$$\begin{aligned} \text{find} \quad & \mathcal{A} \\ \text{such that} \quad & f(X) = \frac{1}{2}\|\mathcal{A}(XX^T - ZZ^T)\|^2 \\ & \nabla f(X) = 0, \quad \nabla^2 f(X) \succeq 0 \\ & \mathcal{A} \text{ satisfies } \delta\text{-RIP}. \end{aligned} \tag{14}$$

If problem (14) is *feasible* for $\delta = \lambda$, then any feasible point is a counterexample that refutes the claim (13). However, if problem (14) is *infeasible* for $\delta = \lambda$, then counterexamples do not exist, so we must accept the claim (13) at face value. In other words, *the inexistence of counterexamples is proof for the original claim.*

The same argument can be posed in an optimization form. Instead of finding any arbitrary counterexample, we will look for the counterexample with the *smallest* RIP constant

$$\begin{aligned} \delta(X,Z) \quad \equiv \quad & \underset{\mathcal{A}}{\text{minimum}} \quad \delta \\ & \text{subject to} \quad f(X) = \frac{1}{2}\|\mathcal{A}(XX^T - ZZ^T)\|^2 \\ & \qquad\qquad \nabla f(X) = 0, \quad \nabla^2 f(X) \succeq 0 \\ & \qquad\qquad \mathcal{A} \text{ satisfies } \delta\text{-RIP.} \end{aligned} \tag{15}$$

Suppose that problem (15) attains its minimum at $\mathcal{A}^\star$. If $\lambda \geq \delta(X,Z)$, then the minimizer $\mathcal{A}^\star$ is a counterexample that refutes the claim (13). On the other hand, if $\lambda < \delta(X,Z)$, then problem (14) is infeasible for $\delta = \lambda$, so counterexamples do not exist, so the claim (13) must be true.

Repeating these arguments over all choices of X and Z yields the following global recovery guarantee.

Lemma 6 (Sharp global guarantee)**.** *Suppose that problem (15) attains its minimum of $\delta(X,Z)$. Define $\delta^\star$ as in*

$$\delta^\star \quad \equiv \quad \underset{X,Z\in\mathbb{R}^{n\times r}}{\text{infimum}} \quad \delta(X,Z) \quad \text{subject to} \quad XX^T \neq ZZ^T. \tag{16}$$

If $\mathcal{A}$ satisfies λ-RIP with $\lambda < \delta^\star$, then $f(X) = \|\mathcal{A}(XX^T - M^\star)\|^2$ with ground truth $M^\star \succeq 0$ and $\operatorname{rank}(M^\star) \leq r$ *satisfies:*

$$\nabla f(X) = 0, \quad \nabla^2 f(X) \succeq 0 \quad \Longleftrightarrow \quad XX^T = M^\star. \tag{17}$$

Moreover, if there exist $X^\star, Z^\star$ such that $\delta^\star = \delta(X^\star, Z^\star)$, then the threshold $\delta^\star$ is sharp.

Proof. To prove (17), we simply prove the claim (13) for $\lambda < \delta^\star$ and every possible choice of $X, Z \in \mathbb{R}^{n\times r}$. Indeed, if $XX^T = ZZ^T$, then X is not a spurious point (as it is a global minimum), whereas if $XX^T \neq ZZ^T$, then $\lambda < \delta^\star \leq \delta(X,Z)$ proves the inexistence of a counterexample. Sharpness follows because the minimum $\delta^\star = \delta(X^\star, Z^\star)$ is attained by the minimizer $\mathcal{A}^\star$ that refutes the claim (13) for all $\lambda \geq \delta^\star$ and $X = X^\star$ and $Z = Z^\star$. □

Repeating the same arguments over an ϵ-local neighborhood of the ground truth yields the following *local* recovery guarantee.

Lemma 7 (Sharp local guarantee)**.** *Suppose that problem (15) attains its minimum of $\delta(X,Z)$. Given $\epsilon > 0$, define $\delta^\star(\epsilon)$ as in*

$$\delta^\star(\epsilon) \quad \equiv \quad \underset{X,Z\in\mathbb{R}^{n\times r}}{\text{infimum}} \quad \delta(X,Z) \quad \text{subject to} \quad XX^T \neq ZZ^T, \ \|XX^T - ZZ^T\|_F \leq \epsilon \|ZZ^T\|_F. \tag{18}$$

If $\mathcal{A}$ satisfies λ-RIP with $\lambda < \delta^\star(\epsilon)$, then $f(X) = \|\mathcal{A}(XX^T - M^\star)\|^2$ with ground truth $M^\star \succeq 0$ and $\operatorname{rank}(M^\star) \leq r$ *satisfies:*

$$\nabla f(X) = 0, \quad \nabla^2 f(X) \succeq 0, \quad \|XX^T - ZZ^T\|_F \leq \epsilon \|ZZ^T\|_F \iff XX^T = M^\star. \tag{19}$$

Moreover, if there exist $X^\star, Z^\star$ such that $\delta^\star = \delta(X^\star, Z^\star)$, then the threshold $\delta^\star$ is sharp.

Our main difficulty with Lemma 6 and Lemma 7 is the evaluation of $\delta(X,Z)$. Indeed, verifying δ-RIP for a fixed $\mathcal{A}$ is already NP-hard in general [Tillmann and Pfetsch, 2014], so it is reasonable to expect that solving an optimization problem (15) with a δ-RIP constraint would be at least NP-hard. Instead, Zhang et al. [2018a] suggests replacing the δ-RIP constraint with a convex *sufficient* condition, obtained by enforcing the RIP inequality (3) over *all* $n \times n$ matrices (and not just rank-$2r$ matrices):

$$(1-\delta)\|M\|_F^2 \le \|\mathcal{A}(M)\|^2 \le (1+\delta)\|M\|_F^2 \qquad \forall M \in \mathbb{R}^{n\times n}. \tag{20}$$

The resulting problem is a linear matrix inequality (LMI) optimization over the kernel operator $\mathcal{H} = \mathcal{A}^T\mathcal{A}$ that yields an upper-bound on $\delta(X,Z)$:

$$\begin{aligned} \mathrm{LMI}(X,Z) \quad \equiv \quad & \underset{\mathcal{H}=\mathcal{A}^T\mathcal{A}}{\text{minimum}} \quad \delta \\ & \text{subject to} \quad f(X) = \frac{1}{2}\|\mathcal{A}(XX^T - ZZ^T)\|^2 \\ & \qquad\qquad\quad \nabla f(X) = 0, \quad \nabla^2 f(X) \succeq 0 \\ & \qquad\qquad\quad (1-\delta)I \preceq \mathcal{A}^T\mathcal{A} \preceq (1+\delta)I \end{aligned} \tag{21}$$

Surprisingly, the upper-bound is *tight*—problem (21) is actually an *exact reformulation* of problem (15).

Theorem 8 (Exact convex reformulation)**.** *Given $X, Z \in \mathbb{R}^{n\times r}$, we have $\delta(X,Z) = \mathrm{LMI}(X,Z)$ with both problems attaining their minima. Moreover, every minimizer $\mathcal{H}^\star$ for the latter problem is related to a minimizer $\mathcal{A}^\star$ for the former problem via $\mathcal{H}^\star = (\mathcal{A}^\star)^T\mathcal{A}^\star$.*

Theorem 8 is the key insight that allows us to establish our main results. When rank $r = 1$, the LMI is sufficiently simple that it can be suitably relaxed and solved in closed-form, as we will soon show in Section 7. But even when $r > 1$, the LMI can still be solved numerically using an interior-point method. This allows us to perform numerical experiments to probe at the true value of $\delta^\star$ and $\delta^\star(\epsilon)$, even when analytical arguments are not available.

Section 5.1 below gives a proof of Theorem 8. A key step of the proof is to establish the following equivalence:

$$\mathrm{LMI}(X,Z) = \mathrm{LMI}(P^TX, P^TZ) \text{ where } P = \mathrm{orth}([X,Z]). \tag{22}$$

For small values of the rank $r \ll n$, equation (22) also yields an efficient algorithm for evaluating $\mathrm{LMI}(X,Z)$ in *linear* time: compute P, P^TX, and P^TZ, and then evaluate $\mathrm{LMI}(P^TX, P^TZ)$. Moreover, the associated minimizer $\mathcal{A}^\star$ can also be efficiently recovered. These practical aspects are discussed in detail in Section 5.2.

5.1 Proof of Theorem 8

Given $X, Z \in \mathbb{R}^{n\times r}$, we define $e \in \mathbb{R}^{n^2}$ and $\mathbf{X} \in \mathbb{R}^{n^2\times nr}$ to satisfy equation (11) with respect to X and $M^\star = ZZ^T$. Then, problem (21) can be explicitly written as

$$\begin{aligned} \mathrm{LMI}(X,Z) \quad = \quad \underset{\delta,\mathbf{H}}{\text{minimum}} \quad & \delta \\ \text{subject to} \quad & \mathbf{X}^T\mathbf{He} = 0, \\ & 2\cdot[I_r \otimes \mathrm{mat}(\mathbf{He})] + \mathbf{X}^T\mathbf{HX} \succeq 0, \\ & (1-\delta)I \preceq \mathbf{H} \preceq (1+\delta)I, \end{aligned} \tag{23}$$

with Lagrangian dual

$$\begin{aligned} \underset{y,U_1,U_2,V}{\text{maximize}} \quad & \mathrm{tr}(U_1 - U_2) \\ \text{subject to} \quad & \mathrm{tr}(U_1+U_2) = 1, \\ & \sum_{j=1}^{r} (\mathbf{X}y - \mathrm{vec}\,(V_{j,j}))\mathbf{e}^T + \mathbf{e}(\mathbf{X}y - \mathrm{vec}\,(V_{j,j}))^T \\ & \qquad - \mathbf{X}V\mathbf{X}^T = U_1 - U_2, \\ & V = \begin{bmatrix} V_{1,1} & \cdots & V_{r,1} \\ \vdots & \ddots & \vdots \\ V_{r,1}^T & \cdots & V_{r,r} \end{bmatrix} \succeq 0, \quad U_1 \succeq 0, \quad U_2 \succeq 0. \end{aligned} \tag{24}$$

The dual problem admits a strictly feasible point (for sufficiently small $\epsilon > 0$, set $y = 0$, $V = \epsilon I$, $U_1 = \eta I - \epsilon W$, and $U_2 = \eta \cdot I + \epsilon W$ where $2\eta = n^{-2}$ and $2W = r[\mathrm{vec}\,(I)\mathbf{e}^T + \mathbf{e}\mathrm{vec}\,(I)^T] - \mathbf{X}\mathbf{X}^T$) and the primal problem is bounded (the constraints imply $\delta \geq 0$). Hence, Slater's condition is satisfied, strong duality holds, and the primal problem attains its optimal value at a minimizer.

It turns out that both the minimizer and the minimum are invariant under an orthogonal projection.

Lemma 9 (Orthogonal projection)**.** *Given $X, Z \in \mathbb{R}^{n\times r}$, let $P \in \mathbb{R}^{n\times q}$ with $q \leq n$ satisfy*

$$P^TP = I_q, \qquad PP^TX = X, \qquad PP^TZ = Z.$$

Let $(\hat{\delta}, \hat{\mathbf{H}})$ be a minimizer for $\mathrm{LMI}(P^TX, P^TZ)$*. Then, $(\delta, \mathbf{H})$ is a minimizer for* $\mathrm{LMI}(X, Z)$*, where $\mathbf{P} = P \otimes P$ and*

$$\delta = \hat{\delta}, \qquad \mathbf{H} = \mathbf{P}\hat{\mathbf{H}}\mathbf{P}^T + (I - \mathbf{P}\mathbf{P}^T).$$

Proof. Choose arbitrarily small $\epsilon > 0$. Strong duality guarantees the existence of a dual feasible point $(\hat{y}, \hat{U}_1, \hat{U}_2, \hat{V})$ with duality gap ϵ. This is a certificate that proves $(\hat{\delta}, \hat{\mathbf{H}})$ to be ϵ-suboptimal for $\mathrm{LMI}(P^TX, P^TZ)$. We can mechanically verify that $(\delta, \mathbf{H})$ is primal feasible and that (y, U_1, U_2, V) is dual feasible, where

$$y = (I_r \otimes P)\hat{y}, \qquad U_1 = \mathbf{P}\hat{U}_1\mathbf{P}^T, \qquad U_2 = \mathbf{P}\hat{U}_2\mathbf{P}^T, \qquad V = (I_r \otimes P)\hat{V}(I_r \otimes P)^T.$$

Then, (y, U_1, U_2, V) is a certificate that proves $(\delta, \mathbf{H})$ to be ϵ-suboptimal for $\mathrm{LMI}(X, Z)$, since

$$\delta - \mathrm{tr}(U_1 - U_2) = \hat{\delta} - \mathrm{tr}(\hat{U}_1 - \hat{U}_2) = \epsilon.$$

Given that ϵ-suboptimal certificates exist for all $\epsilon > 0$, the point $(\delta, \mathbf{H})$ must actually be optimal. The details for verifying primal and dual feasibility are straightforward but tedious; they are included in Appendix A for completeness. $\square$

Recall that we developed an upper-bound $\mathrm{LMI}(X, Z)$ on $\delta(X, Z)$ by replacing δ-RIP with a convex *sufficient* condition (20). The same idea can also be used to produce a lower-bound. Specifically, we replace the δ-RIP constraint with a convex *necessary* condition, obtained by enforcing the RIP inequality (3) over *a subset of* rank-$2r$ matrices (instead of over all rank-$2r$ matrices):

$$(1-\delta)\|PYP^T\|_F^2 \le \|\mathcal{A}(PYP^T)\|^2 \le (1+\delta)\|PYP^T\|_F^2 \qquad \forall Y \in \mathbb{R}^{d\times d} \tag{25}$$

where P is a *fixed* $n \times d$ matrix with $d \le 2r$. The resulting problem is also convex (we write $\mathbf{P} = P \otimes P$)

$$\begin{aligned} \delta(X,Z) \quad \ge \quad & \text{minimize} && \delta \\ & \text{subject to} && \mathbf{X}^T\mathbf{He} = 0, \\ &&& 2\cdot[I_r \otimes \mathrm{mat}(\mathbf{He})] + \mathbf{X}^T\mathbf{HX} \succeq 0, \\ &&& (1-\delta)\mathbf{P}^T\mathbf{P} \preceq \mathbf{P}^T\mathbf{HP} \preceq (1+\delta)\mathbf{P}^T\mathbf{P} \end{aligned} \tag{26}$$

with Lagrangian dual

$$\begin{aligned} \underset{y,U_1,U_2,V}{\text{maximize}} \quad & \mathrm{tr}[\mathbf{P}(U_1 - U_2)\mathbf{P}^T] \\ \text{subject to} \quad & \mathrm{tr}[\mathbf{P}(U_1 + U_2)\mathbf{P}^T] = 1, \\ & \sum_{j=1}^{r} (\mathbf{X}y - \mathrm{vec}\,(V_{j,j}))\mathbf{e}^T + \mathbf{e}(\mathbf{X}y - \mathrm{vec}\,(V_{j,j}))^T \\ & \qquad - \mathbf{X}V\mathbf{X}^T = \mathbf{P}(U_1 - U_2)\mathbf{P}^T, \\ & V = \begin{bmatrix} V_{1,1} & \cdots & V_{r,1} \\ \vdots & \ddots & \vdots \\ V_{r,1}^T & \cdots & V_{r,r} \end{bmatrix} \succeq 0, \quad U_1 \succeq 0, \quad U_2 \succeq 0. \end{aligned} \tag{27}$$

It turns out that for the specific choice of $P = \mathrm{orth}([X, Z])$, the lower-bound in (26) *coincides* with the upper-bound in (23).

Lemma 10 (Tightness)**.** *Define* $P = \mathrm{orth}([X, Z])$*. Let* $(\hat{\delta}, \hat{\mathbf{H}})$ *be a minimizer for* $\mathrm{LMI}(P^TX, P^TZ)$*. Then,* $(\delta, \mathbf{H})$ *is a minimizer for problem (26), where* $\mathbf{P} = P \otimes P$ *and*

$$\delta = \hat{\delta}, \qquad\qquad \mathbf{H} = \mathbf{P}\hat{\mathbf{H}}\mathbf{P}^T.$$

Proof. The proof is almost identical to that of Lemma 9. Again, choose arbitrarily small $\epsilon > 0$. Let $(\hat{y}, \hat{U}_1, \hat{U}_2, \hat{V})$ be a dual feasible point for $\mathrm{LMI}(P^T X, P^T Z)$ with duality gap ϵ. Then, (y, U_1, U_2, V) where

$$y = (I_r \otimes P)\hat{y}, \qquad U_1 = \hat{U}_1, \qquad U_2 = \hat{U}_2, \qquad V = (I_r \otimes P)\hat{V}(I_r \otimes P)^T$$

is a certificate that proves $(\delta, \mathbf{H})$ to be ϵ-suboptimal for problem (26). The details for verifying primal and dual feasibility are included in Appendix B. □

Putting the upper- and lower-bounds together then yields a short proof of Theorem 8.

Proof of Theorem 8. Denote $\delta_{\mathrm{ub}} = \mathrm{LMI}(X, Z)$ as the optimal value to the upper-bound problem (23) and $\mathcal{H}^\star$ as the corresponding minimizer. (The minimizer $\mathcal{H}^\star$ always exists due to the boundedness of the primal problem and the existence of a strictly feasible point in the dual problem.) Denote δ_{lb} as the optimal value to the lower-bound problem (26). For $P = \mathrm{orth}([X, Z])$, the sequence of inclusions

$$\{PYP^T : Y \in \mathbb{R}^{d\times d}\} \subseteq \{M \in \mathbb{R}^{n\times n} : \mathrm{rank}\,(M) \le 2r\} \subseteq \mathbb{R}^{n\times n},$$

implies $\delta_{\mathrm{lb}} \le \delta(X, Z) \le \delta_{\mathrm{ub}}$. However, by Lemma 9 and Lemma 10, we actually have $\delta_{\mathrm{ub}} = \delta_{\mathrm{lb}} = \mathrm{LMI}(P^T X, P^T Z)$, and hence $\delta_{\mathrm{lb}} = \delta(X, Z) = \delta_{\mathrm{ub}}$. Finally, the minimizer $\mathcal{H}^\star$ factors into $(\mathcal{A}^\star)^T \mathcal{A}^\star$, where $\mathcal{A}^\star$ satisfies the sufficient condition (20), and hence also δ-RIP. □

5.2 Efficient evaluation of $\delta(X, Z)$ and $\mathcal{A}^\star$

We now turn to the practical problem of evaluating $\delta(X, Z)$ and the associated minimizer $\mathcal{A}^\star$ using a numerical algorithm. While its exact reformulation $\mathrm{LMI}(X, Z) = \delta(X, Z)$ is indeed convex, naïvely solving it using an interior-point solution can require up to $O(n^{13})$ time and $O(n^8)$ memory (as it requires solving an order-n^2 semidefinite program). In our experiments, the largest instances of (21) that we could accommodate using the state-of-the-art solver MOSEK [Andersen and Andersen, 2000] had dimensions no greater than $n \le 12$.

Instead, we can efficiently evaluate $\delta(X, Z)$ using Algorithm 1. When the rank $r \ll n$ is small, the algorithm evaluates $\delta(X, Z)$ in linear $O(n)$ time and memory, and if desired, also recovers the minimizer $\mathcal{A}^\star$ in $O(n^4)$ time and memory. In practice, our numerical experiments were able to accommodate for rank as large as $r \le 10$.

Proposition 11. *Algorithm 1 correctly outputs the minimum value $\hat{\delta} = \delta(X, Z)$ and the minimizer $\mathcal{A}^\star$. Moreover, Steps 1-2 for $\hat{\delta}$ use*

$$O(nr^2 + r^{13}\log(1/\epsilon)) \text{ time and } O(nr + r^8) \text{ memory,} \tag{28}$$

while Steps 3-6 for $\mathcal{A}^\star$ use

$$O(n^4 + n^2 r^3 + nr^4 + r^6) \text{ time and } O(n^4) \text{ memory.} \tag{29}$$

Algorithm 1 Efficient algorithm for $\delta(X,Z)$ and $\mathcal{A}^\star$.

Input. Choices of $X,Z\in\mathbb{R}^{n\times r}$.
Output. The value $\hat{\delta}=\delta(X,Z)$ and the corresponding minimizer $\mathcal{A}^\star$ (if desired).
Algorithm.

1. Compute $P=\mathrm{orth}([X,Z])\in\mathbb{R}^{n\times d}$ and project $\hat{X}=P^TX$ and $\hat{Z}=P^TZ$.
2. Solve $\hat{\delta}=\mathrm{LMI}(\hat{X},\hat{Z})$ using an interior-point method to obtain minimizer $\hat{\mathbf{H}}$. **Output** $\hat{\delta}$.
3. Compute the orthogonal complement $P_\perp=\mathrm{orth}(I-PP^T)\in\mathbb{R}^{n\times(n-d)}$.
4. Factor $\hat{\mathbf{H}}=\hat{\mathbf{A}}^T\hat{\mathbf{A}}$ using (dense) Cholesky factorization.
5. Analytically factor $(\mathbf{A}^\star)^T\mathbf{A}^\star=\mathbf{H}^\star=\mathbf{P}\hat{\mathbf{H}}\mathbf{P}^T+(I-\mathbf{P}\mathbf{P}^T)$ using the formula
$$(\mathbf{A}^\star)^T=\begin{bmatrix}(P\otimes P)\hat{\mathbf{A}}^T & P\otimes P_\perp & P_\perp\otimes P & P_\perp\otimes P_\perp\end{bmatrix}$$
while using the Kronecker identity $(P\otimes P)\mathrm{vec}\,(U)=\mathrm{vec}\,(PUP^T)$ to evaluate each column of $(P\otimes P)\hat{\mathbf{A}}^T$.
6. Recover the matrices $A_1^\star,\dots,A_m^\star$ associated with the minimizer $\mathcal{A}^\star$ by matricizing each row of $\mathbf{A}^\star$. **Output** $\mathcal{A}^\star$.

Proof. We begin by verifying correctness. The fact that the minimum value $\delta(X,Z)=\mathrm{LMI}(P^TX,P^TZ)$ follows from Lemma 8 and Lemma 9. To prove correctness for the minimizer $\mathcal{A}^\star$, we recall that Algorithm 1 defines $P_\perp\in\mathbb{R}^{n\times(n-d)}$ as the orthogonal complement of $P\in\mathbb{R}^{n\times d}$, and note that

$$\begin{aligned}&PP^T\otimes P_\perp P_\perp^T+P_\perp P_\perp^T\otimes PP^T+P_\perp P_\perp^T\otimes P_\perp P_\perp^T\\=&(PP^T+P_\perp P_\perp^T)\otimes(PP^T+P_\perp P_\perp^T)-PP^T\otimes PP^T\\=&I-\mathbf{P}\mathbf{P}^T,\end{aligned}$$

where $\mathbf{P}=P\otimes P$. Hence, Algorithm 1 produces the minimizer $\mathbf{H}^\star=\mathbf{P}\hat{\mathbf{H}}\mathbf{P}^T+(I-\mathbf{P}\mathbf{P}^T)$ for $\mathrm{LMI}(X,Z)$ in Lemma 9 as desired.

Now, let us quantify complexity. Note that $d\le 2r$ by construction. Step 1 takes $\Theta(nr^2)$ time and $\Theta(nr)$ memory. Step 2 requires solving an order $\theta=O(d^4)$ semidefinite program, and so requires $O(\theta^{6.5}\log(1/\epsilon))$ time and $O(\theta^4)$ memory. Stopping here yields (28). Step 3 uses $\Theta(n^3+n^2r)$ time and $\Theta(n^2)$ memory. Step 4 uses $\Theta(r^6)$ time and $\Theta(r^4)$ memory. Step 5 performs $\Theta(r^2)$ matrix-vector products each costing $O(nr^2+n^2r)$ time and $O(n^4)$ memory, and then filling the rest of $\mathbf{A}$ in $O(n^4)$ time and memory. Step 6 costs $O(n^4)$ time and memory. Summing the terms and substituting $r=O(n)$ in the memory complexity yields the desired figures. □

6 Counterexample with $\delta = 1/2$ for the rank-1 problem

In this section, we use a *family* of counterexamples to prove that δ-RIP with $\delta < 1/2$ is necessary for the exact recovery of *any* arbitrary rank-1 ground truth $M^\star = zz^T$ (and not just the 2×2 ground truth studied by Zhang et al. [2018a]). Specifically, we state a choice of $\mathcal{A}^\star$ that satisfies 1/2-RIP but whose $f^\star(x) = \|\mathcal{A}^\star(xx^T - M^\star)\|^2$ admits a spurious second-order point.

Example 12. Given rank-1 ground truth $M^\star = zz^T \neq 0$, define a set of orthonormal vectors $u_1, u_2, \ldots, u_n \in \mathbb{R}^n$ with $u_1 = z/\|z\|$, and define $m = n^2$ measurement matrices $A_1, A_2, \ldots, A_m$, with

$$A_1 = u_1u_1^T + \frac{1}{2}u_2u_2^T, \qquad A_2 = \frac{\sqrt{3}}{2}(u_1u_2^T + u_2u_1^T),$$
$$A_{n+1} = \frac{1}{\sqrt{2}}(u_1u_2^T - u_2u_1^T), \qquad A_{n+2} = \frac{\sqrt{3}}{2}u_2u_2^T,$$

and the remaining $n^2 - 4$ measurement matrices sequentially assigned as

$$A_k = u_iu_j^T, \quad k = i + n \cdot (j-1), \qquad \forall (i,j) \in \{1, 2, \ldots, n\}^2 \backslash \{1, 2\}^2.$$

Then, the associated operator $\mathcal{A}^\star$ satisfies 1/2-RIP:

$$\left(1 - \frac{1}{2}\right)\|M\|_F^2 \leq \|\mathcal{A}^\star(M)\|^2 \leq \left(1 + \frac{1}{2}\right)\|M\|_F^2 \qquad \forall M \in \mathbb{R}^{n\times n},$$

but the corresponding $f^\star(x) \equiv \|\mathcal{A}^\star(xx^T - M^\star)\|^2$ admits $x = (\|z\|/\sqrt{2})\, u_2$ as a spurious second-order critical point:

$$f^\star(x) = \frac{3}{4}\|M^\star\|_F^2, \qquad \nabla f^\star(x) = 0, \qquad \nabla^2 f^\star(x) \succeq 8xx^T.$$

We derived Example 12 by numerically solving $\delta(x, z)$ with any x satisfying $x^Tz = 0$ and $\|x\| = \|z\|/\sqrt{2}$ using Algorithm 1. The 1/2-RIP counterexample of Zhang et al. [2018a] arises as the instance of Example 12 associated with the 2×2 ground truth $\hat{z}\hat{z}^T$ and $\hat{z} = (1, 0)$:

$$\hat{A}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}, \quad \hat{A}_2 = \begin{bmatrix} 0 & \sqrt{3}/2 \\ \sqrt{3}/2 & 0 \end{bmatrix}, \quad \hat{A}_3 = \begin{bmatrix} 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 \end{bmatrix}, \quad \hat{A}_4 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{3}/2 \end{bmatrix}.$$

The associated operator $\hat{\mathcal{A}} : \mathbb{S}^2 \to \mathbb{R}^4$ is invertible and satisfies 1/2-RIP, but $\hat{x} = (0, 1/\sqrt{2})$ is a spurious second-order point:

$$\hat{f}(\hat{x}) \equiv \|\hat{\mathcal{A}}(\hat{x}\hat{x}^T - \hat{z}\hat{z}^T)\|^2 = \frac{3}{4}, \qquad \nabla \hat{f}(\hat{x}) = 0, \qquad \nabla^2 \hat{f}(\hat{x}) = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}.$$

We can verify the correctness of Example 12 for a general rank-1 ground truth by reducing it down to this specific 2×2 example.

Proof of correctness for Example 12. We can mechanically verify Example 12 to be correct with ground truth $\hat{z}\hat{z}^T$ and $\hat{z} = (1,0)$. Denote $\hat{\mathcal{A}}$, $\hat{f}(\hat{x}) = \|\hat{\mathcal{A}}(\hat{x}\hat{x}^T - \hat{z}\hat{z}^T)\|^2$, and $\hat{x} = (0, 1/\sqrt{2})$ as the corresponding minimizer, nonconvex objective, and spurious second-order critical point.

For a general rank-1 ground truth $M^\star = zz^T$, recall that we have defined a set of orthonormal vectors $u_1, u_2, \ldots, u_n \in \mathbb{R}^n$ with $u_1 = z/\|z\|$. Then, setting $P = [u_1, u_2]$ and $P_\perp = [u_3, \ldots, u_n]$ shows that the matrix version of $\mathcal{A}^\star$ can be permuted row-wise to satisfy

$$(\mathbf{A}^\star)^T = \begin{bmatrix}(P \otimes P)\hat{\mathbf{A}}^T & P \otimes P_\perp & P_\perp \otimes P & P_\perp \otimes P_\perp\end{bmatrix}$$

where $\hat{\mathbf{A}}$ is the matrix version of $\hat{\mathcal{A}}$. Repeating the proof of Proposition 11 shows that

$$(\mathbf{A}^\star)^T \mathbf{A}^\star = \mathbf{P}\hat{\mathbf{A}}^T\hat{\mathbf{A}}\mathbf{P} + (I - \mathbf{P}\mathbf{P}^T)$$

where $\mathbf{P} = P \otimes P$, and so $\mathcal{A}^\star$ also satisfies 1/2-RIP. Moreover, this implies that

$$f^\star(x) \equiv \|\mathcal{A}^\star(xx^T - zz^T)\|^2 = \|z\|^4 \hat{f}(P^T x/\|z\|) + (\|x\|^4 - \|P^T x\|^4).$$

Differentiating yields the following at $x = (\|z\|/\sqrt{2})u_2$:

$$\begin{aligned}
f^\star(x) &= \|z\|^4 \hat{f}(\hat{x}) && = (3/4)\|z\|^4, \\
\nabla f^\star(x) &= \|z\|^3 P \nabla \hat{f}(\hat{x}) && = 0, \\
\nabla^2 f^\star(x) &= \|z\|^2 P \nabla^2 \hat{f}(\hat{x}) P^T + 2\|x\|^2 (I - PP^T) && \succeq 4\|z\|^2 u_2 u_2^T.
\end{aligned}$$

□

7 Closed-form lower-bound for the rank-1 problem

It turns out that the LMI problem (21) in the rank-1 case is sufficiently simple to be suitably relaxed and then solved in closed-form. Our main result in this section is the following lower-bound on $\delta(x,z) = \mathrm{LMI}(x,z)$.

Theorem 13 (Closed-form lower-bound)**.** *Let $x, z \in \mathbb{R}^n$ be arbitrary nonzero vectors, and define their length ratio ρ and incidence angle ϕ:*

$$\rho \equiv \frac{\|x\|}{\|z\|}, \qquad \phi \equiv \arccos\left(\frac{x^T z}{\|x\|\|z\|}\right). \tag{30}$$

Define the following two scalars with respect to ρ and ϕ:

$$\alpha = \frac{\sin^2\phi}{\sqrt{(\rho^2-1)^2 + 2\rho^2\sin^2\phi}}, \qquad \beta = \frac{\rho^2}{\sqrt{(\rho^2-1)^2 + 2\rho^2\sin^2\phi}}.$$

Then, we have $\delta(x,z) \geq \delta_{\mathrm{lb}}(x,z)$, where

$$\delta_{\mathrm{lb}}(x,z) \equiv \sqrt{1-\alpha^2} \qquad \text{if } \beta \geq \frac{\alpha}{1+\sqrt{1-\alpha^2}}, \tag{31}$$

$$\frac{1 - 2\alpha\beta + \beta^2}{1-\beta^2} \qquad \text{if } \beta \leq \frac{\alpha}{1+\sqrt{1-\alpha^2}}. \tag{32}$$

The rank-1 global and local recovery guarantees follow quickly from this theorem, as shown below.

Proof of Theorem 3. The existence of Example 12 already proves that

$$\delta^\star = \min_{x,z\in\mathbb{R}^n} \delta(x,z) \le 1/2. \tag{33}$$

Below, we will show that $\delta_{\text{lb}}(x,z)$ attains its minimum of $1/2$ at any x satisfying $x^T z = 0$ and $\|x\|/\|z\| = 1/\sqrt{2}$, as in

$$1/2 = \min_{x,z\in\mathbb{R}^n} \delta_{\text{lb}}(x,z) \le \delta^\star. \tag{34}$$

Substituting $\delta^\star = 1/2$ into Lemma 6 then completes the proof of our global recovery guarantee in Theorem 3.

To prove (34), we begin by optimizing $\delta_{\text{lb}}(x,z)$ over the region $\beta \ge \alpha/(1+\sqrt{1-\alpha^2})$ using equation (31), and find that the minimum value is attained along the boundary

$$\beta = \frac{\alpha}{1+\sqrt{1-\alpha^2}} = \frac{1-\sqrt{1-\alpha^2}}{\alpha}.$$

Note that the two equations (31) and (32) coincide at this boundary:

$$\left(\frac{1-2\alpha\beta+\beta^2}{1-\beta^2}\right)\left(\frac{\alpha/\beta}{\alpha/\beta}\right) = \frac{(1+\sqrt{1-\alpha^2})-2\alpha^2+(1-\sqrt{1-\alpha^2})}{(1+\sqrt{1-\alpha^2})-(1-\sqrt{1-\alpha^2})} = \sqrt{1-\alpha^2}.$$

Now, we optimize $\delta_{\text{lb}}(x,z)$ over the region $\beta \le \alpha/(1+\sqrt{1-\alpha^2})$ using equation (32). First, substituting the definitions of α and β yields

$$\delta_{\text{lb}}(x,z) = \frac{1-2\alpha\beta+\beta^2}{1-\beta^2} = \frac{(\rho^4+1-2\rho^2\cos^2\phi)-2\rho^2\sin^2\phi+\rho^4}{(\rho^4+1-2\rho^2\cos^2\phi)-\rho^4} = \frac{(\rho^2-1)^2+\rho^4}{1-2\rho^2\cos^2\phi}.$$

This expression is minimized at $\phi = \pm\pi/2$ and $\rho = 1/\sqrt{2}$, with a minimum value of $1/2$. The corresponding point $\alpha = 2/\sqrt{5}$ and $\beta = \alpha/4$ lies in the strict interior $\beta < \alpha/(1+\sqrt{1-\alpha^2})$. This point must be the global minimum, because it dominates the boundary $\beta = \alpha/(1+\sqrt{1-\alpha^2})$, which in turn dominates the other region $\beta > \alpha/(1+\sqrt{1-\alpha^2})$. ☐

Proof of Theorem 4. We will optimize over an ϵ-neighborhood of the ground truth and show that

$$\left(1-\frac{\epsilon^2}{2(1-\epsilon)}\right)^{1/2} \le \min_{x,z\in\mathbb{R}^n}\{\delta_{\text{lb}}(x,z) : \|xx^T - zz^T\|_F \le \epsilon\|zz^T\|_F\} \le \delta^\star(\epsilon). \tag{35}$$

Substituting this lower-bound on $\delta^\star(\epsilon)$ into Lemma 7 then completes the proof of our local recovery guarantee in Theorem 4.

To obtain (35), we first note that the ϵ-neighborhood constraint implies the following

$$\|xx^T - zz^T\|_F \le \epsilon\|zz^T\|_F \qquad \iff \qquad (\rho^2-1)^2 + 2\rho^2\sin^2\phi \le \epsilon^2.$$

This in turn implies $\epsilon^2 \ge (\rho^2-1)^2$ and $\epsilon^2 \ge [(\rho^2-1)^2+2\rho^2]\sin^2\phi$, and hence

$$1-\epsilon \le \rho^2 \le 1+\epsilon, \qquad \sin^2\phi \le \epsilon^2.$$

We wish to derive a threshold $\hat{\epsilon}$ such that if $\epsilon \le \hat{\epsilon}$, then

$$\frac{\beta}{\alpha} = \frac{\rho^2}{\sin^2\phi} \ge \frac{1-\epsilon}{\epsilon^2} \ge 1 \ge \frac{1}{1+\sqrt{1-\alpha^2}},$$

and so $\delta_{\text{lb}}(x,z) = \sqrt{1-\alpha^2}$ as dictated entirely by equation (31). Clearly, this requires solving the quadratic equation $(1-\hat{\epsilon}) = \hat{\epsilon}^2$ for the positive root at $\hat{\epsilon} = (-1+\sqrt{5})/2 \ge 0.618$. Now, we upper-bound α^2 to lower-bound $\sqrt{1-\alpha^2}$:

$$\begin{aligned}\alpha^2 &= \frac{\sin^4\phi}{(\rho^2-1)^2 + 2\rho^2\sin^2\phi} \le \frac{\sin^4\phi}{[(\rho^2-1)^2+2\rho^2]\sin^2\phi} = \frac{\sin^2\phi}{\rho^4+1} \\ &\le \frac{\epsilon^2}{(1-\epsilon)^2+1} = \frac{\epsilon^2}{2-2\epsilon+\epsilon^2} \le \frac{\epsilon^2}{2(1-\epsilon)}\end{aligned}$$

and so

$$\delta(x,z) \ge \delta_{\text{lb}}(x,z) = \sqrt{1-\alpha^2} \ge \sqrt{1 - \frac{\epsilon^2}{2(1-\epsilon)}}.$$

□

Proof of Corollary 5. Under δ-RIP, a point with a small residual must also have a small error:

$$(1-\delta)\|xx^T - M^\star\|_F^2 \le f(x) \le (1-\delta)\epsilon^2\|M^\star\|_F^2. \tag{36}$$

In particular, any point in the level set $f(x) \le f(x_0)$ must also lie in the ϵ-neighborhood:

$$f(x) \le f(x_0) < (1-\delta)\epsilon^2 f(0) \qquad \Longrightarrow \qquad \|xx^T - M^\star\|_F \le \epsilon\|M^\star\|_F.$$

Additionally, note that

$$\epsilon^2 \le 1-\delta^2, \quad \epsilon^2 \le \frac{\sqrt{5}-1}{2} \qquad \Longrightarrow \qquad \delta \le \sqrt{1-\frac{\epsilon^2}{2(1-\epsilon)}}$$

because $2(1-\epsilon) \le 1$. The result then follows by applying Theorem 4. □

The rest of this section is devoted to proving Theorem 13. We begin by providing a few important lemmas in Section 7.1, and then move to the proof itself in Section 7.2.

7.1 Technical lemmas

Given $M \in \mathbb{S}^n$ with eigendecomposition $M = \sum_{i=1}^m \lambda_i v_i v_i^T$, we define its projection onto the semidefinite cone as the following

$$[M]_+ \equiv \arg\min_{S \succeq 0} \|M - S\|_F^2 = \sum_{i=1}^n \max\{\lambda_i, 0\} v_i v_i^T.$$

For notational convenience, we also define a complement projection

$$[M]_- \equiv [-M]_+ = [M]_+ - M,$$

thereby allowing us to decompose every M into a positive and a negative part as in

$$M = [M]_+ - [M]_- \quad \text{where} \quad [M]_+ \succeq 0, \quad [M]_- \succeq 0.$$

Lemma 14. *Given $M \in \mathbb{S}^n$ with $\mathrm{tr}(M) \geq 0$, the following problem*

$$\underset{\alpha,U,V}{\text{minimize}} \quad \mathrm{tr}(V) \quad \text{subject to} \quad \mathrm{tr}(U) = 1, \quad \alpha M = U - V, \quad U, V \succeq 0$$

has minimizer

$$\alpha^\star = 1/\mathrm{tr}([M]_+), \qquad U^\star = \alpha^\star \cdot [M]_+, \qquad V^\star = \alpha^\star \cdot [M]_-.$$

Proof. Write $p^\star$ as the optimal value. Then,

$$\begin{aligned}
p^\star =& \max_{\beta} \min_{\substack{\alpha\in\mathbb{R} \\ U,V\succeq 0}} \{\mathrm{tr}(V) + \beta \cdot [\mathrm{tr}(U) - 1] : \alpha M = U - V\} \\
=& \max_{\beta\geq 0} \min_{\alpha\in\mathbb{R}} \{-\beta + \min_{U,V\succeq 0} \{\mathrm{tr}(V) + \beta \cdot \mathrm{tr}(U) : \alpha M = U - V\}\} \\
=& \max_{\beta\geq 0} \min_{\alpha\in\mathbb{R}} \{-\beta + \alpha \cdot [\mathrm{tr}([M]_-) + \beta \cdot \mathrm{tr}([M]_+)]\} \\
=& \max_{\beta\geq 0} \{-\beta : \mathrm{tr}([M]_-) + \beta \cdot \mathrm{tr}([M]_+) = 0\} \\
=& \mathrm{tr}([M]_-)/\mathrm{tr}([M]_+) = \mathrm{tr}(V^\star).
\end{aligned}$$

The first line converts an equality constraint into a Lagrangian. The second line isolates the optimization over $U, V \succeq 0$ with $\beta \geq 0$, noting that $\beta < 0$ would yield $\mathrm{tr}(U) \to \infty$. The third line solves the minimization over $U, V \succeq 0$ in closed-form. The fourth line views α as a Lagrange multiplier. □

For symmetric indefinite matrices of a particular rank-2 form, the positive and negative eigenvalues can be computed in closed-form.

Lemma 15. *Given $a, b \in \mathbb{R}^n$, the matrix $M = ab^T + ba^T$ has eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ where:*

$$\lambda_i = \begin{cases} +\|a\|\|b\|(1 + \cos\theta) & i = 1 \\ -\|a\|\|b\|(1 - \cos\theta) & i = n \\ 0 & \text{otherwise,} \end{cases}$$

and $\theta \equiv \arccos\left(\frac{a^T b}{\|a\|\|b\|}\right)$ is the angle between a and b.

Proof. Without loss of generality, assume that $\|a\| = \|b\| = 1$. (Otherwise, rescale $\hat{a} = a/\|a\|$, $\hat{b} = b/\|b\|$ and write $M = \|a\|\|b\| \cdot (\hat{a}\hat{b}^T + \hat{b}\hat{a}^T)$.) Decompose b into a tangent and normal component with respect to a, as in

$$b = a \underbrace{a^T b}_{\cos\theta} + \underbrace{(I - aa^T)b}_{c\sin\theta} = a\cos\theta + c\sin\theta,$$

where c is a unit normal vector with $\|c\| = 1$ and $a^T c = 0$. This allows us to write

$$ab^T + ba^T = \begin{bmatrix} a & c \end{bmatrix} \begin{bmatrix} 2\cos\theta & \sin\theta \\ \sin\theta & 0 \end{bmatrix} \begin{bmatrix} a & c \end{bmatrix}^T$$

and hence M is spectrally similar a 2×2 matrix with eigenvalues $\cos\theta \pm \sqrt{\cos^2\theta + \sin^2\theta}$. □

Given $x, z \in \mathbb{R}$, recall that $\mathbf{e}$ and $\mathbf{X}$ are implicitly defined in (11) to satisfy

$$\mathbf{e} = \mathrm{vec}\,(xx^T - zz^T), \qquad \mathbf{X}y = \mathrm{vec}\,(xy^T + yx^T) \quad \forall y \in \mathbb{R}^n.$$

Let us give a preferred orthogonal basis to study these two objects. We define $v_1 = x/\|x\|$ in the direction of x. Then, we decompose z into a tangent and normal component with respect to v_1, as in

$$z = v_1 \underbrace{v_1^T z}_{\|z\|\cos\phi} + \underbrace{(I - v_1 v_1^T) z}_{v_2 \|z\| \sin\phi} = \|z\| \cdot (v_1 \cos\phi + v_2 \sin\phi). \tag{37}$$

Here, ϕ is the incidence angle between x and z, and v_2 is the associated unit normal vector with $\|v_2\| = 1$ and $v_1^T v_2 = 0$. Using the Gram-Schmidt process, we complete v_1, v_2 with the remaining $n-2$ set of orthonormal unit vectors $v_3, v_4, \ldots, v_n$. This results in a set of right singular vectors for $\mathbf{X}$.

Lemma 16. *The matrix* $\mathbf{X} \in \mathbb{R}^{n^2 \times n}$ *has singular value decomposition* $\mathbf{X} = \sum_{i=1}^n \sigma_i u_i v_i^T$ *where* v_i *are defined as above, and*

$$\sigma_i = \begin{cases} 2\|x\| & i = 1 \\ \sqrt{2}\|x\| & i > 1 \end{cases}, \qquad u_i = \begin{cases} v_1 \otimes v_1 & i = 1 \\ \frac{1}{\sqrt{2}}(v_i \otimes v_1 + v_1 \otimes v_i) & i > 1 \end{cases}.$$

Proof. It is easy to verify that

$$\begin{aligned} \mathbf{X}y = y \otimes x + x \otimes y &= \|x\|(y \otimes v_1 + v_1 \otimes y) \\ &= \|x\| \cdot \sum_{i=1}^n (v_i^T y)(v_i \otimes v_1 + v_1 \otimes v_i). \end{aligned}$$

Normalizing the left singular vectors then yields the designed u_i and σ_i. $\square$

We can also decompose $\mathbf{e}$ into a tangent and normal component with respect to range($\mathbf{X}$) as in

$$\mathbf{e} = \underbrace{\mathbf{X}\mathbf{X}^\dagger \mathbf{e}}_{\hat{e}_1 \|\mathbf{e}\| \cos\theta} + \underbrace{(I - \mathbf{X}\mathbf{X}^\dagger)\mathbf{e}}_{\hat{e}_2 \|\mathbf{e}\| \sin\theta} = \|\mathbf{e}\| \cdot (\hat{e}_1 \cos\theta + \hat{e}_2 \sin\theta) \tag{38}$$

where $\mathbf{X}^\dagger = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ is the usual pseudoinverse. The following Lemma gives the exact values of $\hat{e}_2$ and $\sin\theta$ (thereby also implicitly giving $\hat{e}_1$ and $\cos\theta$).

Lemma 17. *Define* ϕ *and* v_2 *as in (37), we have*

$$(I - \mathbf{X}\mathbf{X}^\dagger)\mathbf{e} = -(v_2 \otimes v_2)(\|z\| \sin\phi)^2$$

and hence $\hat{e}_2 = v_2 \otimes v_2$ *and* $\sin\theta = (\|z\|\sin\phi)^2 / \|\mathbf{e}\|$.

Proof. We solve the projection problem

$$\begin{aligned} \|(I - \mathbf{X}\mathbf{X}^\dagger)\mathbf{e}\| = \min_y \|\mathbf{e} - \mathbf{X}y\| &= \|(xx^T - zz^T) - (xy^T + yx^T)\|_F \\ &= \min_{\alpha,\beta} \left\| \begin{bmatrix} \|x\|^2 - \|z\|^2 \cos^2\phi & -\|z\|^2 \sin\phi\cos\phi \\ -\|z\|^2 \sin\phi \cos\phi & -\|z\|^2 \sin^2\phi \end{bmatrix} - \begin{bmatrix} 2\alpha & \beta \\ \beta & 0 \end{bmatrix} \right\| \\ &= \|z\|^2 \sin^2\phi \end{aligned}$$

in which the second line makes a change of bases to v_1 and v_2. Clearly, the minimizer is in the direction of $-v_2 \otimes v_2$. $\square$

Using these properties of $\mathbf{X}$ and $\mathbf{e}$, we can now solve the following problem in closed-form.

Lemma 18. *Define $\alpha = (\|z\| \sin\phi)^2/\|e\| = \sin\theta$ and $\beta = \|x\|^2/\|e\|$. Then, the following optimization problem*

$$\begin{aligned} \psi(\gamma) \quad \equiv \quad & \underset{y,W}{\text{maximum}} && \mathbf{e}^T[\mathbf{X}y - \text{vec}\,(W)] \\ & \text{subject to} && \|\mathbf{e}\| \cdot \|\mathbf{X}y - \text{vec}\,(W)\| = 1 \\ &&& \text{tr}(\mathbf{X}W\mathbf{X}^T) = 2\beta \cdot \gamma \\ &&& W \succeq 0 \end{aligned}$$

is feasible if and only if $0 \le \gamma \le 1$ with optimal value

$$\psi(\gamma) = \gamma\alpha + \sqrt{1-\gamma^2}\sqrt{1-\alpha^2}.$$

Proof. The case of $\gamma < 0$ is infeasible as it would require $\text{tr}(W) < 0$ with $W \succeq 0$. For $\gamma \ge 0$, we begin by relaxing the norm constraint into an inequality, as in $\|\mathbf{e}\| \cdot \|\mathbf{X}y - \text{vec}\,(W)\| \le 1$. Solving the resulting convex optimization over y with a fixed W yields

$$y^\star = \mathbf{X}^\dagger[\tau \cdot \mathbf{e} + \text{vec}\,(W)], \qquad \|\mathbf{e}\|\|\mathbf{X}y^\star - \text{vec}\,(W)\| = 1, \qquad \tau \ge 0. \tag{39}$$

The problem is feasible if and only if $\|\mathbf{e}\|\|(I - \mathbf{X}\mathbf{X}^\dagger)\text{vec}\,(W)\| \le 1$. Whenever feasible, the relaxation is tight, and equality is attained. The remaining problem over W reads (after some rearranging):

$$\underset{W \succeq 0}{\text{minimize}} \quad \mathbf{e}^T(I - \mathbf{X}\mathbf{X}^\dagger)\text{vec}\,(W) \quad \text{subject to} \quad \langle \mathbf{X}^T\mathbf{X}, W\rangle = 2\beta \cdot \gamma$$

and this reduces to the following using Lemma 16 and Lemma 17:

$$\underset{W \succeq 0}{\text{minimize}} \quad -(\|e\| \sin\theta)\langle v_2 v_2^T, W\rangle \quad \text{subject to} \quad 2\|x\|^2\langle I + 2v_1v_1^T, W\rangle = 2\beta \cdot \gamma$$

with minimizer

$$\text{vec}\,(W^\star) = \frac{2\beta \cdot \gamma}{2\|x\|^2}(v_2 \otimes v_2) = \frac{\gamma}{\|e\|}\hat{e}_2.$$

Clearly, we have feasibility $\|e\|\|(I - \mathbf{X}\mathbf{X}^\dagger)\text{vec}\,(W^\star)\| \le 1$ if and only if $\gamma \le 1$. Substituting this particular $W^\star$ into (39) yields

$$\mathbf{X}y^\star - \text{vec}\,(W^\star) = \frac{\sqrt{1-\gamma^2}}{\|e\|}\hat{e}_1 + \frac{\gamma}{\|e\|}\hat{e}_2.$$

Substituting (38) yields

$$\mathbf{e}^T[\mathbf{X}y^\star - \text{vec}\,(W^\star)] = \sqrt{1-\gamma^2}\cos\theta + \gamma\sin\theta$$

as desired. $\square$

7.2 Proof of Theorem 13

We consider the condition number optimization problem from Zhang et al. [2018a]:

$$\eta(x,z) \quad \equiv \quad \max_{\eta,\mathbf{H}} \left\{\eta \quad : \quad \mathbf{X}^T\mathbf{H}\mathbf{e} = 0, \quad 2\mathrm{mat}(\mathbf{H}\mathbf{e}) + \mathbf{X}^T\mathbf{H}\mathbf{X} \succeq 0, \quad \eta I \preceq \mathbf{H} \preceq I\right\}. \tag{40}$$

Its optimal value satisfies the following identity with respect to our original LMI in (21):

$$\delta(x,z) = \mathrm{LMI}(x,z) = \frac{1-\eta(x,z)}{1+\eta(x,z)} = 1 - \frac{2}{1+1/\eta(x,z)}. \tag{41}$$

The latter equality shows that $\delta(x,z)$ is a decreasing function of $\eta(x,z)$. This allows us to lower-bound $\delta(x,z)$ by upper-bounding $\eta(x,z)$.

Next, we relax (40) to the following problem

$$\eta_{\mathrm{ub}}(x,z) \quad \equiv \quad \max_{\eta,\mathbf{H}} \left\{\eta \quad : \quad \mathbf{X}^T\mathbf{H}\mathbf{e} = 0, \quad 2\mathrm{mat}(\mathbf{H}\mathbf{e}) + \mathbf{X}^T\mathbf{X} \succeq 0, \quad \eta I \preceq \mathbf{H} \preceq I\right\}. \tag{42}$$

This yields an upper-bound $\eta_{\mathrm{ub}}(x,z)$ on $\eta(x,z)$ because $\mathbf{H} \succeq I$ implies $\mathbf{X}^T\mathbf{X} \succeq \mathbf{X}^T\mathbf{H}\mathbf{X}$. Problem (42) has Lagrangian dual (we write $v = \mathrm{vec}\,(V)$ to simplify notation)

$$\begin{aligned} \underset{y,U_1,U_2,V=\mathrm{mat}(v)}{\text{minimize}} \quad & \mathrm{tr}(U_2) + \langle \mathbf{X}^T\mathbf{X}, V\rangle \\ \text{subject to} \quad & (\mathbf{X}y - v)\mathbf{e}^T + \mathbf{e}(\mathbf{X}y-v)^T = U_1 - U_2 \\ & \mathrm{tr}(U_1) = 1, \quad U_1, U_2, V \succeq 0. \end{aligned} \tag{43}$$

The dual is strictly feasible (for sufficiently small ϵ, set $y=0$, $V=\epsilon I$, $U_1 = \eta I - \epsilon W$, and $U_2 = \eta \cdot I_{n^2} + \epsilon W$ with suitable η and W), so Slater's condition is satisfied, strong duality holds, and the objectives coincide. We will implicitly solve the primal problem (42) by solving the dual problem (43).

In the case that $x = 0$, problem (43) yields a trivial solution $y = 0$, $V = zz^T/(2\|z\|^4)$, $U_1 = (z\otimes z)(z\otimes z)^T/\|z\|^4$, and $U_2 = 0$ with objective value $\eta_{\mathrm{ub}}(0,z) = 0$.

In the case that $x \neq 0$, we define $\alpha = (\|z\|\sin\phi)/\|\mathbf{e}\|$ and $\beta = \|x\|^2/\|\mathbf{e}\| > 0$ and make a number of reductions on the dual problem (43). First, we use Lemma 14 to optimize over U_1 and U_2 and the length of y to yield

$$\underset{y,V=\mathrm{mat}(v)\succeq 0}{\text{minimize}} \quad \frac{\mathrm{tr}([M]_-) + \langle \mathbf{X}^T\mathbf{X}, V\rangle}{\mathrm{tr}([M]_+)} \quad \text{where} \quad M = (\mathbf{X}y - v)\mathbf{e}^T + \mathbf{e}(\mathbf{X}y-v)^T. \tag{44}$$

Here, we have divided the objective by the constraint $\mathrm{tr}([M]_+) = 1$ noting that the problem is homogenous over y and V. Substituting explicit expressions for the eigenvalues of M in Lemma 15 yields

$$\underset{y,V=\mathrm{mat}(v)\succeq 0}{\text{minimize}} \quad \frac{\langle \mathbf{X}^T\mathbf{X}, V\rangle + \|\mathbf{e}\|\|\mathbf{X}y - v\|(1-\cos\theta)}{\|\mathbf{e}\|\|\mathbf{X}y-v\|(1+\cos\theta)} \quad \text{where} \quad \cos\theta = \frac{\mathbf{e}^T(\mathbf{X}y-v)}{\|\mathbf{e}\|\|\mathbf{X}y-v\|}. \tag{45}$$

This is a multi-objective optimization over two competing trade-offs: minimizing $\langle \mathbf{X}^T\mathbf{X}, V\rangle$ and maximizing $\cos\theta$. To balance these two considerations, we parameterize over a fixed $\gamma =$

$\langle \mathbf{X}^T\mathbf{X}, V\rangle/2\beta$ and use Lemma 18 to maximize $\cos\theta$. The resulting univariate optimization reads

$$\eta_{\text{ub}}(x,z) \quad = \quad \min_{0\le\gamma\le\alpha} \Psi(\gamma) \equiv \frac{2\beta\cdot\gamma + [1-\psi(\gamma)]}{1+\psi(\gamma)} \tag{46}$$

where the function $\psi(\gamma) = \gamma\alpha + \sqrt{1-\gamma^2}\sqrt{1-\alpha^2}$ defined on Lemma 18 takes on the role of the best choice of $\cos\theta$. Here, one limit $\gamma = 0$ sets $\langle \mathbf{X}^T\mathbf{X}, V\rangle = 0$, while the other $\gamma = \alpha$ sets $\cos\theta = 1$. We cannot have $\gamma < 0$ because $V \succeq 0$. Any choice of $\gamma > \alpha$ will be strictly dominated by $\gamma = \alpha$, because $\gamma = \alpha$ already maximizes $\cos\theta$.

The univariate problem (46) is *quasiconvex*. This follows from the *concavity* of $\psi(\gamma)$ over this range:

$$\psi'(\gamma) = \alpha - \gamma\frac{\sqrt{1-\alpha^2}}{\sqrt{1-\gamma^2}}, \qquad \psi''(\gamma) = -\frac{\sqrt{1-\alpha^2}}{\sqrt{1-\gamma^2}} - \gamma^2\frac{\sqrt{1-\alpha^2}}{(1-\gamma^2)^{3/2}}.$$

Hence, the level sets of $\Psi(\gamma) \ge 0$ are convex:

$$\Psi(\gamma) \le c \quad \iff \quad 2\beta\cdot\gamma + (1-c) \le (1+c)\psi(\gamma).$$

We will proceed to solve the problem in closed-form and obtain

$$\Psi(\gamma^\star) = \min_{0\le\gamma\le\alpha}\Psi(\gamma) \quad = \quad \frac{1-\sqrt{1-\alpha^2}}{1+\sqrt{1-\alpha^2}} \qquad \text{if } \beta \ge \frac{\alpha}{1+\sqrt{1-\alpha^2}}, \tag{47}$$

$$\frac{\beta(\beta-\alpha)}{\beta\alpha-1} \qquad \text{if } \beta \le \frac{\alpha}{1+\sqrt{1-\alpha^2}}. \tag{48}$$

Substituting $\Psi(\gamma^\star) = \eta_{\text{ub}}(x,z)$ into $\eta_{\text{ub}}(x,z) \ge \eta(x,z)$ and using $\eta(x,z)$ to lower-bound $\delta(x,z)$ via (41) completes the proof of the lemma. (Note that setting $x = 0$ sets $\beta = 0$ and yields $\Psi(\gamma^\star) = \eta_{\text{ub}}(0,z) = 0$ as desired.)

First, we verify whether the optimal solution $\gamma^\star$ lies on the boundary of the search interval $[0,\alpha]$, that is $\gamma^\star \in \{0,\alpha\}$. Taking derivatives yields

$$\Psi'(\gamma) = \frac{[2\beta - \psi'(\gamma)](1+\psi(\gamma)) - \psi'(\gamma)[2\beta\cdot\gamma + 1 - \psi(\gamma)]}{(1+\psi(\gamma))^2}.$$

For $\gamma = 0$ to be a stationary point, we require $\Psi'(0) \ge 0$, and hence

$$[2\beta - \psi'(0)](1+\psi(0)) \ge \psi'(0)[2\beta\cdot 0 + 1 - \psi(0)]$$
$$\iff \qquad \beta \ge \frac{\alpha}{1+\sqrt{1-\alpha^2}}.$$

In this case, we have $\Psi(0) = (1-\sqrt{1-\alpha^2})/(1+\sqrt{1-\alpha^2})$, which is the expression in (47). The choice $\gamma = \alpha$ cannot be stationary, because $\Psi'(\alpha) \le 0$ would imply

$$[2\beta - \psi'(\alpha)](1+\psi(\alpha)) - \psi'(\alpha)[2\beta\cdot\alpha + 1 - \psi(\alpha)] \le 0,$$
$$\iff \qquad 2\beta(1+1) \le 0,$$

which is impossible as we have $\beta = \|x\|^2/\|e\| > 0$ by hypothesis.

Otherwise, the optimal solution $\gamma^\star$ lies in the interior of the search interval $[0, \alpha]$, that is $\gamma^\star \in (0, \alpha)$. In this case, we simply relax the bound constraints on γ and solve the unconstrained problem as a linear fractional conic program

$$\min_{|\gamma|\le 1} \Psi(\gamma) = \min_{\|\xi\|\le 1} \left\{ \frac{1 + (c-d)^T \xi}{1 + d^T \xi} \right\}$$

where

$$c = \begin{bmatrix} 2\beta \\ 0 \end{bmatrix}, \qquad d = \begin{bmatrix} \alpha \\ \sqrt{1-\alpha^2} \end{bmatrix}, \qquad \xi = \begin{bmatrix} \gamma \\ \sqrt{1-\gamma^2} \end{bmatrix}.$$

(Note that the relaxation $\|\xi\| \le 1$ is always tight, because the linear fractional objective is always monotonous with respect to scaling of ξ.) Defining $q = \xi/(1 + d^T\xi)$ and $q_0 = 1/(1 + d^T\xi) \ge 0$ rewrites this as the second-order cone program

$$\Psi(\gamma^\star) = \min_{\|q\|\le q_0} \left\{ \begin{bmatrix} c-d \\ 1 \end{bmatrix}^T \begin{bmatrix} q \\ q_0 \end{bmatrix} : \begin{bmatrix} d \\ 1 \end{bmatrix}^T \begin{bmatrix} q \\ q_0 \end{bmatrix} = 1 \right\}$$

that admits a strictly feasible point $q = 0$ and $q_0 = 1$. Accordingly, the Lagrangian dual has zero duality gap:

$$\Psi(\gamma^\star) = \max_\lambda \{\lambda : \|c - (1+\lambda)d\| \le (1-\lambda)\}.$$

If the maximum $\lambda^\star$ exists, then it must attain the inequality, as in

$$\|c - (1+\lambda^\star)d\|^2 = (1-\lambda^\star)^2.$$

We can simply solve this quadratic equation

$$c^T c - 2c^T d(1+\lambda^\star) + (1+\lambda^\star)^2 d^T d = 4 - 4(\lambda^\star + 1) + (\lambda^\star + 1)^2$$

for the optimal $\lambda^\star = \Psi(\gamma^\star)$. Noting that $d^T d = 1$, we actually have just a single root

$$1 + \lambda^\star = \frac{(c^T c - 4)}{2(c^T d - 2)} = \frac{\beta^2 - 1}{\alpha\beta - 1} = 1 + \frac{\beta(\beta - \alpha)}{\alpha\beta - 1},$$

and this yields the expression (48).

8 Numerical Results

An important advantage of our formulation is that $\delta(X, Z)$ can be evaluated numerically in cases where an exact closed-form solution does not exist (or is too difficult to obtain). In this section, we augment our analysis with a numerical study. In the rank $r = 1$ case, we exhaustively evaluate $\delta(x, z)$ over its two degrees of freedom to gain insight on its behavior, and also to quantify the conservatism of the lower-bound in Theorem 13. In the rank $r \ge 1$ case, we sample $\delta(X, Z)$ uniformly at random over X and Z, in order to understand its distribution and hypothesize on higher-rank versions of our recovery guarantees.

In our experiments, we implement Algorithm 1 in MATLAB. We parse the LMI problem using YALMIP [Lofberg, 2004] and solve it using MOSEK [Andersen and Andersen, 2000]. All algorithms parameters (e.g. accuracy, iterations, etc.) are left at their default values.

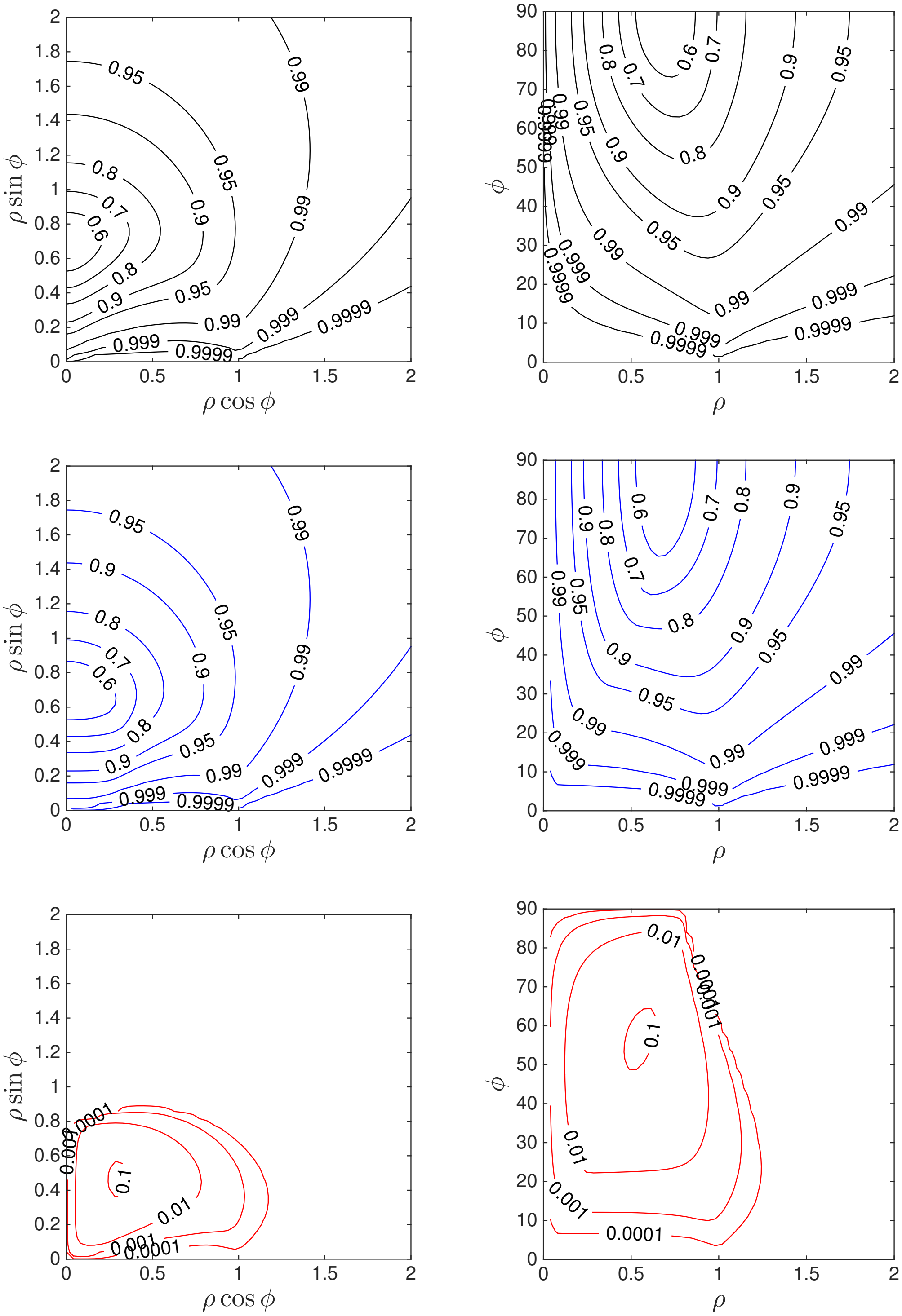

Figure 1: The function $\delta(x,z)$ and its lower bound $\delta_{\mathrm{lb}}(x,z)$ visualized with respect to the length ratio $\rho = \|x\|/\|z\|$ and the incidence angle $\phi = \arccos(x^T z/\|x\|\|z\|)$: **(top)** the function $\delta(x,z)$; **(middle)** the lower-bound $\delta_{\mathrm{lb}}(x,z)$; **(bottom)** the error $\delta(x,z) - \delta_{\mathrm{lb}}(x,z)$; **(left)** rectangular coordinates; **(right)** polar coordinates.

8.1 Visualizing $\delta(x,z)$ and $\delta_{\mathrm{lb}}(x,z)$ for rank $r=1$

Using a suitable orthogonal projector P, we can reduce the function $\delta(x,z)$ down to two underlying degrees of freedom: the length ratio $\rho = \|x\|/\|z\|$ and the incidence angle $\phi = \arccos(x^T z/\|x\|\|z\|)$. First, without loss of generality, we assume that the ground truth $M^\star = zz^T$ has unit norm $\|M^\star\|_F = 1$. (Otherwise, we can suitably rescale all arguments below.) Then, the following projector P satisfies $PP^T x = x$ and $PP^T z = z$ with

$$P = \begin{bmatrix} z & \frac{(I - zz^T)x}{\|(I-zz^T)x\|} \end{bmatrix}, \qquad P^T x = \begin{bmatrix} \rho\cos\phi \\ \rho\sin\phi \end{bmatrix}, \qquad P^T z = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{49}$$

Applying this particular P to Lemma 9 yields the following

$$\delta(x,z) = \delta(P^T x, P^T z) = \delta\left(\begin{bmatrix} \rho\cos\phi \\ \rho\sin\phi \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right). \tag{50}$$

In fact, this two-variable function is symmetric over its four rectangular quadrants

$$\delta\left(\begin{bmatrix} \rho\cos\phi \\ \rho\sin\phi \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \delta\left(\begin{bmatrix} \pm\rho\cos\phi \\ \pm\rho\sin\phi \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \tag{51}$$

because either $\pm z$ corresponds to the same ground truth, and because the second column of P can point in either $\pm(I - zz^T)x$.

Accordingly, we can use (50) and (51) to visualize $\delta(x,z)$ as a two-dimensional graph, either in rectangular coordinates over $(\rho\cos\phi, \rho\sin\phi) \in [0, \rho_{\max}]^2$, or in polar coordinates over $(\rho,\phi) \in [0,\rho_{\max}] \times [0, \pi/2]$. Moreover, we can plot our closed-form lower-bound $\delta_{\mathrm{lb}}(x,z)$ on the same axes, in order to quantify its conservatism $\delta(x,z) - \delta_{\mathrm{lb}}(x,z)$.

The top row of Figure 1 plots $\delta(x,z)$ in rectangular and polar coordinates. The plot shows $\delta(x,z)$ as a smooth function with a single basin at $\rho = 1/\sqrt{2}$ and $\phi = 90^\circ$. Outside of a narrow region with $1/2 \le \rho \le 1$ and $\phi \ge 45^\circ$, we have $\delta(x,z) \ge 0.9$. For smaller RIP constants, spurious local minima must appear in a narrow region—they cannot occur arbitrarily anywhere. Excluding this region—as in our local guarantee in Theorem 4—allows much larger RIP constants δ to be accommodated.

The middle and bottom rows of Figure 1 plots $\delta_{\mathrm{lb}}(x,z)$ and $\delta(x,z) - \delta_{\mathrm{lb}}(x,z)$ in rectangular and polar coordinates. The two functions match within 0.01 for either $\rho \ge 1$ or $\phi \le 30^\circ$, and fully concur in the asymptotic limits $\rho \to \{0, +\infty\}$ and $\phi \to \{0^\circ, 90^\circ\}$. The greatest error of around 0.1 occurs at $\rho \approx 0.5$ and $\phi \approx 55^\circ$. We conclude that $\delta_{\mathrm{lb}}(x,z)$ is a high quality approximation for $\delta(x,z)$.

8.2 Distribution of $\delta(X,Z)$ for rank $r \ge 1$

In the high-rank case, a simple characterization of $\delta(X,Z)$ is much more elusive. Given a fixed rank-r ground truth $M^\star$, let its corresponding eigendecomposition be written as $M^\star = V\Lambda V^T$ where $V \in \mathbb{R}^{n\times r}$ is orthogonal and Λ is diagonal. By setting $Z = V\Lambda^{1/2}$ and suitably selecting an orthogonal projector P, it is always possible to satisfy

$$\delta(X,Z) = \delta(P^T X, P^T Z) = \delta\left(\begin{bmatrix} \hat{X}_1 \\ \hat{X}_2 \end{bmatrix}, \begin{bmatrix} \Lambda^{1/2} \\ 0 \end{bmatrix} \right), \tag{52}$$

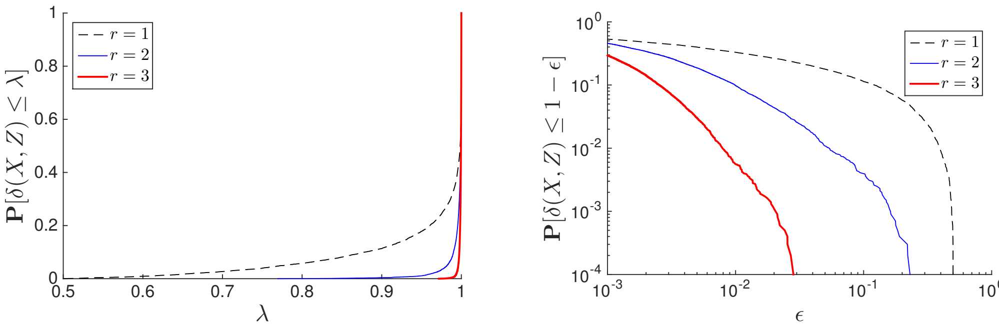

Figure 2: Empirical cumulative distribution of $\delta(X,Z)$ over $N = 10^4$ samples of $X, Z \in \mathbb{R}^{n\times r}$ where $X_{i,j}, Z_{i,i} \sim \text{Gaussian}(0,1)$: **(left)** linear plot of $\mathbf{P}[\delta(X,Z) \le \lambda]$ over $\lambda \in [1/2, 1]$; **(right)** logarithmic plot of $\mathbf{P}[\delta(X,Z) \le 1-\epsilon]$ over the tail $\epsilon \in [10^{-3}, 10^{0}]$.

where $\hat{X}_1, \hat{X}_2 \in \mathbb{R}^{r\times r}$. While (52) bares superficial similarities to (50), the equation now contains at least $2r^2 + r - 1$ degrees of freedom. Even $r = 2$ results in 9 degrees of freedom, which is too many to visualize.

Instead, we sample $\delta(X,Z)$ uniformly at random over its underlying degrees of freedom. Specifically, we select all elements in $X \in \mathbb{R}^{n\times r}$ and only the diagonal elements of $Z \in \mathbb{R}^{n\times r}$ independently and identically distributed from the standard Gaussian, as in $X_{i,j}, Z_{i,i} \sim \text{Gaussian}(0,1)$. We then use Algorithm 1 to evaluate $\delta(X,Z)$.

Figure 2 plots the empirical cumulative distributions for $r \in \{1,2,3\}$ from $N = 10^4$ samples. We see that each increase in rank r results in a sizable reduction in the distribution tail. The rank $r = 1$ trials yielded $\delta(x,z)$ arbitrarily close to the minimum value of $1/2$, but the rank $r = 2$ trials were only able to find $\delta(X,Z) \approx 0.8$. The rank $r = 3$ trials were even more closely concentrated about one, with the minimum at $\delta(X,Z) \approx 0.97$. These results suggest that higher rank problems are generically easier to solve, because larger RIP constants are sufficient to prevent the points from being spurious local minima. They also suggest that $\delta(X,Z) \ge 1/2$ over *all* rank $r \ge 1$, though this is not guaranteed, because "bad" choices of X, Z can always exist on a lower-dimensional zero-measure set.

9 Conclusions

The low-rank matrix recovery problem is known to contain *no spurious local minima* under a restricted isometry property (RIP) with a sufficiently small RIP constant δ. In this paper, we introduce a proof technique capable of establishing RIP thresholds that are both *necessary* and *sufficient* for exact recovery. Specifically, we define $\delta(X,Z)$ as the *smallest* RIP constant associated with a counterexample with *fixed* ground truth $M^\star = ZZ^T$ and *fixed* spurious point X, and define $\delta^\star = \min_{X,Z} \delta(X,Z)$ as the smallest RIP constant over all counterexamples. Then, δ-RIP low-rank matrix recovery contains no spurious local minima if and only if $\delta < \delta^\star$.

Our key insight is to show that $\delta(X,Z)$ has an *exact* convex reformulation. In the rank-1 case, the resulting problem is sufficiently simple that it can be relaxed and solved in closed-form. Using this closed-form bound, we prove that $\delta < 1/2$ is both necessary and sufficient for exact recovery from any arbitrary initial point. For larger RIP constants $\delta \geq 1/2$, we show that an initial point x_0 satisfying $f(x_0) \leq (1-\delta)^2 f(0)$ is enough to guarantee exact recovery using a descent algorithm.

Acknowledgements

We are grateful to Salar Fattahi for a meticulous reading and detailed comments, and to Salar Fattahi and Cédric Josz for fruitful discussions. This work was supported by the ONR Awards N00014-17-1-2933 and ONR N00014-18-1-2526, NSF Award 1808859, DARPA Award D16AP00002, and AFOSR Award FA9550- 17-1-0163.

A Detailed proof of Lemma 9

Given $X, Z \in \mathbb{R}^{n\times r}$, define $\mathbf{e} \in \mathbb{R}^{n^2}$ and $\mathbf{X} \in \mathbb{R}^{n^2 \times nr}$ to satisfy the following with respect to X and Z

$$\mathbf{e} = \mathrm{vec}\,(XX^T - ZZ^T), \qquad \mathbf{X}\mathrm{vec}\,(Y) = \mathrm{vec}\,(XY^T + YX^T) \qquad \forall Y \in \mathbb{R}^{n\times r}, \tag{53}$$

Let $P \in \mathbb{R}^{n\times d}$ with $d \leq n$ satisfy

$$P^T P = I_d, \qquad PP^T X = X, \qquad PP^T Z = Z$$

and define $\mathbf{P} = P \otimes P$ and the projections $\hat{X} = P^T X$ and $\hat{Z} = P^T Z$. Define $\hat{\mathbf{e}} \in \mathbb{R}^{d^2}$ and $\hat{\mathbf{X}} \in \mathbb{R}^{d\times dr}$ to satisfy (53) with X, Z replaced by $\hat{X}, \hat{Z}$.

Our goal is to show that

$$\delta = \hat{\delta}, \qquad \mathbf{H} = \mathbf{P}\hat{\mathbf{H}}\mathbf{P}^T + (I - \mathbf{P}\mathbf{P}^T),$$

satisfy the primal feasibility equations

$$\mathbf{X}^T \mathbf{H}\mathbf{e} = 0, \tag{54a}$$

$$2[I_r \otimes \mathrm{mat}(\mathbf{H}\mathbf{e})] + \mathbf{X}^T\mathbf{H}\mathbf{e} \succeq 0, \tag{54b}$$

$$(1-\delta)I \preceq \mathbf{H} \preceq (1+\delta)I, \tag{54c}$$

and that

$$y = (I_r \otimes P)\hat{y}, \qquad U_1 = \mathbf{P}\hat{U}_1\mathbf{P}^T, \qquad U_2 = \mathbf{P}\hat{U}_2\mathbf{P}^T, \qquad V = (I_r \otimes P)\hat{V}(I_r \otimes P)^T$$

satisfy the dual feasibility equations

$$\sum_{j=1}^{r} (\mathbf{X}y - \mathrm{vec}\,(V_{j,j}))\mathbf{e}^T + \mathbf{e}(\mathbf{X}y - \mathrm{vec}\,(V_{j,j}))^T - \mathbf{X}V\mathbf{X}^T = U_1 - U_2, \tag{55a}$$

$$\mathrm{tr}(U_1 + U_2) = 1, \tag{55b}$$

$$V \succeq 0, \quad U_1 \succeq 0, \quad U_2 \succeq 0, \tag{55c}$$

under the hypothesis that $(\hat{\delta}, \hat{\mathbf{H}})$ and $(\hat{y}, \hat{U}_1, \hat{U}_2, \hat{V})$ satisfy (54) and (55) with $\mathbf{e}, \mathbf{X}$ replaced by $\hat{\mathbf{e}}, \hat{\mathbf{X}}$.

We can immediately verify (54c), (55b), and (55c) using the orthogonality of $\mathbf{P}$. To verify the remaining equations, we will use the following identities.

Claim 19. We have

$$\mathbf{e} = \mathbf{P}\hat{\mathbf{e}}, \qquad \mathbf{X}(I_r \otimes P) = \mathbf{P}\hat{\mathbf{X}} \qquad \mathbf{P}^T\mathbf{X} = \hat{\mathbf{X}}(I_r \otimes P)^T.$$

Proof. For all $Y \in \mathbb{R}^{n\times r}$ and $\hat{Y} \in \mathbb{R}^{d\times r}$, we have

$$\begin{aligned}
\mathbf{e} &= \mathrm{vec}\,(XX^T - ZZ^T) = \mathrm{vec}\,[P(\hat{X}\hat{X}^T - \hat{Z}\hat{Z}^T)P^T] = (P\otimes P)\hat{\mathbf{e}},\\
\mathbf{X}(I_r\otimes P)\mathrm{vec}\,(\hat{Y}) &= \mathbf{X}\mathrm{vec}\,(P\hat{Y}) = \mathrm{vec}\,[P(\hat{X}\hat{Y}^T + \hat{Y}\hat{X}^T)P^T] = \mathbf{P}\hat{\mathbf{X}}\mathrm{vec}\,(\hat{Y}),\\
\mathbf{P}^T\mathbf{X}\mathrm{vec}\,(Y) &= \mathrm{vec}\,[(P^TX)(P^TY)^T + (P^TY)(P^TX)^T] = \hat{\mathbf{X}}\mathrm{vec}\,(P^TY) = \hat{\mathbf{X}}(I_r\otimes P)^T\mathrm{vec}\,(Y).
\end{aligned}$$

□

Now, we have (55a) from

$$\begin{aligned}
\mathbf{X}y - \mathrm{vec}\,(V_{j,j}) &= \mathbf{X}(I_r\otimes P)\hat{y} - \mathbf{P}\mathrm{vec}\,(\hat{V}_{j,j}) = \mathbf{P}(\hat{\mathbf{X}}\hat{y} - \mathrm{vec}\,(\hat{V}_{j,j})),\\
\mathbf{X}V\mathbf{X}^T &= \mathbf{X}(I_r\otimes P)\hat{V}(I_r\otimes P)^T\mathbf{X} = \mathbf{P}(\hat{\mathbf{X}}\hat{V}\hat{\mathbf{X}}^T)\mathbf{P}^T.
\end{aligned}$$

To prove (54a), we use

$$\mathbf{X}^T\mathbf{H}\mathbf{e} = \mathbf{X}^T\mathbf{H}(\mathbf{P}\hat{\mathbf{e}}) = \mathbf{X}^T(\mathbf{P}\hat{\mathbf{H}})\hat{\mathbf{e}} = (I_r\otimes P)\hat{\mathbf{X}}^T\hat{\mathbf{H}}\hat{\mathbf{e}}.$$

Lastly, to prove (54b), we define

$$\mathbf{S} = 2\cdot[I_r\otimes \mathrm{mat}(\mathbf{H}e)] + \mathbf{X}^T\mathbf{H}\mathbf{X}$$

and $P_\perp$ as the orthogonal complement of P. Then, observe that

$$I_r\otimes \mathrm{mat}(\mathbf{H}e) = I_r\otimes (P\,\mathrm{mat}(\hat{\mathbf{H}}\hat{e})\,P^T) = (I_r\otimes P)(I_r\otimes \mathrm{mat}(\hat{\mathbf{H}}\hat{e}))(I_r\otimes P)^T,$$

and that

$$\mathbf{X}^T\mathbf{H}\mathbf{X}(I_r\otimes P) = \mathbf{X}^T\mathbf{H}(\mathbf{P}\hat{\mathbf{X}}) = \mathbf{X}^T(\mathbf{P}\hat{\mathbf{H}})\hat{\mathbf{X}} = (I_r\otimes P)\hat{\mathbf{X}}^T\hat{\mathbf{H}}\hat{\mathbf{X}}.$$

Hence, we have

$$\begin{aligned}
(I_r\otimes P)^T\mathbf{S}(I_r\otimes P) &= 2\cdot[I_r\otimes\mathrm{mat}(\hat{\mathbf{H}}\hat{e})] + \hat{\mathbf{X}}\hat{\mathbf{H}}\hat{\mathbf{X}} \succeq 0,\\
(I_r\otimes P_\perp)^T\mathbf{S}(I_r\otimes P_\perp) &= (I_r\otimes P_\perp)^T\mathbf{X}^T\mathbf{H}\mathbf{X}(I_r\otimes P_\perp) \succeq 0,\\
(I_r\otimes P_\perp)^T\mathbf{S}(I_r\otimes P) &= 0,
\end{aligned}$$

and this shows that $\mathbf{S}\succeq 0$ as desired.

B Detailed proof of Lemma 10

Given $X, Z \in \mathbb{R}^{n\times r}$, let $P = \mathrm{orth}([X,Z])$ and $\mathbf{P} = P\otimes P$. Our goal is to show that

$$\delta = \hat{\delta}, \qquad \mathbf{H} = \mathbf{P}\hat{\mathbf{H}}\mathbf{P}^T$$

satisfy the primal feasibility equations

$$\mathbf{X}^T\mathbf{He} = 0, \tag{56a}$$

$$2[I_r \otimes \mathrm{mat}(\mathbf{He})] + \mathbf{X}^T\mathbf{He} \succeq 0, \tag{56b}$$

$$(1-\delta)I \preceq \mathbf{P}^T\mathbf{HP} \preceq (1+\delta)I, \tag{56c}$$

and that

$$y = (I_r \otimes P)\hat{y}, \qquad U_1 = \hat{U}_1, \qquad U_2 = \hat{U}_2, \qquad V = (I_r \otimes P)\hat{V}(I_r \otimes P)^T$$

satisfy the dual feasibility equations

$$\sum_{j=1}^{r}(\mathbf{X}y - \mathrm{vec}\,(V_{j,j}))\mathbf{e}^T + \mathbf{e}(\mathbf{X}y - \mathrm{vec}\,(V_{j,j}))^T - \mathbf{X}V\mathbf{X}^T = \mathbf{P}(U_1 - U_2)\mathbf{P}^T, \tag{57a}$$

$$\mathrm{tr}[\mathbf{P}(U_1 + U_2)\mathbf{P}^T] = 1, \tag{57b}$$

$$V \succeq 0, \quad U_1 \succeq 0, \quad U_2 \succeq 0, \tag{57c}$$

under the hypothesis that $(\hat{\delta}, \hat{\mathbf{H}})$ and $(\hat{y}, \hat{U}_1, \hat{U}_2, \hat{V})$ satisfy (54) and (55) with $\mathbf{e}, \mathbf{X}$ replaced by $\hat{\mathbf{e}}, \hat{\mathbf{X}}$. The exact steps for verifying (56) and (57) are identical to the proof of Lemma 9, and are omitted for brevity.

References

Yonatan Amit, Michael Fink, Nathan Srebro, and Shimon Ullman. Uncovering shared structures in multiclass classification. In *Proceedings of the 24th international conference on Machine learning*, pages 17–24. ACM, 2007.

Erling D Andersen and Knud D Andersen. The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm. In *High performance optimization*, pages 197–232. Springer, 2000.

Sohail Bahmani and Justin Romberg. Phase retrieval meets statistical learning theory: A flexible convex relaxation. In *Artificial Intelligence and Statistics*, pages 252–260, 2017.

Srinadh Bhojanapalli, Anastasios Kyrillidis, and Sujay Sanghavi. Dropping convexity for faster semi-definite optimization. In *Conference on Learning Theory*, pages 530–582, 2016a.

Srinadh Bhojanapalli, Behnam Neyshabur, and Nati Srebro. Global optimality of local search for low rank matrix recovery. In *Advances in Neural Information Processing Systems*, pages 3873–3881, 2016b.

Nicolas Boumal, P-A Absil, and Coralia Cartis. Global rates of convergence for nonconvex optimization on manifolds. *IMA Journal of Numerical Analysis*, page drx080, 2018. doi: 10.1093/imanum/drx080. URL `http://dx.doi.org/10.1093/imanum/drx080`.

T Tony Cai and Anru Zhang. Sharp RIP bound for sparse signal and low-rank matrix recovery. *Applied and Computational Harmonic Analysis*, 35(1):74–93, 2013.

Emmanuel J Candes and Yaniv Plan. Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements. *IEEE Transactions on Information Theory*, 57(4):2342–2359, 2011.

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

Emmanuel J Candès and Terence Tao. The power of convex relaxation: Near-optimal matrix completion. *IEEE Transactions on Information Theory*, 56(5):2053–2080, 2010.

Emmanuel J Candes, Thomas Strohmer, and Vladislav Voroninski. Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66(8):1241–1274, 2013.

Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval via wirtinger flow: Theory and algorithms. *IEEE Transactions on Information Theory*, 61(4):1985–2007, 2015.

Coralia Cartis, Nicholas IM Gould, and Ph L Toint. Complexity bounds for second-order optimality in unconstrained optimization. *Journal of Complexity*, 28(1):93–108, 2012.

Tony F Chan. Rank revealing qr factorizations. *Linear algebra and its applications*, 88: 67–82, 1987.

Pengwen Chen, Albert Fannjiang, and Gi-Ren Liu. Phase retrieval with one or two diffraction patterns by alternating projections with the null initialization. *Journal of Fourier Analysis and Applications*, 24(3):719–758, 2018.

Yuxin Chen and Emmanuel Candes. Solving random quadratic systems of equations is nearly as easy as solving linear systems. In *Advances in Neural Information Processing Systems*, pages 739–747, 2015.

Andrew R Conn, Nicholas IM Gould, and Ph L Toint. *Trust region methods*, volume 1. Siam, 2000.

Simon S Du, Chi Jin, Jason D Lee, Michael I Jordan, Aarti Singh, and Barnabas Poczos. Gradient descent can take exponential time to escape saddle points. In *Advances in Neural Information Processing Systems*, pages 1067–1077, 2017.

Rong Ge, Furong Huang, Chi Jin, and Yang Yuan. Escaping from saddle points–online stochastic gradient for tensor decomposition. In *Conference on Learning Theory*, pages 797–842, 2015.

Rong Ge, Jason D Lee, and Tengyu Ma. Matrix completion has no spurious local minimum. In *Advances in Neural Information Processing Systems*, pages 2973–2981, 2016.

Rong Ge, Chi Jin, and Yi Zheng. No spurious local minima in nonconvex low rank problems: A unified geometric analysis. In *International Conference on Machine Learning*, pages 1233–1242, 2017.

Tom Goldstein and Christoph Studer. Phasemax: Convex phase retrieval via basis pursuit. *IEEE Transactions on Information Theory*, 2018.

Gene H Golub and Charles F Van Loan. *Matrix computations*. JHU Press, 3 edition, 1996.

Prateek Jain, Praneeth Netrapalli, and Sujay Sanghavi. Low-rank matrix completion using alternating minimization. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 665–674. ACM, 2013.

Chi Jin, Rong Ge, Praneeth Netrapalli, Sham M Kakade, and Michael I Jordan. How to escape saddle points efficiently. In *International Conference on Machine Learning*, pages 1724–1732, 2017.

Raghunandan H Keshavan, Andrea Montanari, and Sewoong Oh. Matrix completion from a few entries. *IEEE Transactions on Information Theory*, 56(6):2980–2998, 2010a.

Raghunandan H Keshavan, Andrea Montanari, and Sewoong Oh. Matrix completion from noisy entries. *Journal of Machine Learning Research*, 11(Jul):2057–2078, 2010b.

Jason D Lee, Max Simchowitz, Michael I Jordan, and Benjamin Recht. Gradient descent only converges to minimizers. In *Conference on Learning Theory*, pages 1246–1257, 2016.

Johan Lofberg. Yalmip: A toolbox for modeling and optimization in matlab. In *Computer Aided Control Systems Design, 2004 IEEE International Symposium on*, pages 284–289. IEEE, 2004.

Marco Mondelli and Andrea Montanari. Fundamental limits of weak recovery with applications to phase retrieval. *Foundations of Computational Mathematics*, pages 1–71, 2018.

Katta G Murty and Santosh N Kabadi. Some np-complete problems in quadratic and nonlinear programming. *Mathematical programming*, 39(2):117–129, 1987.

Yurii Nesterov and Boris T Polyak. Cubic regularization of newton method and its global performance. *Mathematical Programming*, 108(1):177–205, 2006.

Praneeth Netrapalli, Prateek Jain, and Sujay Sanghavi. Phase retrieval using alternating minimization. In *Advances in Neural Information Processing Systems*, pages 2796–2804, 2013.

Dohyung Park, Anastasios Kyrillidis, Constantine Carmanis, and Sujay Sanghavi. Non-square matrix sensing without spurious local minima via the Burer-Monteiro approach. In *Artificial Intelligence and Statistics*, pages 65–74, 2017.

Dohyung Park, Anastasios Kyrillidis, Constantine Caramanis, and Sujay Sanghavi. Finding low-rank solutions via nonconvex matrix factorization, efficiently and provably. *SIAM Journal on Imaging Sciences*, 11(4):2165–2204, 2018.

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

Jasson DM Rennie and Nathan Srebro. Fast maximum margin matrix factorization for collaborative prediction. In *Proceedings of the 22nd international conference on Machine learning*, pages 713–719. ACM, 2005.

Sujay Sanghavi, Rachel Ward, and Chris D White. The local convexity of solving systems of quadratic equations. *Results in Mathematics*, 71(3-4):569–608, 2017.

Ruoyu Sun and Zhi-Quan Luo. Guaranteed matrix completion via non-convex factorization. *IEEE Transactions on Information Theory*, 62(11):6535–6579, 2016.

Andreas M Tillmann and Marc E Pfetsch. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *IEEE Transactions on Information Theory*, 60(2):1248–1259, 2014.

Gang Wang, Georgios B Giannakis, and Yonina C Eldar. Solving systems of random quadratic equations via truncated amplitude flow. *IEEE Transactions on Information Theory*, 64(2):773–794, 2018.

HuiMin Wang and Song Li. The bounds of restricted isometry constants for low rank matrices recovery. *Science China Mathematics*, 56(6):1117–1127, 2013.

Richard Y Zhang, Cédric Josz, Somayeh Sojoudi, and Javad Lavaei. How much restricted isometry is needed in nonconvex matrix recovery? In *Advances in Neural Information Processing Systems*, 2018a. arXiv:1805.10251.

Richard Y Zhang, Javad Lavaei, and Ross Baldick. Spurious critical points in power system state estimation. In *Hawaii International Conference on System Sciences (HICSS)*, 2018b.

Tuo Zhao, Zhaoran Wang, and Han Liu. A nonconvex optimization framework for low rank matrix estimation. In *Advances in Neural Information Processing Systems*, pages 559–567, 2015.

Qinqing Zheng and John Lafferty. A convergent gradient descent algorithm for rank minimization and semidefinite programming from random linear measurements. In *Advances in Neural Information Processing Systems*, pages 109–117, 2015.

Zhihui Zhu, Qiuwei Li, Gongguo Tang, and Michael B Wakin. Global optimality in low-rank matrix optimization. *IEEE Transactions on Signal Processing*, 66(13):3614–3628, 2018.